The optimization of targeted B cell receptor-antigen binding: a Reinforcement Learning model

Technical Document

1.State of the Art:

The diverse repertoire of B-cells obtained from VDJ recombination and Junctional Diversity has been studied quite widely [1,2]. Since computational tools are faster than experimental profiling, several structural and probabilistic bioinformatic techniques have been developed in the past to analyze the antibody diversity. Homology modelling [3] and protein-protein docking [4] are some of the structural computational methods employed in the past. Tools like EnsembleDock [5] and SnugDock [6] used structural conformations and docking of the antibody-antigen to determine the best-fit antibody model. There have also been computational sequence based predictors that rely on the physio-chemical properties of the proteins such as hydrophobicity distribution [7]. Techniques based on the charge-balance of the residues of the antibodies have also been used in pharmacokinetics [8].

Computational analysis of the Complementarity Determining Regions (CDRs) of the antibodies has been in practice for design and modelling of antibodies. PARATOME is a web-based tool for predicting the antigen binding sites based on the CDRs of the antibodies [9-11]. Tools such as OptCDR [12] , AbDesign [13], OptMaven [14] are some other tools that combine docking and binding energy evaluation for antibody design and prediction.

Since the VDJ recombinations are probabilistic, Hidden Markov Models (HMM) were employed to model them. Some of the HMM implementations include iHMMune [15] and SoDA [16] where the hidden nodes represented omitted nucleotides or gene segments and these were combined with different transition probabilities. Recently, another HMM compiler based implementation “partis” was released and it includes certain parameters based on more detailed patterns [17]. The authors in [17] compare the previous techniques with “partis” using Hamming distance measures between the naive and inferred sequences and report that their technique outperforms the previous methods. In [18], an Information theory based maximum entropy model is employed to predict that there is higher order correlation between the repertoire of antibodies in zebrafish.

There have been several studies for modelling and design of antibodies based on VDJ combinations and structural binding. However, there have not been any theoretic modelling study on the Somatic Hypermutations (SHM) and affinity maturation processes which lead to further improvements in binding with the antigen. We propose a model using Reinforcement Learning (RL) where the antibodies learn to conform to the ideal structure based on feedback from the binding site in order to bind with minimum energy loss with the antigen . In [19,20], the authors explore adaptive immunity using RL briefly with patterns of antigens and antibodies but do not delve into the bio-physical processes underlying them. We combine ideas from some of the above studies such as Probabilistic Decoding and energy minimization and apply RL algorithms suitably for a critical understanding of the process.

2. Origin of the proposal:

The process by which B cells (and T cells) learn to recognize the antigen epitopes, bind to it and perform targeted killing of the antigen cells is a complex series of operations and is a stellar example of a “biological learning system”. As part of VDJ Recombination, enzymes such as RAG1/RAG2 bring together different combinations of V, D and J gene segments (bringing about a combinatorial diversity (in the order of 10^{13}) among the naive BCR (B cell Receptors) [22]. The BCR combination that can bind to the antigen with even partial affinity is activated. In this process in which one of the probabilistic VDJ combinations binds to the target antigen is modelled as an Information theoretic/Communication problem: When a received signal is a corrupted version of an input signal (antigen), decoding algorithms are employed to decode the correct sent signal (the matching VDJ combination antibody). The Hamming distance between the patterns of the antigen and the BCR combination will aid in the detection process. Monte Carlo

simulations that increase the probability of detection of the correct VDJ combination using the Hamming distance measure will also be employed in our work. Our experience with our previous work involving Information theory based modeling and decoding of chromatin inheritance [23] will be beneficial during the proposed work.

