BIO_20 - Evolutionary algorithms are capable of *in silico* antibody optimization: a software for protein engineering using Genetic Algorithms

João Sartori[1]; Eduardo Krempser[2]; Ana Carolina Ramos Guimarães[1]; Lucas de Almeida Machado[3].
[1]Laboratory for Applied Genomics and Bioinnovations – Oswaldo Cruz Institute
[2]Institutional Platform for Biodiversity and Wildlife Health Oswaldo Cruz Foundation - Fiocruz
[3]National Institute of Women Children and Adolescents Health Fernandes Figueira/Fiocruz

Introduction: Monoclonal antibodies (mAbs) are crucial for therapy and diagnosis due to their antigen specificity. They serve as a basis for cost-effective antibody fragments like single chain fragment variables (scFv), ideal for penetrating challenging tissues like tumors. To fine-tune mAb-antigen affinity, adjustments are occasionally needed either boosting it for enhanced effects or - in the case of CAR T cell development - reducing it to minimize off-tumor effects. Achieving these modifications usually involves mutating complementarity- determining regions (CDR). However, navigating the vast mutational space is resource-intensive. This way, evolution based methods such as a Genetic Algorithm (GA) could help explore this large sequence space more efficiently.

Objectives: Here, we aim at enhancing the affinity of a scFv to CD19 using a GA through an iterative evolution of its CDRs.

Methodology: We started with the cryo-EM structure (PDB code: 7urv) of FMC63 scFv with CD19. Then, we initiate the GA testing 3 initial populations: i. scFvs with random mutations on the CDRs, ii. the initial scFv + random + Rosetta Design (RD) on the CDRs, iii. initial scFv + Rosetta Design on the CDRs. The algorithm initializes by evaluating scFv ΔG_{bind} to CD19 using pyRosetta's ref2015_cart score function. Tournament selection then samples the population and picks individuals with the highest scores, subjecting them to recombination and random mutations at a 2.5% rate. This process repeats until the new population reaches the desired size, restarting the cycle.

Results: Preliminary results indicate GA capacity to optimize the scFv, even when starting from high ΔG_{bind} fragments. For the random scFv population, the highest scFv ΔG_{bind} went from 1640.85 Rosetta Energy Units (REU) to 486.92 REU, while the lowest achieved -26.26 REU. Meanwhile, population mixed with random + RD the worst scFv goes from 1654.63 REU to -49.70 REU, and the best scFv -60.85 REU. For pure Rosetta Designs, the algorithm was still not able to optimize it, maintaining a ΔG_{bind} around -50 REU.

Conclusion: Here, we show the potential of a GA in optimizing a scFv binding to its antigen, through an iterative evolution of its CDRs, comparing different initial populations.

Keywords: Genetic Algorithm; scFv; Affinity optimization