

# **B cell multi-epitope senovaccine for healthy aging and tackling age associated pathologies**

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE AWARD OF THE DEGREE  
OF

MASTER OF SCIENCE  
IN  
**BIOTECHNOLOGY**

Submitted by:

**MOHD FARDEEN HUSAIN SHAHANSHAH**

**2K21/MSCBIO/27**

Under the supervision of

DR. ASMITA DAS



**DEPARTMENT OF BIOTECHNOLOGY**  
**DELHI TECHNOLOGICAL UNIVERSITY**  
(Formerly Delhi College of Engineering)  
Bawana Road, Delhi-110042

MAY, 2023

DELHI TECHNOLOGICAL UNIVERSITY  
(Formerly Delhi College of Engineering)  
Bawana Road, Delhi-110042

**CANDIDATE'S DECLARATION**

I, Mohd Fardeen Husain Shahanshah, 2K21/MSCBIO/27 of M.Sc. Biotechnology, hereby declare that the project Dissertation titled "B cell multiepitope senovaccine for healthy aging and tackling age associated pathologies" which is submitted by me to the Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Science, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi

**MOHD FARDEEN HUSAIN SHAHANSHAH**

Date: 30.05.2023

DEPARTMENT OF BIOTECHNOLOGY  
DELHI TECHNOLOGICAL UNIVERSITY  
(Formerly Delhi College of Engineering)  
Bawana Road, Delhi-110042

**CERTIFICATE**

I hereby certify that the Project Dissertation titled "B cell multiepitope senovaccine for healthy aging and tackling age associated pathologies" which is submitted by Mohd Fardeen Husain Shahanshah, 2K21/MSCBIO/27, Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Science, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

**DR. ASMITA DAS**

**SUPERVISOR**

Department of Biotechnology  
Delhi Technological University

**PROF.PRAVIR KUMAR**

**HEAD OF THE DEPARTMENT**

Department of Biotechnology  
Delhi Technological University

## **ACKNOWLEDGEMENT**

Countless people have supported my effort for this thesis, from close and from afar. I would like to thank my supervisor, Dr. Asmita Das, for her support and guidance. Their expertise and constructive feedback have been invaluable to me and have played an instrumental role in the success of this thesis. I have benefited greatly from your meticulous prompting towards thinking through the details of this project. I am also grateful to Delhi Technological University for providing me with the opportunity and technical knowledge base to conduct this novel research.

Finally, I would like to express my sincerest gratitude to my family and friends for their unconditional, unequivocal and unwavering support throughout the completion of my master's program. Their love, motivation, and belief in my abilities have been my constant source of strength.

## ABSTRACT

Age-associated illnesses are a consequence of accumulating senescent cells within the body. These non-proliferative derivatives of normal cells evade cytotoxic immune clearance and supplement disease pathogenesis and aging. Since most senolytic drugs show short-lived, off-targeted effects with high toxicity, a search for a relatively safer and highly specific modality is warranted. A preemptive approach to stall the pathognomonic signs of aging can be achieved through prophylactics called senovaccines, that trigger the immune system to specifically target and eliminate the senescent cells. Characteristic cellular markers of senescent cells such as urokinase plasminogen activator receptor (uPAR) and glycoprotein nonmetastatic melanoma protein B (GPNMB) can be used as promising senoantigens for fabricating senovaccines. In this research, a novel B cell multiepitope senovaccine has been proposed that can potentially elicit a long-lasting humoral immune response. Five highly antigenic B-cell epitopes were predicted and combined with a built-in adjuvant beta-defensin using suitable linkers. The senovaccine construct fulfilled the criteria of nonallergenicity, nontoxicity, solubility and stability. Molecular docking and simulation analysis revealed that the senovaccine construct can form productive and stable complexes with the variable region of anti-uPAR antibody. The computationally designed B-cell multiepitope senovaccine provides us with a novel plausible model that can be explored further for the development of efficacious senovaccines that support healthy aging.

# CONTENTS

<b>Candidate's Declaration</b>	ii
<b>Certificate</b>	iii
<b>Acknowledgement</b>	iv
<b>Abstract</b>	v
<b>Contents</b>	vi
<b>List of Figures</b>	ix
<b>List of Tables</b>	x
<b>List of Symbols, abbreviations</b>	xi
<b>CHAPTER 1 INTRODUCTION</b>	1
1.1 Aging	1
1.2 Cellular Senescence	2
1.3 Senolytics as an Anti-aging intervention	3
1.4 Theme of the research	4
<b>CHAPTER 2 LITERATURE REVIEW</b>	5
2.1 Age related pathologies	5
2.1.2 Atherosclerosis	5
2.2.2 Chronic kidney disease	5
2.2.3 Cancer	6
2.2.4 Neurodegenerative diseases	6

2.2 Senoantigens	7
2.2.1 uPAR	7
2.2.2 GPNMB	7
2.2.3 Rationale for selection	7
2.3 Research Hypothesis	8
<b>CHAPTER 3 METHODOLOGY</b>	9
3.1 Insilico vaccine designing	9
3.2 Research pipeline	10
3.2.1 Sequence retrieval and domain identification	10
3.2.2 Linear B-cell epitope prediction	10
3.2.3 Evaluation of predicted linear B-cell epitopes	10
3.2.4 Visualization of the linear B-cell epitopes	11
3.2.5 Construction of a B-cell multiepitope senovaccine and determination of its features	11
3.2.6 Determination of physicochemical properties	11
3.2.7 Secondary structure prediction	11
3.2.8 Tertiary structure prediction	12
3.2.9 Structural refinement and validation of the senovaccine	12
3.2.10 Molecular Docking of vaccine construct on uPAR antibody	12
3.2.11 Molecular dynamics simulation	12
3.2.12. Vaccine optimization and insilico cloning	13
<b>CHAPTER 4 RESULTS</b>	14

4.1. Prediction and screening of linear B -cell epitopes	14
4.2. Vaccine designing and feature prediction of the construct	17
4.3. Physicochemical analysis of the senovaccine construct.	17
4.4. Protein structure prediction and validation of the vaccine construct.	18
4.5. Molecular docking of the senovaccine on anti-uPAR antibody	22
4.6 Molecular Dynamics of the senovaccine and anti-uPAR antibody complex	24
4.7 Codon adaptation and in silico cloning	26
<b>CHAPTER 5 DISCUSSION</b>	27
<b>CHAPTER 6 CONCLUSION</b>	31
<b>Appendices</b>	32
Appendix 1: Linear B-cell epitope prediction tools	32
Appendix 2: Results from the epitope prediction tool for the uPAR antigen	33
Appendix 3: Results from the epitope prediction tool for the GPNMB antigen	34
Appendix 4: Predicted epitopes for the uPAR antigen	35
Appendix 5: Predicted epitopes for the GPNMB antigen	43
Appendix 6: Results of Galaxy Refine	56
<b>List of Publications</b>	57
<b>References</b>	58



## LIST OF FIGURES

- Figure 1.1 A cell displaying the typical hallmarks of senescence.
- Figure 4.1 Visualization of the most antigenic linear B-cell epitopes
- Figure 4.2 A graphical representation of the 347 AA long B-cell multiepitope senovaccine.
- Figure 4.3 Secondary structure prediction of the senovaccine construct
- Figure 4.4 Tertiary structure and validation of the refined protein structure of the senovaccine construct
- Figure 4.5 Molecular docking of senovaccine on anti-uPAR antibody AT-658.
- Figure 4.6 Protein-protein interactions between senovaccine construct and anti-uPAR antibody AT-658.
- Figure 4.7 Results from the NMA molecular dynamics simulation conducted on iMODs.
- Figure 4.8 In silico cloning map of the B cell multiepitope senovaccine sequence inserted into the pET28a(+) vector.