The activated B cell (with a given VDJ combination) further undergoes Somatic Hypermutation (SHM) with the help of AID and UNG enzymes, in which the B cell clones itself, and the clones mutate repeatedly with a signal from the binding site, until an optimized binding with the antigen is attained [24]. It is known that the rate of mutations in the SHM phase is very high (1 in 1000 bases per generation) and that certain motifs around the V region are preferentially targeted [25]. We therefore model the SHM and affinity maturation as a RL algorithm (using Q-Learning). The feedback signal in the RL algorithm would be a factor of the Gibbs free energy released by the non-covalent binding of the antigen-BCR and this would guide the BCR to optimize the mutations for the best possible fit with the antigen. The different parameters and processes associated with the SHM such as the enzymes AID, UNG, MMR (Mismatch repair enzyme), AP-endonuclease, Isotope switching and Non homologous end joining will also be taken into consideration in the model[26]. Our previous work on the parametrized path optimization of the ant food forage using Reinforcement Learning would serve as a preliminary RL based study[27] for this work.

We would be validating our work based on publicly available Covid-19 human data [28,29,30] using the antibody conformations at different timepoints. Collaborators will also experimentally verify the model using Next-gen-sequencing and/or BCR -sequencing on B cells from the mice immunized with Covid19 vaccines.
As part of future work, we would study how SHM is linked to memory B cells in the context of vaccines as an extension to RL. This would help in prediction by ML and Deep Learning as to how for a given individual, the immune system would optimize its response to a specific pathogen in a given time.

Collaborators:

Dr. Shashi Gujar (the Immunology expert from Dalhousie University, Canada),
Dr. Sanjay Chandrasekharan (Reinforcement Learning expert from Homi Bhabha Center for Science Education, TIFR, Mumbai)
Dr. Tripti Bameta, ACTREC - Tata Memorial Hospital, Mumbai (experimentalist)

Research Plan:

The main aim of the proposed work is to model the best fit binding of the BCR to the antigen. This happens during the process of Somatic Hypermutations and affinity maturation and this is modeled as an RL system. To narrow down the activated BCR pattern from the large repertoire of VDJ combinations, we propose to decode the BCR combination selected through Information Theory decoding algorithms/ Monte Carlo simulations as the initial part.

My previous work on Computational Biology in various genetic, epigenetic datasets and Information theoretic modeling of epigenetic inheritance in chromatin and recent work on Reinforcement Learning on ant-foraging optimisation would collectively enable me to scale the challenge of the proposed study successfully. My collaborators **Dr. Shashi Gujar (the Immunology expert from Dalhousie University, Canada), Dr. Sanjay Chandrasekharan (Reinforcement Learning expert from Homi Bhabha Center for Science Education, TIFR, Mumbai) and Dr. Tripti Bameta, ACTREC - Tata Memorial Hospital, Mumbai (experimentalist)** will provide additional support during the course of the development and validation of the project. The model will be implemented in R/Python using scientific libraries on a multi-core processor at IBAB, Bengaluru, with the help of a Research Fellows, hired for this project under my supervision. As a team, we are uniquely placed to implement and validate the proposed model.

For the first part of choosing the activating VDJ combination, as can be seen from Fig 1, we use Communication Decoding algorithms such as Viterbi decoding in combination with Monte Carlo simulations. The antigen and BCR structural conformations would be represented using a suitable code such as octal or hexadecimal in order to accommodate the millions of VDJ combinations. Since the simulations may be computationally intensive, we would use a multi-core CPU at IBAB and or a cloud service. The validation would be performed using the publicly available Covid-19 databases [28,29,30] for checking the antibody repertoire during recovery time. At the end of the first year, we would have completed this part of the project and the manuscript submitted to a bioinformatic journal.

The inputs from this part of the work would serve as the initial template for the Reinforcement Learning algorithm (implemented using Q-Learning) as shown in Fig 2. The agent (activated BCR) would clone, and the clones (becoming multiple agents) mutate and try to bind to the antigen with more affinity. They receive the feedback from the binding site in the form of Gibbs Free energy :

The Gibbs free energy change (ΔG) associated with the binding process is given by the equation [30]:

$$\Delta G = \Delta H - T\Delta S \qquad \text{(Eq 1)}$$

where ΔH (enthalpy) represents the energy released or absorbed during the formation of non-covalent interaction between the antigen and the BCR clone. Temperature (T) is kept constant in our simulations , so the change in the entropy (ΔS) determines the effectiveness of the binding. The reward function for the RL system can therefore be modelled as a negative exponential of ΔG.

