**Modelling Somatic Hypermutations with Reinforcement Learning for PD1 and Pembrolizumab**

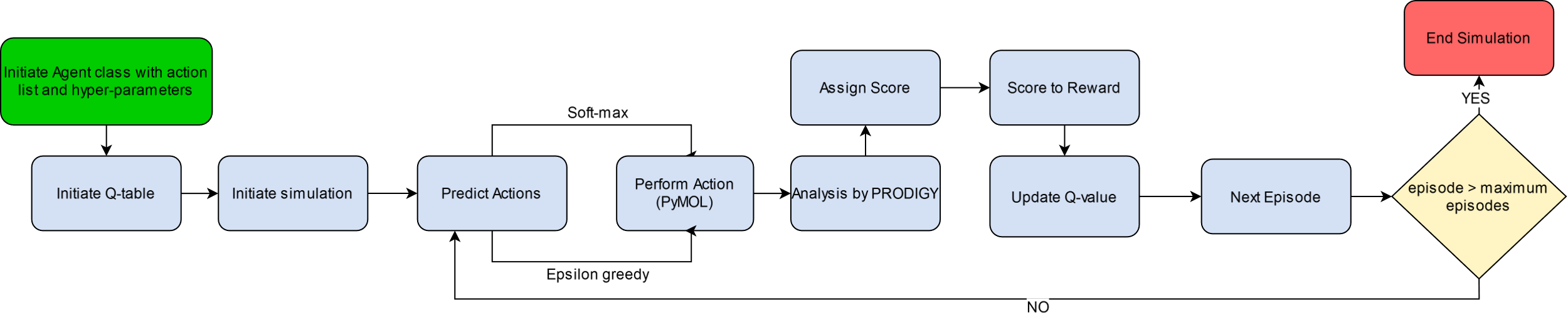
**Anmol Singh,** Subhashini Srinivasan, Nithya Ramakrishnan

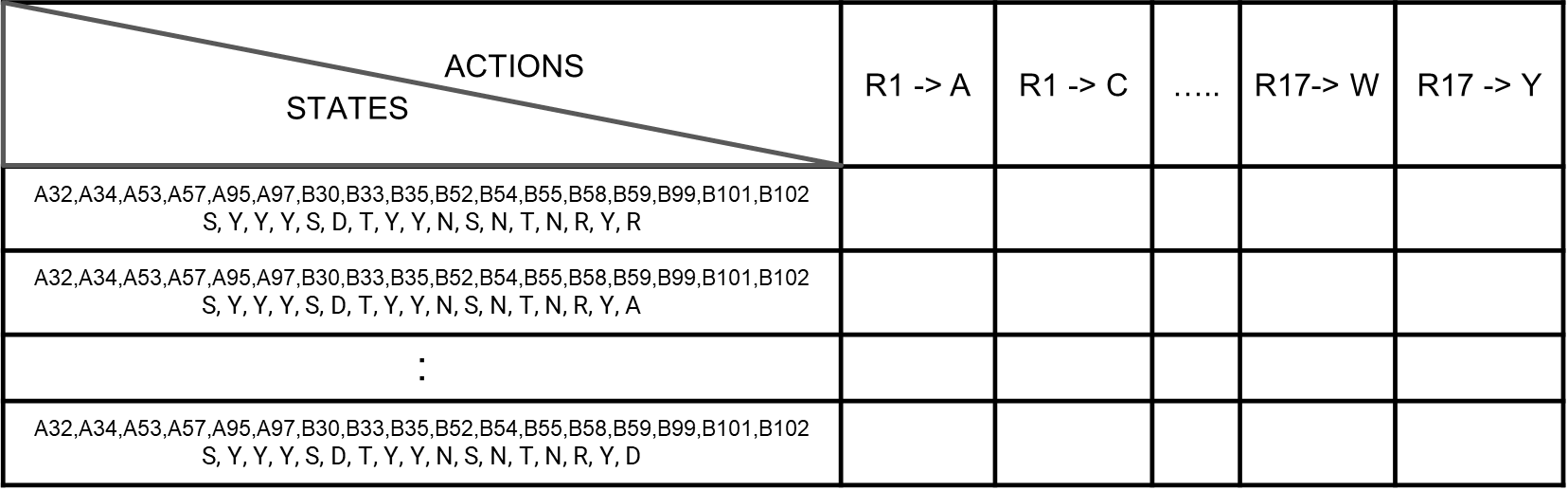
Abstract

The process of engineering antibodies with high affinity towards an antigen is very expensive and time-consuming in traditional wet approaches. Bioinformatic Tools and computational approaches can be used for this antibody design, to reduce the time and money required. Reinforcement learning is similar to natural selection, as in by trial-and-error method, the better actions (mutations) remain in the population and those harmful are removed over time.