## LIST OF TABLES

Table 4.1	Predicted B cell epitopes of uPAR and GPNMB
Table 4.2	Features of the B-cell multiepitope vaccine construct
Table 4.3	Experimentally determined and predicted CDRs of anti-uPAR antibody ATN-658

## LIST OF SYMBOLS, ABBREVIATIONS, ACRONYMS

SC	Senescent Cell
SASP	Senescence-associated secretory phenotypes
uPAR	Urokinase-type plasminogen activator receptor
GPNMB	Glycoprotein Nonmetastatic Melanoma Protein B
DPP4	Dipeptidyl peptidase 4
ER	Endoplasmic reticulum
PDL	Programmed death ligand
CAR	Chimeric antigen receptor
IEDB	Immune Epitope Database and Analysis resource
MHC	Major Histocompatibility Complex
SA-beta-gal	Senescence associated beta-galactosidase
MALP-2	Macrophage activating lipopeptide
AA	Amino acids
Ldlr <sup>-/-</sup>	low-density lipoprotein receptor-deficient

# CHAPTER 1: INTRODUCTION

## 1.1 Aging

Aging is regarded as a nonlinear biological process which is typically accompanied with crippling comorbidities such as cancer, chronic kidney disease, diabetes, Alzheimer's and atherosclerosis, that diminish the quality of life and survivability of an individual [1-4]. It is presumed to emerge from accumulated cellular and genetic damage that manifests as a state of gradual decline in overall fitness along with increased susceptibility to illnesses that may ultimately result in death [5].

Even though aging in itself remains a bit of inconclusive mystery, it has been speculated that this pathophysiology can be a consequence of an array of intrinsic as well as extrinsic intermediaries. Common hallmarks that are typically correlated with the organismal aging process are: cellular senescence, genomic instability, epigenetic alterations, telomere attrition, loss of proteostasis, mitochondrial dysfunction, dysregulated nutrient-sensing, stem cell exhaustion, and altered intercellular communication [6].

As people can now live longer because of improvements in healthcare and medical technology, an unprecedented demographic shift is being observed wherein the number and proportion of people aged 60 years and above in the population is growing. WHO reports that by 2030, 1 out of every 6 individuals on the planet will be aged 60 years or over. Following the current trend, experts predict that by 2050 these figures are anticipated to double to 2.1 billion [5].

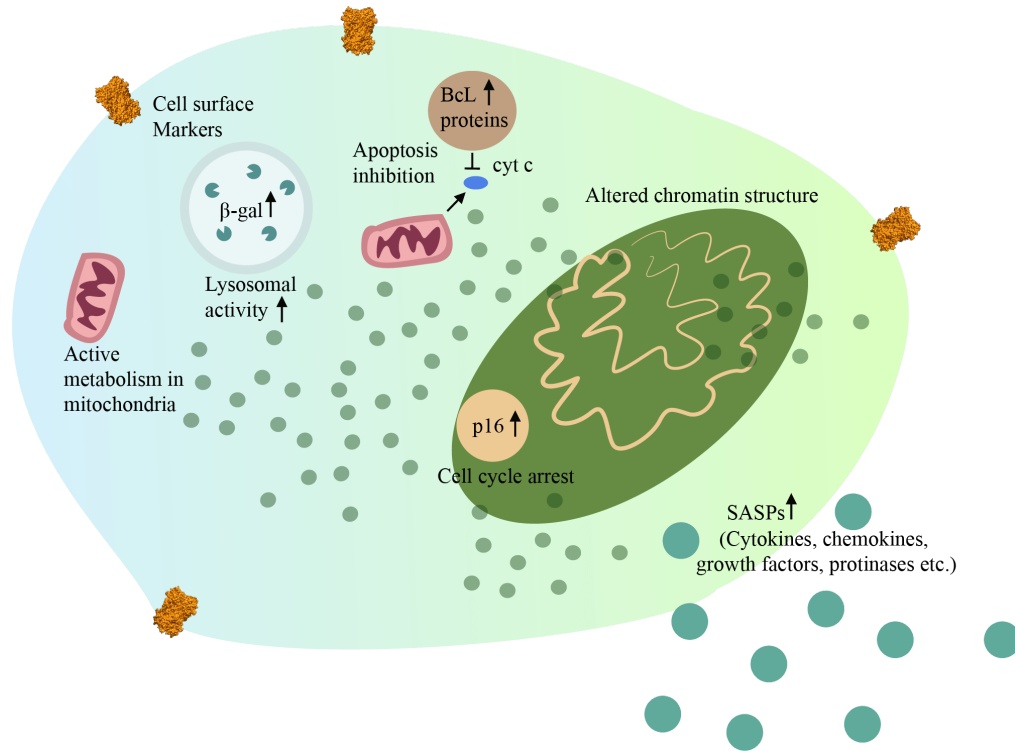
Generally, aging presents both opportunities and challenges. While it will drive up the demand for long-term primary healthcare and exacerbate the requirement for more age-appropriate environments, it will also allow aging individuals to continue serving as productive members of society and enrich our communities. However the latter will only be possible if aging individuals remain disease free.

## **1.2 Cellular Senescence**

A key catalyst that is supposed to facilitate this geroconversion is cellular senescence [4,5]. Senescence is a type of proliferation arrest that cells adopt in response to stressful stimuli like telomere shortening, nutritional disruptions, oxidative damage, endoplasmic reticulum stress, and genotoxic stress. The state of dormancy is characterized with the overexpression of cellular markers such as p16, senescence-associated-beta-gal, urokinase-type plasminogen activator receptor (uPAR), glycoprotein nonmetastatic melanoma protein B (GPNMB) and immunosuppressive ligands like programmed death ligand-1 (PDL-1) and nonclassical major histocompatibility complexes (MHC) along with secretion of effector molecules known as senescence-associated secretory phenotypes (SASPs) [2-4].

Phenotypic adaptations such as altered chromatin patterns, cytoskeleton remodeling, increased cellular size and granularity, upregulation of lysosomal enzymes and a metabolic shift to glycolysis from fatty acid catabolism are also observed among senescent cells [2,7,8]. Figure 1.1 briefly summarizes and illustrates the biology of a typical senescent cell.

In most cases these static cells can be easily identified through universal markers like SA  $\beta$ -gal, CDK4/6 inhibitor p16<sup>INK4a</sup>/p16, uPAR and dipeptidyl peptidase 4 (DPP4/CD26) [2]. However, evidence also suggests that SCs may display hysteresis wherein there are marked variations and strong phenotypic heterogeneity among the transcriptional and secretory profiles (SASPs) of SCs, based on their anatomical location or mode of senescence induction [2,7].



**Figure 1.1: A cell displaying the typical hallmarks of senescence.** The altered chromatin structure results in upregulation of p16 and BcL proteins that trigger a state of dormancy and apoptosis inhibition respectively. Elevated levels of beta-gal enzyme increase lysosomal activity. The senescent cell actively secretes molecules called SASPs that supplement its effector functions. There is also overexpression of senescent cell specific markers on its surface. Despite all these changes, the senescent cell retains its metabolism in the mitochondria as shown in the figure.

### 1.3 Senolytics as an Anti-aging intervention

Since the breakthrough revelation of this unique cellular process by Hayflick and Moorhead in 1961 that challenged the established archetype of endless cell division, our comprehension of cellular senescence has evolved drastically [2,7,8]. Senescence is generally believed to be an irreversible process that typically occurs in normal cells. Recent investigations, however, have revealed that even tumor cells can experience senescence when exposed to suitable stimuli, such as cancer therapies.

Additionally, numerous studies have been reported that reaffirm the role of senescent cells in progression of age associated ailments. Reversal of the aging phenotype, tumor cessation and chronic disease suppression has been observed in Phase 1 clinical studies, animal models and cell lines that were treated with senolytic agents [1-4].

Senolytic therapy involving Dasatinib and Quercetin (D + Q) caused substantial reduction in the levels of SCs in mice suffering from age-associated maladies like osteoporosis and frailty [7]. In the context of cancer, D+Q has been effective at eliminating radiation-induced skin ulceration and pulmonary fibrosis [9].