The proposed RL model would provide the best fitting antibody BCR clone based on energy minimization principles. Since SHMs go on for a longer period of time after the infection, we can validate our results using the publicly available Covid-19 vaccination databases with the antibody repertoires at different time points post recovery [31] . We would also experimentally validate this using BCR-sequencing or RNA-seq (by amplifying the BCRs) with the equipment in ACTREC or Dalhousie University, Canada. By the end of the second year, we propose to complete this work and prepare the manuscript for peer-reviewed journals.

As part of future work, we aim to use Machine Learning algorithms such as Random Forests , SVM and/ or Deep Neural Networks to predict if for a given individual, a specific antigen pattern would invoke the required immune response in time for recovery. The research plan timeline is captured in fig 3. We would also aim to build capacities and train next-generation reseachers on ML algorithms for Immunology and Cancer therapy in India as part of this project.

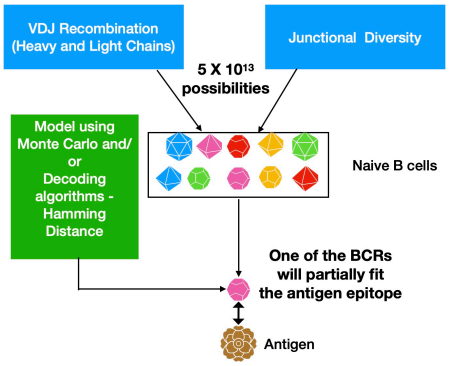

Fig 1: Modeling of the selection of the BCR that fits to the antigen after VDJ recombination and Junctional Diversity. This is modelled as an Information theoretic problem and can be solved using Decoding algorithms and Monte Carlo simulations with Hamming distance as the measure of error.

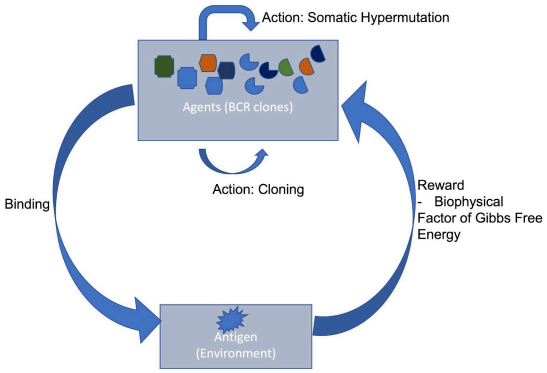

Fig 2: Modeling of the Somatic Hypermutations and selection of the best fitting antibody as a Reinforcement Learning algorithm. The agents are the BCR clones that compete for best-fit with the antigen

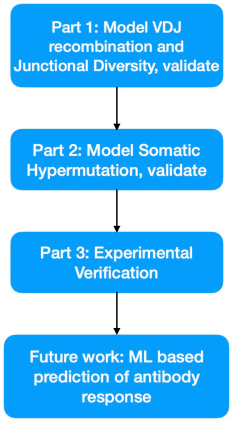

Fig 3: Research plan and timeline

4. Key publications of the Investigator during the last 5 years

1. Chandroth, Aamir Sahil, **Ramakrishnan, Nithya** and Chandrasekharan, Sanjay. “The self-organization of selfishness: Reinforcement Learning shows how selfish behavior can emerge from agent-environment interaction dynamics”, arxiv.org, (2023)

2. **Nithya Ramakrishnan**, SRB Pillai, R. Padinhateeri, “High fidelity epigenetic inheritance: Information theoretic model predicts threshold filling of histone modifications post replication”, PLoS Comput Biol 18(2): e1009861 (2022).

