Artificial Intelligence-Based Design of Antibody-like Engineered Protein Scaffolds

André R. F. Salgueiro [1,2], Pedro Moreira [1,3], Cláudio M. Soares [1], Leonardo Vanescchi [2] and Diana Lousa [1]

1. Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal
2. NOVA Information Management School, Universidade Nova de Lisboa, Lisboa, Portugal
3. Centro de Engenharia Biológica, Universidade do Minho, Braga, Portugal

Throughout humanity history, viral outbreaks caused devastating epidemics, like the recent COVID-19, caused by the SARS-CoV-2 virus originating in China. As of December 13, 2023, over 772 million confirmed COVID-19 cases, with more than 6.9 million deaths, have been reported globally to the World Health Organization. Over time, SARS-CoV-2 has evolved into different strains with varied characteristics, including immune system evasion and increased infectivity. This highlights the unpredictable nature of future pandemics, needing the development of new methodologies to combat a wide range of viruses. Therapeutic monoclonal antibodies, with their adaptability, show potential in targeting a broad range of viruses. However, their intricate and costly development processes hinder widespread use. Engineered protein scaffolds, such as monobodies1–3, which are smaller and simpler than monoclonal antibodies, offer a promising alternative. Their expression in bacteria makes their production and development processes simple and cost-effective, presenting exciting prospects for innovative treatments in cancer, infectious diseases, and autoimmune disorders. Although these engineered protein scaffolds have the potential to revolutionize medicine, we need efficient strategies to enable the design of molecules with tailor-made properties. The combination of machine learning methods like ProtGPT24, RFdiffusion5 and MSA Transformer6 and molecular dynamics (MD) simulations can be a powerful strategy to address this problem. This work aims to create a computational framework integrating machine learning techniques and MD simulations to streamline the development of engineered protein scaffolds that combine a high affinity for the target with optimal developability properties (including efficient production in bacteria and high physical and chemical stability). As proof of concept, we are focusing on designing protein scaffolds, including monobodies and helical miniproteins, to neutralize SARS-CoV-2. This approach holds promise for the development of biopharmaceuticals that can be adapted to a broad range of pathogenic agents, contributing to the ongoing battle against infectious diseases.