#### **1.4 Theme of the research**

To ensure that our future geriatric population ages healthily and is not reliant on long-term assisted care, interventions can be devised that either reverse the signs of aging or slow down the rate of aging. The cascade leading on to aging can be interrupted and adjusted. A growing body of scientific work has proved that it's possible to regulate the speed of aging and sustain vitality with the help of senolytics. Targeting SCs to impede aging can be one possible therapeutic approach against aging and age related pathologies.

## **CHAPTER 2: LITERATURE REVIEW**

### **2.1 Age related pathologies**

#### 2.1.2 Atherosclerosis

Atherosclerosis is a disease of the arteries in which gradual plaque buildup obstructs the blood flow and results in serious complications like stroke, heart attack and kidney malfunction[10]. Pre clinical studies on atherosclerotic/transgenic (Ldlr<sup>-/-</sup>) mice models have disclosed the contribution of senescence in disease pathogenesis. At the onset of atherosclerosis, accumulation of senescent-foamy macrophages within the subendothelial space is observed wherein they drive plaque buildup by enhancing the expression of inflammatory and atherogenic chemokines and cytokines. As the disease progresses, SCs residing within the advanced lesions promote plaque instability by increasing the production of metalloproteinases that degrade elastic fibers and thin down the fibrous caps [10].

#### 2.2.2 Chronic kidney disease

In chronic kidney disease, the gradual loss of renal function is supplemented with accumulation of SCs within different parts of the kidney such as the medulla and cortex along with vascular, glomerular, tubular and interstitial cells. These senescent cells facilitate renal fibrosis by stimulating proinflammatory signaling pathways that cause oxidative stress which further contributes to the loss of renoprotective factors as well as vascular rarefaction. Experimental studies have revealed that on administration of senolytic agents, kidney function is better preserved during aging [11].



### 2.2.3 Cancer

Cancer pathogenesis entails a string of complex intrinsic processes that favors their establishment and persistence within the body. One such process that adds on to tumor burden is the cross talk of SCs to their environment. Several studies have reported that a fraction of SASPs synthesized by the resident SCs of the TME encourage the uncontrollable growth, invasiveness and immune evasion of cancer cells [12]. It has been highlighted that this response may manifest in different manners and to varying degrees in a tumor-type dependent fashion. Bearing in mind the pro-tumorigenic potential of SCs and how detrimental it can be, treatment can include strategies that either specifically eliminate SCs (senolytics) or abrogate SASPs expression [12].

### 2.2.4 Neurodegenerative diseases

The sporadic incidences of neurodegenerative diseases increase with advancing age and usually manifests at a much older age. There is evidence that insinuates that cellular senescence in neurons and glial cells may predispose a person to develop age-related neurodegenerative diseases like Alzheimer's disease, Parkinson's disease and multiple sclerosis. The chronic disease generally manifests progressively where there is loss of neurons and impairment of synaptic connections, which ultimately results in functional and cognitive decline [13].

Through their SASPs, senescent cells promote chronic inflammation and deplete the pool of progenitor cells by inducing senescence within them. The cell cycle arrest of neuronal cells also results in silencing of genes that play a crucial role in conduction of nerve impulses. Furthermore, cerebral hypoperfusion and blood-brain barrier (BBB) dysfunction may arise due to changes in the cerebral microvascular structure [13].

## **2.2 Senoantigens**

### 2.2.1 uPAR

uPAR or CD87 encoded by PLAUR is an integral part of the urokinase-type plasminogen activator (uPA) system, which is engaged in normal physiological events such as tissue degradation and reorganization. The uPA system also plays a major role in inflammatory responses, tumorigenesis, metastasis, and embryonic development [14-17,20,22].

### 2.2.2 GPNMB

GPNMB is a membrane protein which is typically expressed on melanocytes, macrophages, dendritic cells, osteoclasts, and osteoblasts. GPNMB overexpression has been correlated with several aggressive forms of breast cancer, melanoma, and bone cancer [18-21].

### 2.2.3 Rationale for selection

Analysis of gene expression profiles of senescent and young human umbilical vein endothelial cells (HUVECs) revealed that uPAR and GPNMB transcripts were highly upregulated among senescent HUVEC cells. Independent in vivo studies on uPAR mice knockouts and GPNMB mice knockouts showed that the mice models retained their normal physiology and viability, thereby suggesting that both these proteins function autonomously without interfering with any signaling pathways critical for survival [20-23].

Owing to their remarkable senescent cell specificity and clinical relevance, uPAR and GPNMB senoantigens are now being used for preferential targeting and eliminating SCs. Recently, a GPNMB peptide based senovaccine was successful at clearing SCs in mice and reversing disease/aging phenotypes. The GPNMB immunized mice displayed reduced atherogenesis as well as improved life span. This correction of metabolic abnormalities, along with extended longevity attests to the prowess of senolytic vaccines

[15]. Additionally, Amor et al. used uPAR-specific Chimeric antigen receptor T-cells (CAR-T) for senolysis of SC population from mice suffering from lung adenocarcinoma. The restoration of liver homeostasis and enhanced survivability of mice validated the potency of uPAR as a potential target for senolytic treatment [4].

### **2.3 Research Hypothesis**

Traditionally used senolytics have proven to be efficacious, but their adverse effects, such as off-target toxicity and bystander killing of normal cells makes them less desirable and safe. Therefore, using preemptive, long-lasting measures like vaccines that specifically target SCs and facilitate their removal through non-apoptotic immune-mediated pathways can be an optimal substitute to senolytic drugs.

To counter the aforementioned challenges, we have computationally designed and cloned a novel uPAR and GPNMB based B cell multiepitope senovaccine that can specifically target senescent cells. Our immunoinformatics pipeline involved the prediction of prospective antigenic linear B-cell epitopes derived from the extracellular domain of uPAR and GPNMB followed by protein-protein docking to ascertain the binding efficacy of our senovaccine constructs with the Fab region of an anti-uPAR antibody.

The results of our *in silico* experiment serve as proof of concept for using uPAR-GPNMB based B cell epitope senovaccine, that not only resolves the vices of senolytics but may also be used as a prophylactic that may potentially tackle age-related pathologies and enhance the quality of life of the aging population.

## **CHAPTER 3: METHODOLOGY**

### **3.1 Insilico vaccine designing**

Insilico vaccine designing is a bioinformatics oriented strategy wherein novel vaccine constructs are created using sophisticated softwares and complex computing machinery. Different techniques like epitope prediction, molecular docking, structure prediction and modeling (either ab initio or homology), sequence alignment, molecular dynamic simulations etc. are included in this immunoinformatics process.

The fundamental idea behind in-silico vaccine designing is to predict and simulate the interaction between two or more molecules herein, a target protein/receptor and the vaccine construct on a computer. A favorable interaction between the two molecules facilitates formation of a stable adduct. Such outcomes lay the ground for further investigations where the aim is to validate and replicate the same results and accurately determine the immunogenicity of the vaccine construct in in-vitro systems and within animal models as well.

Computational methods support the rational design of potent and safe vaccine candidates and offer a quicker, foolproof alternative to immunologists while overcoming the challenges of the conventional vaccine designing process.

Application of in-silico methods of vaccine designing has taken a new arc in this digital era and is destined to prosper with the increasing need for better and safer prophylactics.

## 3.2 Research pipeline

### 3.2.1 Sequence retrieval and domain identification

Protein sequences of uPAR (UPAR\_HUMAN, UniProt ID: Q03405) and GPNMB (GPNMB\_HUMAN, UniProt ID: Q14956) were retrieved from the UniProt database [24] and analyzed for their protein topology on TMHMM2.0 [25,26]. The extracellular domains of uPAR and GPNMB protein were found to be located between 22-335 AA and 1-496 AA, respectively.