3. **Nithya Ramakrishnan** and R. Bose, “Analysis of healthy and tumor DNA methylation distributions in kidney-renal-clear-cell-carcinoma using Kullback–Leibler and Jensen–Shannon distance measures”, IET Systems Biology, vol. 11, no. 3, pp. 99-104, (2017)

4. **Nithya Ramakrishnan**, Mayuri Rege, Dibyendu Das, Sibiraj B.Pillai and Ranjith P., “Computational Analysis of Histone Post-translational Modification Pairs and their Influence on Genes”, EMBO Conference on Histone Chaperones, Crete, Greece, Oct 2019.

5. Bibliography:

1. Tonegawa, Susumu. "Somatic generation of antibody diversity." *Nature* 302.5909 (1983): 575-581.

2. Bassing, Craig H., Wojciech Swat, and Frederick W. Alt. "The mechanism and regulation of chromosomal V (D) J recombination." *Cell* 109.2 (2002): S45-S55.

3. Fiser, András, and Andrej Šali. "Modeller: generation and refinement of homology-based protein structure models." *Methods in enzymology*. Vol. 374. Academic Press, 2003. 461-491.

4. Brenke, Ryan, et al. "Application of asymmetric statistical potentials to antibody–protein docking." *Bioinformatics* 28.20 (2012): 2608-2614.

5. Chaudhury, Sidhartha, and Jeffrey J. Gray. "Conformer selection and induced fit in flexible backbone protein–protein docking using computational and NMR ensembles." *Journal of molecular biology* 381.4 (2008): 1068-1087.

6. Sircar, Aroop, and Jeffrey J. Gray. "SnugDock: paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models." *PloS computational biology* 6.1 (2010): e1000644.

7. Gallet, Xavier, et al. "A fast method to predict protein interaction sites from sequences." *Journal of molecular biology* 302.4 (2000): 917-926.

8. Datta-Mannan, Amita, et al. "Balancing charge in the complementarity-determining regions of humanized mAbs without affecting pI reduces non-specific binding and improves the pharmacokinetics." *MAbs*. Vol. 7. No. 3. Taylor & Francis, 2015.

9. Kunik, Vered, Shaul Ashkenazi, and Yanay Ofran. "Paratome: an online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure." *Nucleic acids research* 40.W1 (2012): W521-W524.

10. Ofran, Yanay, Avner Schlessinger, and Burkhard Rost. "Automated identification of complementarity determining regions (CDRs) reveals peculiar characteristics of CDRs and B cell epitopes." *The Journal of Immunology* 181.9 (2008): 6230-6235.

11. Kunik, Vered, Bjoern Peters, and Yanay Ofran. "Structural consensus among antibodies defines the antigen binding site." *PLoS computational biology* 8.2 (2012): e1002388.

12. Pantazes, R. J., and Costas D. Maranas. "OptCDR: a general computational method for the design of antibody complementarity determining regions for targeted epitope binding." *Protein Engineering, Design & Selection* 23.11 (2010): 849-858.

13. Lapidoth, Gideon D., et al. "Abdesign: A n algorithm for combinatorial backbone design guided by natural conformations and sequences." *Proteins: Structure, Function, and Bioinformatics* 83.8 (2015): 1385-1406.

14. Li, Tong, Robert J. Pantazes, and Costas D. Maranas. "OptMAVEn–a new framework for the de novo design of antibody variable region models targeting specific antigen epitopes." *PloS one* 9.8 (2014): e105954.

15. Gaëta, Bruno A., et al. "iHMMune-align: hidden Markov model-based alignment and identification of germline genes in rearranged immunoglobulin gene sequences." *Bioinformatics* 23.13 (2007): 1580-1587.

