

Nucleic acid (NA)-protein interactions in the cell constitute vital roles in various biological processes such as protein synthesis, gene expression, RNA processing, viral replication, cellular defense, and developmental regulation. The exploration of this interaction in therapeutic studies is primarily limited due to the unavailability of a scoring function that can understand the different components in the macromolecular complex such as protein, DNA, RNA, and ligands in a uniform representation. The goal of this project is to apply coarse-grained modeling to NA-protein-ligand complexes, first-of-such-kind effort and examine the preferential interactions at a macroscopic level. The innovative component of this project is to train this interaction potential in a deep-learning framework named SimNPL which constitutes the potential to score the NA-protein interacting ligands and prioritizes the better scoring molecules in an experimental setup. The SimNPL program will be tested on select complexes: HIV-1 reverse transcriptase-DNA complex (potent HIV-1 drug), SARS-CoV-2 with nsp13 (potent SARS-CoV-1 drug), DNA helicases (potent cancer drug). The project envisions immediate and long-lasting impact in rapidly promoting the therapeutic landscape of NA-targeting small molecules with valuable implications in drug discovery programmes.