### 3.2.2 Linear B-cell epitope prediction

To identify potentially antigenic uPAR and GPNMB epitopes, different B-cell epitope prediction tools offered by Immune Epitope Database and Analysis Resource (IEDB) were used [27]. 9 consensus epitope sequences were shortlisted using a combination of prediction tools<sup>1</sup> such as BepiPred 2.0 (sequential B-cell epitope prediction, threshold : 0.500) [28], Chou and Fasman (beta-turn prediction, threshold: 1.048) [29], Emini (surface accessibility, threshold: 1.000) [30], Karplus and Schulz (flexibility, threshold : 1.003) [31], Parker (hydrophilicity, threshold: 2.314) [32], and Kolaskar and Tongaonkar (antigenicity, threshold : 1.033) [33] .

### 3.2.3 Evaluation of predicted linear B-cell epitopes

The predicted epitopes were validated for their antigenicity on Vaxijen v2.0 against the tumor model set at a threshold of 0.5. Vaxigen adopts an alignment-independent approach wherein peptide sequences are classified into probable antigens on the basis of their physicochemical properties [34, 35]. Allergenicity was tested on AllergenFp [36] which transforms input sequences into uniform vectors and tests them for their physicochemical properties such as hydrophobicity, size, etc that are defined within the five e-descriptors. The toxigenicity of the predicted epitopes was determined on ToxinPred server that uses a SwissProt based trained SVM classifier [37].

---

<sup>1</sup> Refer to **Appendix 1** for more information on the linear B-cell epitope prediction tools

#### 3.2.4 Visualization of the linear B-cell epitopes

Pymol was used to visualize the location and orientation of the shortlisted linear B-cell epitopes on their respective protein structures, uPAR (PDB ID: 3U74) and GPNMB (AlphaFold: AF-Q14956-F1) [38].

#### 3.2.5 Construction of a B-cell multiepitope senovaccine and determination of its features

The epitope candidates that reported the highest antigenicity were joined together in an array using GPGPG linker peptides. Adjuvant human beta-defensin-1 (Uniprot ID: P60022) was added using (EAAAK)<sub>2</sub> linkers at the N-terminus. The final senovaccine construct was assessed for its antigenicity on Vaxijen v2.0 [34, 35], allergenicity on AllergenFp [36], and toxigenicity on ToxinPred [37].

#### 3.2.6 Determination of physicochemical properties

The physicochemical properties of the senovaccine construct were determined using ProtParam [39] of the ExPASy server. Parameters like the AA composition, theoretical pI, molecular weight, instability index, aliphatic index, grand average of hydropathy (GRAVY) and estimated half-life were computed using this tool. The solubility of the senovaccine construct was predicted on Protein-Sol [40], a web-tool algorithm that calculates for 35 sequence features and compares predicted solubility to the solubility of the population average for the experimental dataset (threshold: 0.45).

#### 3.2.7 Secondary structure prediction

The secondary structure of the senovaccine construct was predicted on PSIPRED 4.0 [41] workbench which evaluates the position specific scoring matrices of the query sequence via a two stage neural network. Self-optimized prediction tool called SOPMA [42] was used for determining the distribution of the various secondary structures within the vaccine. The analysis was carried out at default parameters- similarity threshold:8, number of conformational states:4 and window width:17.

### 3.2.8 Tertiary structure prediction

The tertiary structure prediction of the senovaccine construct was performed on I-TASSER [43-45], an iterative protein threading assembly algorithm that takes both sequence homology and structural information in account.

### 3.2.9 Structural refinement and validation of the senovaccine

GalaxyRefine [46,47] was used to improve the quality of the predicted tertiary structure through successive structural perturbation and relaxation simulations. The parameters of refined structure were computed and validated on MolProbity [48] using the Ramachandran plot. The overall model quality, energy plot and Z-score were further validated using ProSA [49, 50].

### 3.2.10 Molecular Docking of vaccine construct on uPAR antibody

As the structure and sequence of the immunological B-cell receptor against uPAR and GPNMB were unavailable/unknown, we chose to perform a protein-protein docking of our senovaccine against a well characterized anti-uPAR antibody ATN-658 (PDB ID: 4K23) [51], that has previously been used for uPAR epitope mapping and cancer treatment, in order to assess the molecular affinity of our senovaccine construct.

The antibody mode on ClusPro [52] was used for docking the senovaccine construct on the Fab region of the anti-uPAR antibody ATN-658 (PDB ID: 4K23) [53]. To identify the best senovaccine-Ab model, the generated clusters were screened and analyzed for the following parameters: protein-protein interface residues (determined using PDBSum [54] and visualized on PyMol [38]), cluster size, and lowest energy coefficients. The best fit was selected for further analysis. ParaPred [55] was used to identify the CDRs of the anti-uPAR antibody AT-658.

### 3.2.11 Molecular dynamics simulation

Coarse graining C $\alpha$ -NMA (Normal Mode Analysis) simulation of the best docking model/pose was performed on iMODS [56] online server to determine the overall

stability of the senovaccine-anti-uPAR antibody complex. C $\alpha$ -NMA simulation model predicts the collective functional motion and flexibility of the macromolecule by using internal coordinates of the dihedral angles. Plots for B factor per residue, deformability, eigenvalues and covariance were computed and analyzed. Covariance map and elastic network was also assessed.

### 3.2.12. Vaccine optimization and insilico cloning

Back translation of the aa sequence of the multiepitope senovaccine was done using the gene infinity server [57]. The generated coding sequence was analyzed for rare codon usage and values for GC content, CAI and CPD were determined on GeneScript [58].

For efficient expression of the senovaccine construct within a heterologous host, *E.coli* plasmid pET-28a(+) was chosen as an expression vector. The restriction enzyme cleavage sites of the vector and the coding sequence were identified and prepared using NEBcutter [59]. Designing and visualization of the *in-silico* vaccine carrying expression vector/clone was done on SnapGene6.2.2 Viewer [60].



## CHAPTER 4: RESULTS

### 4.1. Prediction and screening of linear B-cell epitopes

On TMHMM analysis, the extracellular domains of uPAR and GPNMB protein were found to be located between 22-335 AA and 1-496 AA, respectively. The extracellular domains of these senoantigens were then used as a query sequence for the prediction of linear B-cell epitopes on IEDB. A total of 1238 epitopes were predicted for the uPAR antigen, and 2467 epitopes were predicted for the GPNMB antigen. The location of the top scorers lied between 100-220 amino acids for the uPAR antigen and between ranges 20-70, 100-170 and 320-370 amino acids for the GPNMB antigen. These ranges served as the lower and upper limits for subsequent analyses<sup>2</sup>.

After identifying the top scorers and eliminating peptides using the threshold limits of each program, the number of epitopes came down to 505 for uPAR and 1035 for GPNMB<sup>3</sup>. Out of this cohort, we finally identified 9 consensus peptide sequences that were highly antigenic (Vaxijen, threshold: 0.500) and fulfilled the criteria of non allergenicity (AllergenFp) and non toxicity (ToxinPred), as illustrated in Table 4.1. The top five highly antigenic epitopes were visualized on PyMol (Figure 4.1).

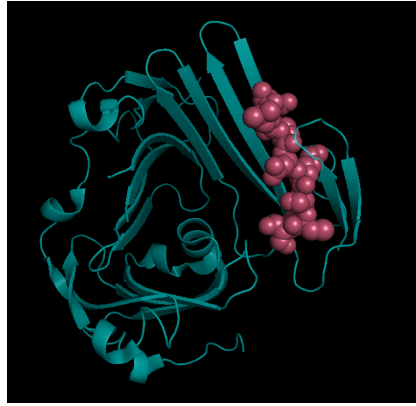
---

<sup>2</sup> Refer to **Appendix 2 and 3** for the graphs obtained from IEDB B-cell epitope prediction tools.