16. Volpe, Joseph M., Lindsay G. Cowell, and Thomas B. Kepler. "SoDA: implementation of a 3D alignment algorithm for inference of antigen receptor recombinations." *Bioinformatics* 22.4 (2006): 438-444.

17. Ralph, Duncan K., and Frederick A. Matsen IV. "Consistency of VDJ rearrangement and substitution parameters enables accurate B cell receptor sequence annotation." *PLoS computational biology* 12.1 (2016): e1004409.

18. Mora, Thierry, et al. "Maximum entropy models for antibody diversity." *Proceedings of the National Academy of Sciences* 107.12 (2010): 5405-5410.

19. Kato, Takuya, and Tetsuya J. Kobayashi. "Understanding adaptive immune system as reinforcement learning." *Physical Review Research* 3.1 (2021): 013222.

20. Ito, Jungo, et al. "Adaptive immunity based reinforcement learning." *Artificial Life and Robotics* 13 (2008): 188-193.

21. McKean, David, et al. "Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin." *Proceedings of the National Academy of Sciences* 81.10 (1984): 3180-3184.

22. Weigert, Martin G., et al. "Variability in the lambda light chain sequences of mouse antibody." *Nature* 228 (1970): 1045-1047.

23. Ramakrishnan, Nithya, Sibi Raj B. Pillai, and Ranjith Padinhateeri. "High fidelity epigenetic inheritance: Information theoretic model predicts threshold filling of histone modifications post replication." *PLOS Computational Biology* 18.2 (2022): e1009861.

24. O'brien, R. L., R. L. Brinster, and U. Storb. "Somatic hypermutation of an immunoglobulin transgene in K transgenic mice." *Nature* 326.6111 (1987): 405-409.

25. Rogozin, Igor B., and Nikolai A. Kolchanov. "Somatic hypermutagenesis in immunoglobulin genes: II. Influence of neighbouring base sequences on mutagenesis." *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1171.1 (1992): 11-18.

26. Li, Ziqiang, et al. "The generation of antibody diversity through somatic hypermutation and class switch recombination." *Genes & development* 18.1 (2004): 1-11.

27. Chandroth, Aamir Sahil, Ramakrishnan, Nithya and Chandrasekharan, Sanjay. “The self-organization of selfishness: Reinforcement Learning shows how selfish behavior can emerge from agent-environment interaction dynamics”, arxiv.org, 2023.

28. Schultheiß, Christoph, et al. "Next-generation sequencing of T and B cell receptor repertoires from COVID-19 patients showed signatures associated with severity of disease." *Immunity* 53.2 (2020): 442-455.

29. Galson, Jacob D., et al. "Deep sequencing of B cell receptor repertoires from COVID-19 patients reveals strong convergent immune signatures." *Frontiers in immunology* 11 (2020): 605170.

30. Gowthaman, Ragul, et al. "CoV3D: a database of high resolution coronavirus protein structures." *Nucleic acids research* 49.D1 (2021): D282-D287.

31. Phillips, Rob, et al. Physical biology of the cell. Garland Science, 2012.

32. Levin, Einav G., et al. "Waning immune humoral response to BNT162b2 Covid-19 vaccine over 6 months." *New England Journal of Medicine* 385.24 (2021): e84.

6. Equipment available with the Institute/ Group/ Department/Other Institutes for the project

Equipment Available	Generic name of the Equipment	Model,Make and year of purchase	Remarks including accessories available and current usage of equipment
IBAB Computing facility	High Performance Computing server with storage	IBM iDataplex HPC Server with IBM V3700 100TB Storage	Node Cluster for MPI based applications + 2 Master Compute Nodes
IBAB Computing facility	High Performance Computing server with storage	HP 14 Node HPC Cluster with Apollo 2000 100TB Storage	128 GB DDR IV for each node
IBAB Computing facility	NGS analysis server	Tyan Big Data Analysis Server	4 CPU, 8 cores per CPU,