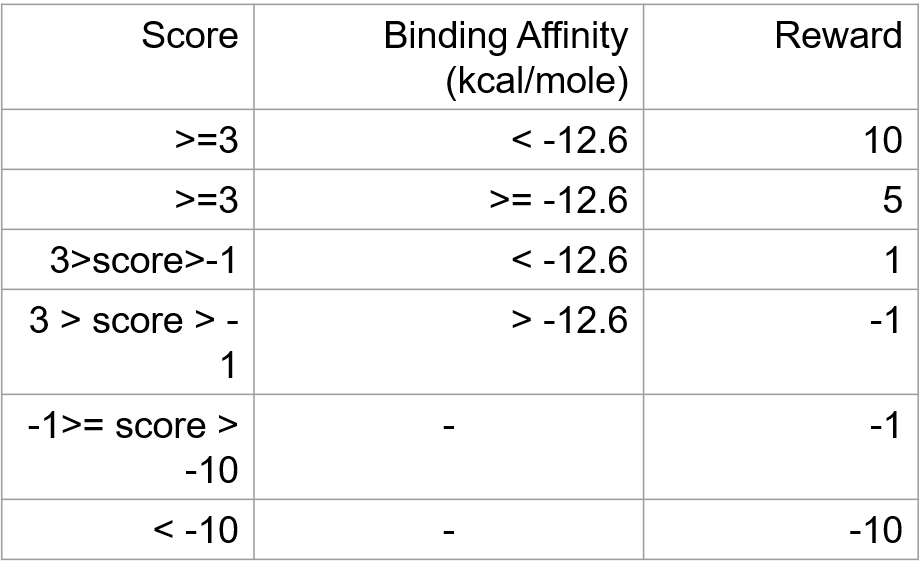
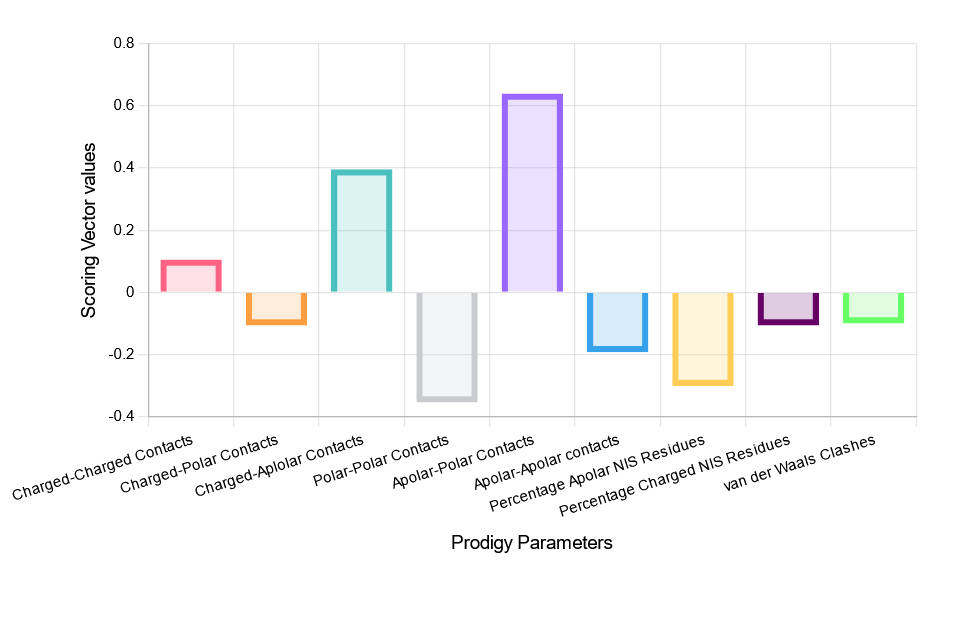
We present a reinforcement learning (RL) model for Somatic Hypermutations (SHM), which mimics the natural selection process in a very short time scale. In this model, the agent can learn to preferentially mutate amino acids in the antibody, leading to affinity maturation. The model thus predicts a higher binding affinity antibody than the initial antibody-antigen complex. We have used the Pembro-PD1 (5b8c) complex to create the model, as Pembro (Pembro) is widely used in immunotherapy.

We use Q-Learning in RL to model SHM on a reduced state space to provide better binding affinity antibodies. We validated the structure of the antibodies predicted by the RL model using AlphaFold2 and inter-residue distance plots to check for proper folding of chains and protein-protein interactions. This study provides a proof of concept that RL can be used for modelling the biological process of SHM and can further be employed for creating novel antibodies.

A

B

C b Table to reward function

C a

**Figure.** **A** Flowchart of Basic Q-learning, **B** Representation of Q-Table, **C** **a** Scoring vector **b** Score to reward function