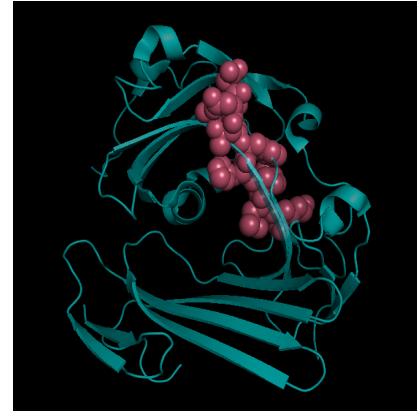
<sup>3</sup> Refer to **Appendix 4 and 5** for the list of shortlisted uPAR and GPNMB epitopes.

**Table 4.1: Predicted B cell epitopes of uPAR and GPNMB**

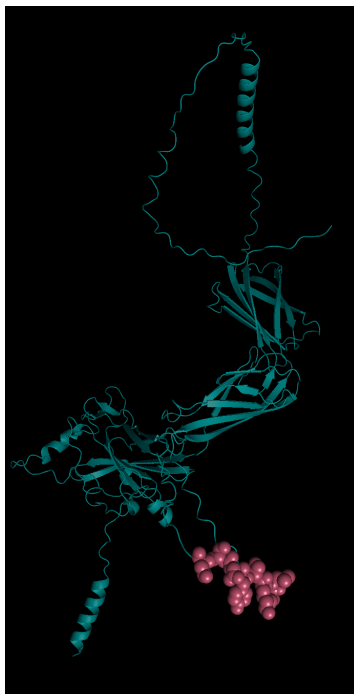
<b>uPAR senoepitopes</b>					
<b>S.no</b>	<b>Consensus B cell epitopes</b>	<b>Location and Length</b>	<b>Antigenicity (Threshold: 0.5)</b>	<b>Allergenicity (Tanimoto coefficient)</b>	<b>Toxicity (SVM Scores)</b>
1.	ELVEKSCT*	61-68 8 AA	1.2874	Non-Allergen (0.72)	Non-Toxin (-0.79)
2.	TLSYRTGLK	76-84 9 AA	0.8347	Non-Allergen (0.72)	Non-Toxin (-1.33)
3.	NNDTFHFLK*	183-191 9 AA	1.1406	Non-Allergen (0.75)	Non-Toxin (-0.76)
4.	LENLPQNGR	206-214 9 AA	0.8762	Non-Allergen (0.69)	Non-Toxin (-0.67)
<b>GPNMB senoepitopes</b>					
1.	VLGNERP	28-34 7 AA	0.9069	Non-Allergen (0.73)	Non-Toxin (-1.33)
2.	KNSWKGG*	70-76 7 AA	1.5934	Non-Allergen (0.77)	Non-Toxin (-0.66)
3.	EAGLSADP	123-130 8 AA	0.9489	Non-Allergen (0.71)	Non-Toxin (-0.69)
4.	NGTGQSHHNV*	146-155 10 AA	1.7441	Non-Allergen (0.73)	Non-Toxin (-0.63)
5.	TLKSYDSN*	342-349 8 AA	1.0162	Non-Allergen (0.75)	Non-Toxin (-1.18)
<p><b>* Highlights the epitopes selected for the multiepitope vaccine construct</b>  <b>SVM Score:</b> A negative SVM score implies non toxigenicity.  Non-toxin &lt;0.00 &lt; Toxin  <b>Tanimoto coefficient:</b> A quantitative metric used to describe the level of similarity between the training dataset and the input query.</p>					



A.



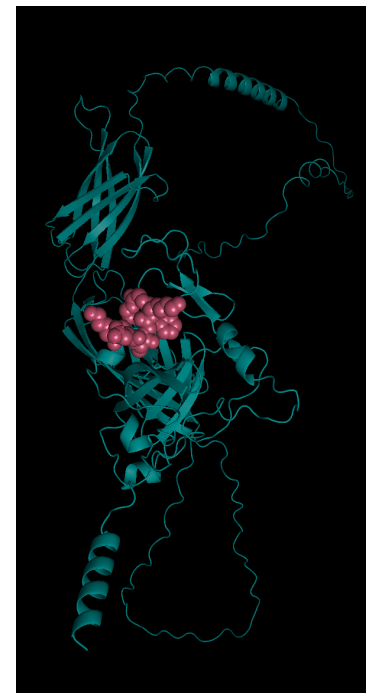
B.



C.



D.



E.

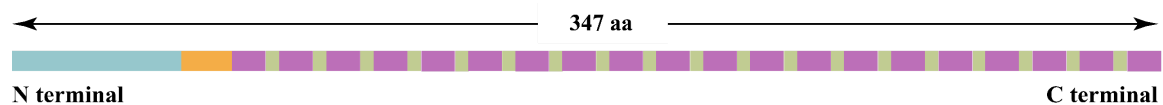
**Figure 4.1: Visualization of the most antigenic linear B-cell epitopes**

- A. Visualization of ELVEKSCT epitope (marked in pink) on uPAR protein domain (PDB ID: 3U74).
- B. Visualization of NNDFHFLK epitope (marked in pink) on uPAR protein domain (PDB ID: 3U74).
- C. Visualization of TLKSYDSN epitope (marked in pink) on GPNMB (Alpha Fold: AF-Q14956-F1).
- D. Visualization of NGTGQSHHNV epitope (marked in pink) on GPNMB (Alpha Fold: AF-Q14956-F1).
- E. Visualization of KNSWKGG epitope (marked in pink) on GPNMB (Alpha Fold: AF-Q14956-F1).

## 4.2. Vaccine designing and feature prediction of the construct

Five epitopes with the highest antigenicity (>1.0000) were selected and joined together with GPGPG linkers to construct a linear B-cell multiepitope senovaccine. Adjuvant human beta-defensin-1 (Uniprot ID: P60022) was added using (EAAAK)<sub>2</sub> linkers at the N-terminus to increase the immunogenicity of the senovaccine (Figure 4.2). To ensure direct activation of the antagonistic B cell clones without any T cell intervention, the aforementioned senoepitopes were repeated multiple times throughout the vaccine construct. We hypothesize that such a repetition of epitopes within our senovaccine would trigger B-cell receptor clustering, which in turn would facilitate the generation of a much more productive humoral immune response against senescent cells.

The designed B cell multi-epitope senovaccine was 347 AA long and showed excellent antigenicity of 0.8402 (Vaxijen, threshold: 0.500). It was classified as a non allergen with a Tanimoto index of 0.78 by AllergenFp server. The vaccine construct also fulfilled the criteria of non toxicogenicity and was classified as a non toxin by ToxinPred.



**Figure 4.2:** A graphical representation of the 347 AA long B-cell multiepitope senovaccine. The linear B cell epitopes (marked in purple) are successively joined together by GPGPG linker (marked in green). The adjuvant (in blue) is located at the N terminal end and linked to the epitopes via (EAAAK)<sub>2</sub> linkers (marked in orange).

## 4.3. Physicochemical analysis of the senovaccine construct.

The physicochemical properties were determined using ProtParam tool, the molecular weight and the theoretical pI of the senovaccine were 34.43kDa and 8.81, respectively.

