Pre-emptive aptamer design using SARS-CoV-2 evolution to inform epitope selection

Lead Group: Nick Goldman (EMBL-EBI)
Partner Group: Gian Gaetano Tartaglia (IIT Genova)
External Collaborator: Michele Vendruscolo (University of Cambridge)

Abstract: We will create bioinformatics resources permitting fast and pre-emptive design of aptamers for SARS-CoV-2 proteins to facilitate vaccine design and other therapeutic needs in the global response to COVID-19. The methods developed will have applications with other RNA viruses.

Rationale & Hypothesis: The overall goal of this project is to develop a strategy to enable a rapid therapeutic response to the SARS-CoV-2 mutations that are emerging as a consequence of the extensive spreading of the virus in the human population. We will use the evolutionary history of the virus to predict mutations that are likely to appear in the coming months, in particular those that may enable the virus to escape binding to therapeutic aptamers. We will thus identify specific positions on the virus spikes that cannot easily be mutated for functional constraints and use them as epitopes for the rational design of RNA aptamers. This strategy will enable the generation of panels of pre-emptive aptamers to target future virus variants, which will be ready for immediate use. The techniques developed will be applicable to aptamer design for other RNA viruses.

Aims: *Phase 1 - Goldman Group, EMBL-EBI.* The objective of this phase is to use SARS-CoV-2 sequence information from the EBI COVID-19 Data Portal and GISAID to find site- and nucleotide-specific mutation probabilities. This will be based on phylogenetic analyses that make allowance for the evolutionary relationships of sampled viruses, biased sampling globally etc., topics in which the Goldman Group is expert [1,2]. This process may benefit from automation facilitating periodic (daily, weekly) updates based on new information uploaded to relevant databases. Based upon the mutation probabilities, we will select sites of low mutational frequency (prioritizing the spike protein, but also other candidates of SARS-CoV-2 proteome would be considered) that may serve as viable epitopes for aptamer design. We will identify multiple candidate sites with the understanding that during phase 2 some of these sites may prove to be inviable. Furthermore, using the inferred mutation probabilities and the past and extant variation observed at the selected epitopes, we will predict which new variants of these epitopes are more likely to emerge in the future, which will help us inform a long-term effective aptamer design strategy. *Phase 2 - Tartaglia Group, IIT Genova.* The goal of this phase will be to design RNA aptamers that bind to these epitopes [3,4]. The Tartaglia Group has pioneered the development of methods for the rational design of RNA-protein interactions and a number of RNAs targeting specific proteins have been previously tested in the lab [3,4]. As a parallel strategy, in collaboration with the group of Prof. Michele Vendruscolo at the Centre for Misfolding Diseases in Cambridge, we will also use the same epitopes to design antibodies using the rational *in silico* approach that he has recently developed for this purpose [5]. Depending on the successful candidate's abilities, experimental verifications could be performed by them or by others in the Tartaglia Group.

Significance & Impact: The world-wide diffusion of COVID-19 is enabling SARS-CoV-2 to evolve rapidly. The emergence of variants is expected to compromise the efficacy of the current vaccines. It is therefore important to develop techniques to respond rapidly to the new variants. The pre-emptive strategy that we propose will enable the generation of panels of aptamers to target future virus variants, in principle even before they appear, thus shortening the timescale of the response. Aptamers would be valuable therapeutic tools as they are non-immunogenic, thermally stable and soluble.This project could have a significant impact in permitting pharmaceutical companies to respond faster to changing needs in global strategies for both active vaccination (Phase 1) and passive vaccination (Phase 2).

Integration of Expertise of Partners: The Tartaglia Group (IIT) has developed a ground-breaking technique for the rational design of RNA-protein interactions. The Goldman Group (EMBL-EBI) provides advanced DNA and protein sequence phylogenetic analyses that are essential for this project for the identification of viral conserved epitopes and prediction of future mutations. The successful candidate will gain experience in both of these areas, combining those skills in this novel project

Person Specification: We expect the ETPOD fellow to be based primarily at IIT Genova. Some time would be spent working with the Goldman Group at the EMBL-European Bioinformatics Institute, Cambridge. Precise details will depend on the fellow's existing skills and training needs.

References:

[1] N. De Maio et al. Mutation rates and selection on synonymous mutations in SARS-CoV-2. Genome Biol. Evol. 13:evab087 (2021).

[2] H. S. Vohringer et al. Genomic reconstruction of the SARS-CoV-2 epidemic across England from September 2020 to May 2021. medRxiv https://www.medrxiv.org/content/10.1101/2021.05.22.21257633v2 (2021).

[3] A. Armaos et al. *cat*RAPID *omics v2.0*: going deeper and wider in the prediction of protein-RNA interactions. Nucl. Acids Res. https://doi.org/10.1093/nar/gkab393 (2021).

[4] A. Vandelli et al. Structural analysis of SARS-CoV-2 genome and predictions of the human interactome. Nucl. Acids Res. 48:11270-11283 (2020).

[5] F. A. Aprile et al. Rational design of a conformation-specific antibody for the quantification of Aβ oligomers. Proc. Natl. Acad. Sci. USA 117:13509-13518 (2020).