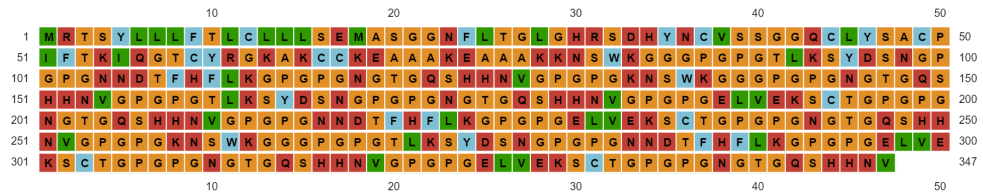
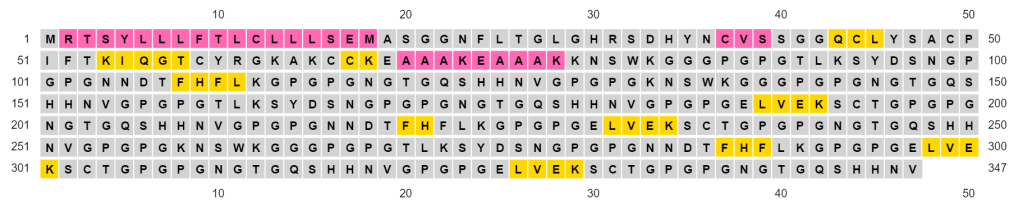
It was noted that the senovaccine construct is relatively stable, with a low instability index (II) score of 22.68 (>40: unstable). The aliphatic index was 37.35% which suggests modest thermostability of the protein, and the GRAVY index was found to be -0.835, illustrating its hydrophilic properties. The ProteinSol server gave a predicted scaled solubility value of 0.585 against the population average of 0.45, indicating that the senovaccine construct was highly soluble. Table 4.2 summarizes the features of the B-cell multi-epitope senovaccine.

**Table 4.2: Features of the B-cell multi-epitope vaccine construct**

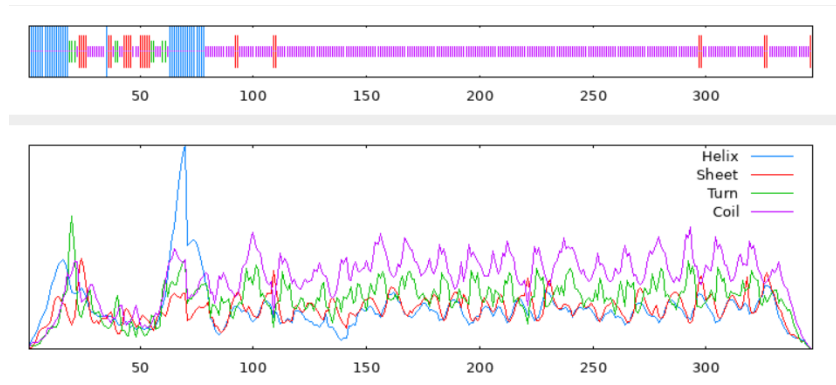
S.no	Property	Insilico tool	Value	Result
1.	Antigenicity	Vaxigen (threshold: 0.500)	0.8402	Probable Antigen
2.	Allergenicity	Allergenfp	0.78	Non-allergen
3.	Toxicity	ToxinPred	-	Non-toxin
4.	Instability index	ProtParam	22.68	Stable
5.	Solubility	Protein-sol (threshold: 0.45)	0.585	Highly soluble
6.	Molecular Weight	ProtParam	34439.63 Da	Probable immunogen

#### 4.4. Protein structure prediction and validation of the vaccine construct.

PSIPRED and SOPMA predicted that the vaccine construct was abundant in random coils (80.40%) and had smaller stretches of alpha helices (10.09%), beta strands (6.92%) and beta turns (2.59%). Alpha helices were located within the built-in beta-defensin adjuvant and the EAAAK linker. Beta strands were found to be formed within the repetitive units of uPAR epitopes “ELVEKSCT” and “NNDTFHFLK”. Furthermore, the analysis also revealed that the senovaccine construct was prevalent in small non-polar amino acids due to the presence of GPGPG linkers (Figure 4.3).



A.



Sequence length : 347

SOPMA :

Alpha helix	(Hh)	:	35 is	10.09%
$\beta_{10}$ helix	(Gg)	:	0 is	0.00%
Pi helix	(Ii)	:	0 is	0.00%
Beta bridge	(Bb)	:	0 is	0.00%
Extended strand	(Ee)	:	24 is	6.92%
Beta turn	(Tt)	:	9 is	2.59%
Bend region	(Ss)	:	0 is	0.00%
Random coil	(Cc)	:	279 is	80.40%
Ambiguous states (?)		:	0 is	0.00%
Other states		:	0 is	0.00%

B.

**Figure 4.3: Secondary structure prediction of the senovaccine construct.**

**A.** Secondary structure prediction of the senovaccine construct on PSIPRED

**B.** Secondary structure prediction of the senovaccine construct on SOPMA

The tertiary structure prediction of the senovaccine construct was performed on I-TASSER that generated 5 protein structure models, out of which the most suitable model had the C-score of -2.78 (highest amongst the predicted models), a TM score of  $0.40 \pm 0.13$  and a RMSD of  $13.2 \pm 4.1$ . A high C-score value corresponds to a model with a high prediction confidence. A TM score  $> 0.17$  signifies that the predicted model does not share any random similarity with the native structures of the protein. Furthermore, the number of decoys (low temperature replicas) generated for this particular model were 1465, forming the largest cluster with a cluster density of 0.400. A greater cluster density indicates that the predicted tertiary structure occurs more frequently in the simulation trajectory and hence can be regarded as the optimal model.

The quality of the I-TASSER predicted model was further refined on GalaxyRefine<sup>4</sup> through successive structural perturbation and relaxation simulations. Out of the 5 refined models generated, the best model had a GDT-HA value of 0.8818, MolProbity of 1.977, an RMSD value of 0.599, and a clash score of 8.6 (Figure 4.4). The structural refinement also resulted in a substantial increase in the percentage of AA residues lying within the energetically favorable regions (Rama favored), from 56.5% to 88.7%. These scores were validated using MolProbity and ProSA.

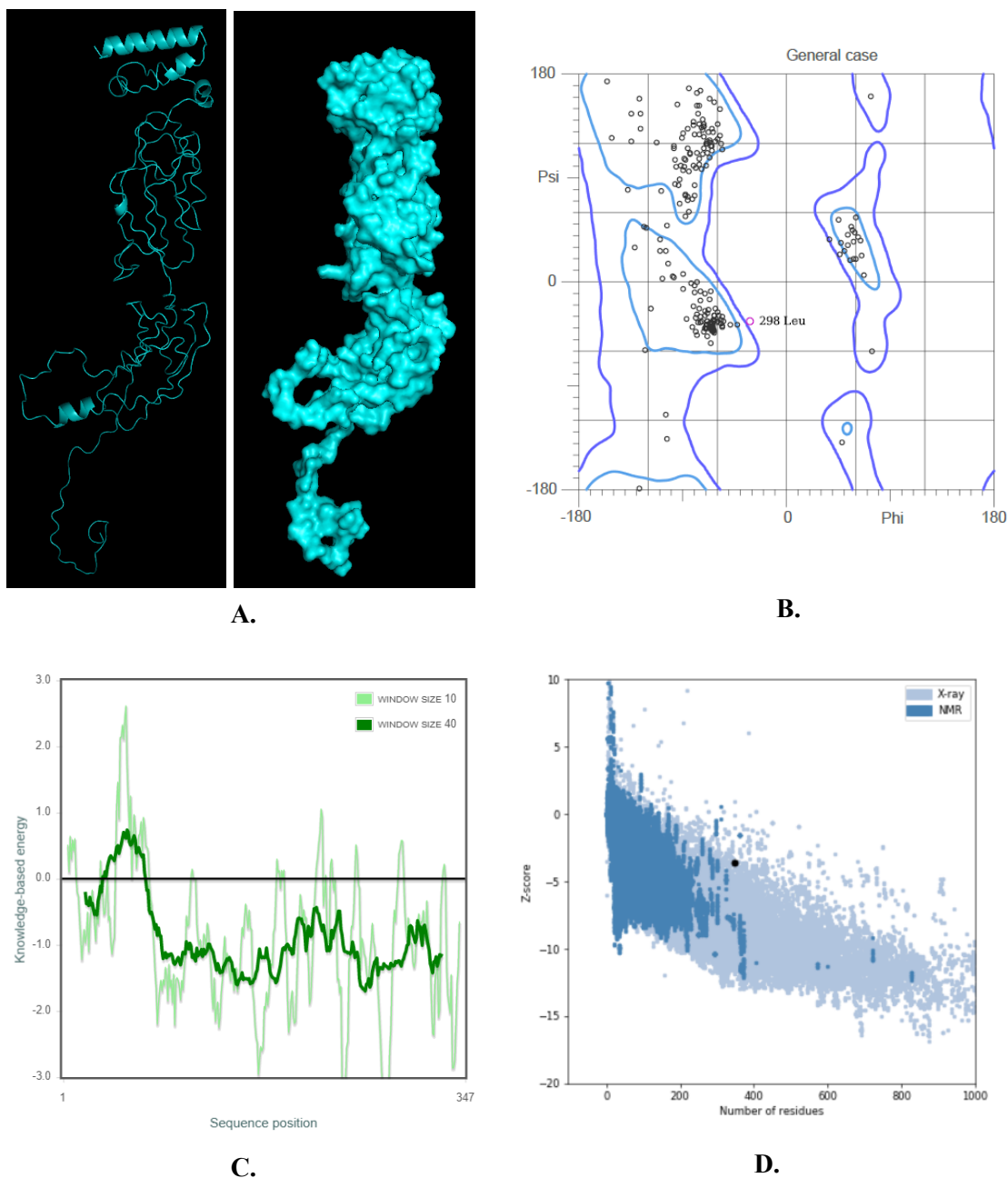
Ramachandran analysis on MolProbity confirmed that 88.7% of all AA residues resided within the favored region, while 98.3% of all AA residues were in the allowed regions. 6 outliers were identified that influenced the protein geometry and contributed to the sub-optimal tertiary structure prediction results. These outliers mainly consisted of glycine and proline residues of the GPGPG linkers within the vaccine construct (Figure 4.4).

ProSA computed a Z-score of -3.6 for the refined model, which as per the Z-score plot resides within the acceptable ranges of experimentally determined Z-score values. The Z-score reflects the overall model quality, which in this case is suboptimal due to the presence of certain erroneous regions. N-terminal of the vaccine containing the adjuvant

---

<sup>4</sup> Refer to **Appendix 6** for the GalaxyRefine results.

sequence has amino acid residues with higher energy values, while senso-peptides fall under the region of lower and non-offending energies (Figure 4.4).



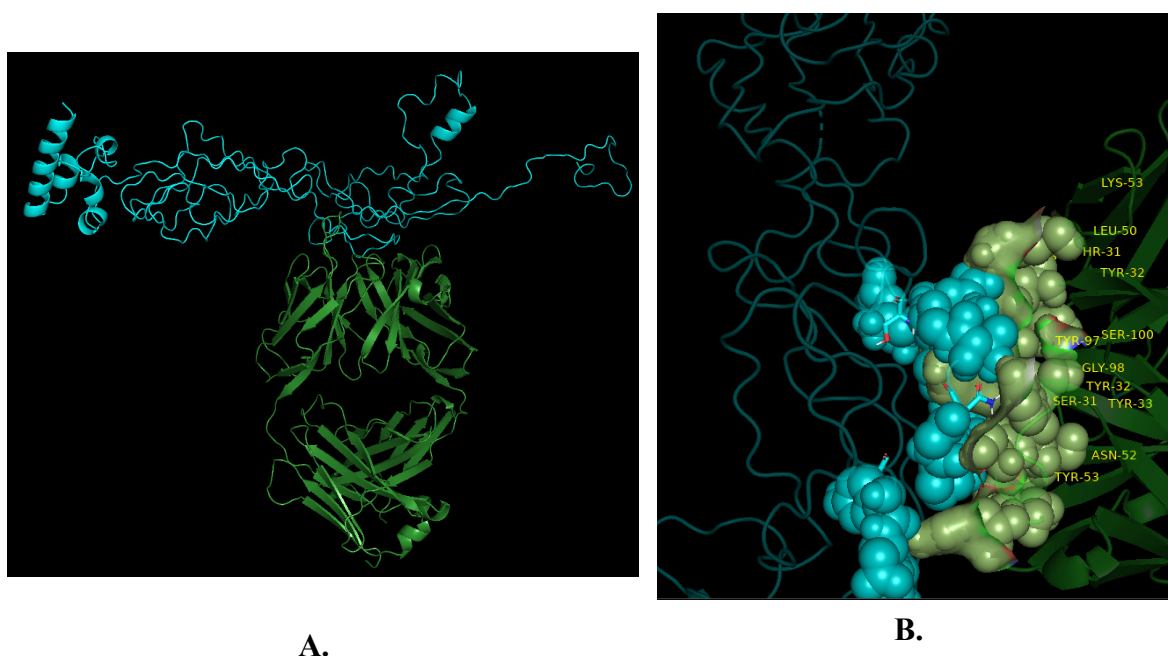
**Figure 4.4 : Tertiary structure and validation of the refined protein structure of the senovaccine construct** **A.** Visualization of the refined tertiary structure of the senovaccine construct using PyMol. **B.** Ramachandran plot from MolProbity illustrating the location of the constituent AA residues of the vaccine construct. **C.** Energy plot from ProSA of the predicted structure of the senovaccine construct. **D.** Z-score plot from ProSA showing the overall model quality of the refined protein structure with a Z-score of -3.6.



#### 4.5. Molecular docking of the senovaccine on anti-uPAR antibody

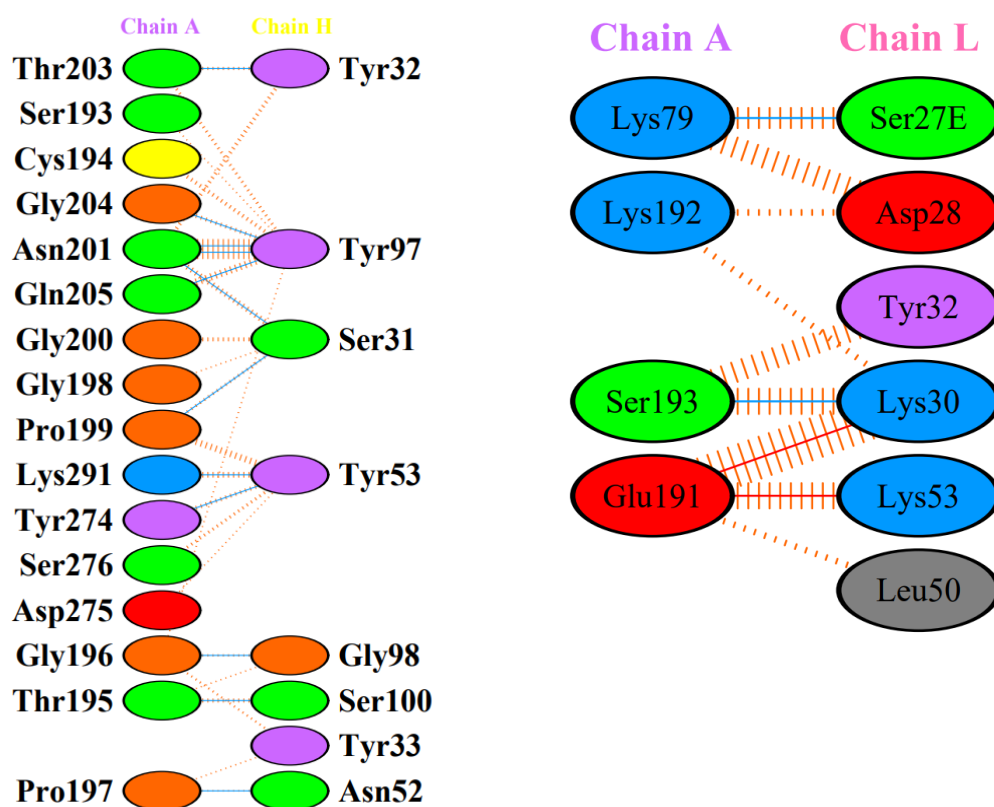
The results of the protein-protein docking on ClusPro confirmed that our senovaccine construct has a propensity to bind to the Fab region of a corresponding anti-uPAR antibody (ATN-658). Out of the 29 clusters generated, the most favorable senovaccine-Ab complex belonged to the largest cluster which had 225 members and a weighted lowest energy score of -337.0 Kcal/mol.

On analyzing the docked pose on PDBSum and PyMol, it was found that the interacting interface residues of the anti-uPAR antibody overlapped with the predicted and experimentally validated CDR regions (Table 4.3)(Figure 4.5). Furthermore, PDBSum protein-protein interactions also revealed that the interface residues of the senovaccine involved in antibody interactions were emerging for the uPAR epitope “ELVEKSCT”, the GPNMB epitope “TLKSYDSN” and the GPGPG linker (Figure 4.6).



**Figure 4.5 Molecular docking of senovaccine on anti-uPAR antibody AT-658.**

**A.** Visualization of the senovaccine and anti-uPAR antibody AT-658 complex generated after molecular docking. The senovaccine is represented in blue and the anti-uPAR antibody is represented in green. **B.** The residues labeled are the interface residues of the antibody involved in the vaccine-Ab complex.



Interface statistics						
Chains	No. of interface residues	Interface area (Å <sup>2</sup> )	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A:H	16:8	527:570	-	-	12	97
A:L	4:6	238:207	2	-	2	33

**Figure 4.6. Protein-protein interactions between senovaccine construct and anti-uPAR antibody AT-658.** A. PDBSum prot-prot interaction between senovaccine and anti-uPAR antibody AT-658. **Chain A:** Senovaccine construct **Chain H:** Heavy chain of the anti-uPAR antibody **Chain L:** Light chain of the anti-uPAR antibody

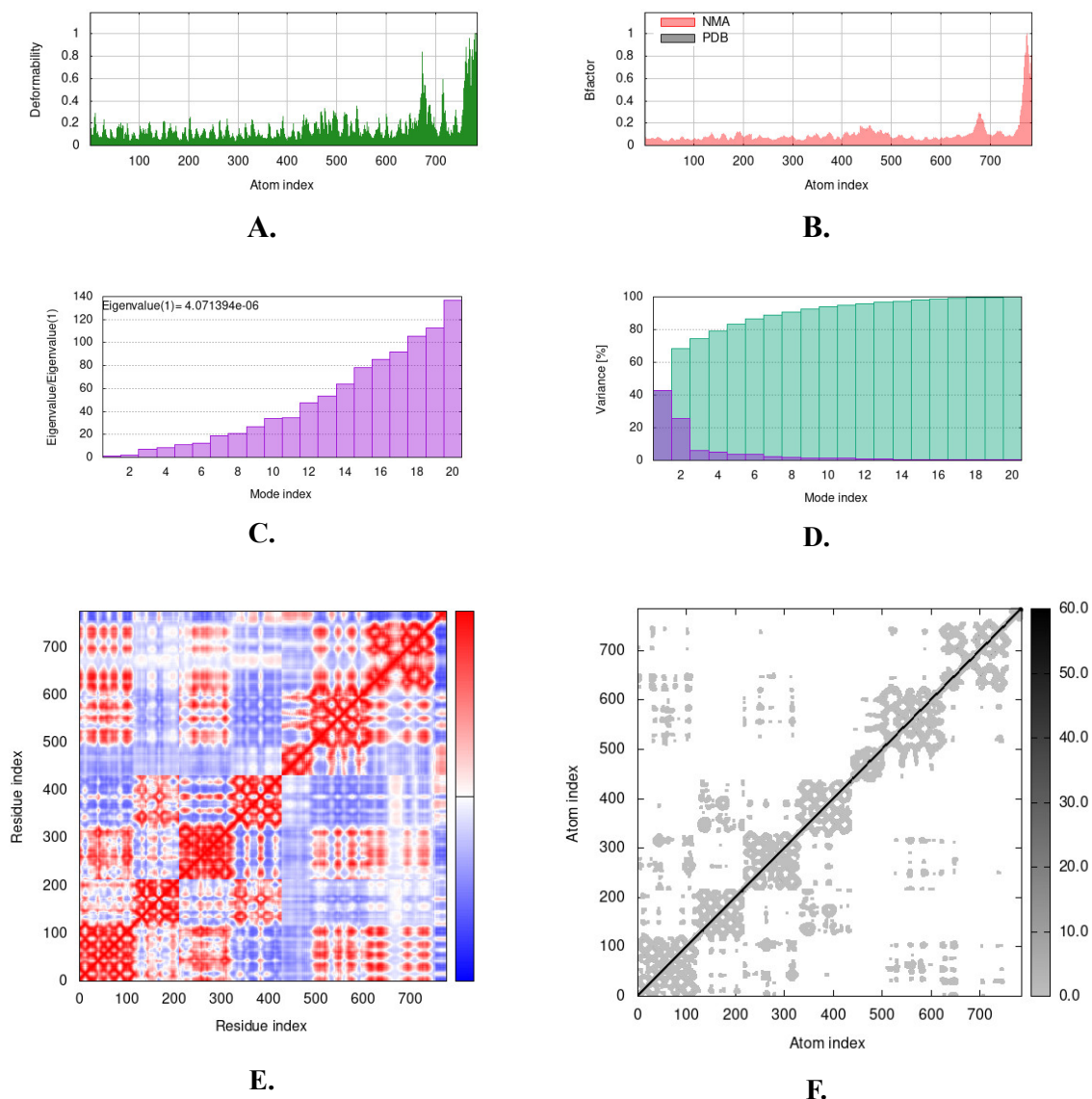
**Table 4.3: Experimentally determined and predicted CDRs of anti-uPAR antibody ATN-658**

<b>Experimentally determined CDRs of anti-uPAR antibody ATN-658 (Xu et al, 2014)</b>			
<b>Heavy Chain</b>	<b>Location</b>	<b>Light Chain</b>	<b>Location</b>
<b>CDR 2: YNQ-K</b>	59-62	<b>CDR1: LDSD</b>	27C-28
<b>CDR 3: YGHSVL</b>	97-101	<b>CDR3: GTHF</b>	91-94
<b>ParaPred prediction of CDRs of anti-uPAR antibody ATN-658</b>			
<b>CDR 1: ASGYSFTSYYM</b>	24-34	<b>CDR1: SCKSSQSLDSD GKTYLNWL</b>	22-34
<b>CDR 2: EINPYNGGAS</b>	50-59	<b>CDR2: IYLVSKLDSGV</b>	53-63
<b>CDR 3: ARSIYGHSVLDY WG</b>	97-110	<b>CDR3: YCWQGTHFPLTF G</b>	92-104

#### **4.6 Molecular Dynamics of the senovaccine and anti-uPAR antibody complex**

Coarse graining-NMA (Normal Mode Analysis) simulation performed on iMODS online server revealed that our senovaccine and anti-uPAR antibody complex is stable and minimally deformable. However, there were certain amino acid residues of the senovaccine (K235, N277, E325, S330, G337, N338, Q342, N347) that showed a high degree of deformability. These residues are represented as peaks in the deformability graph and are termed as “hinges”. It can be inferred that since these AA residues are not involved in the protein-protein interaction (refer PDBSum plots), they show a higher propensity to distort when the equilibrium of the complex is disturbed. The B-factor per residue reflects the average RMSDof atoms. The peaks for the deformability plot overlap with the peaks of the B-factor per residue plot, therefore implying that the high deformability regions of the complex have a greater B-factor value and thermal mobility. The Eigenvalue of the complex was computed to be at  $4.071394 \times 10^{-06}$ , indicating that a higher force and energy may be required to perturb the complex. Low eigenvalue favors

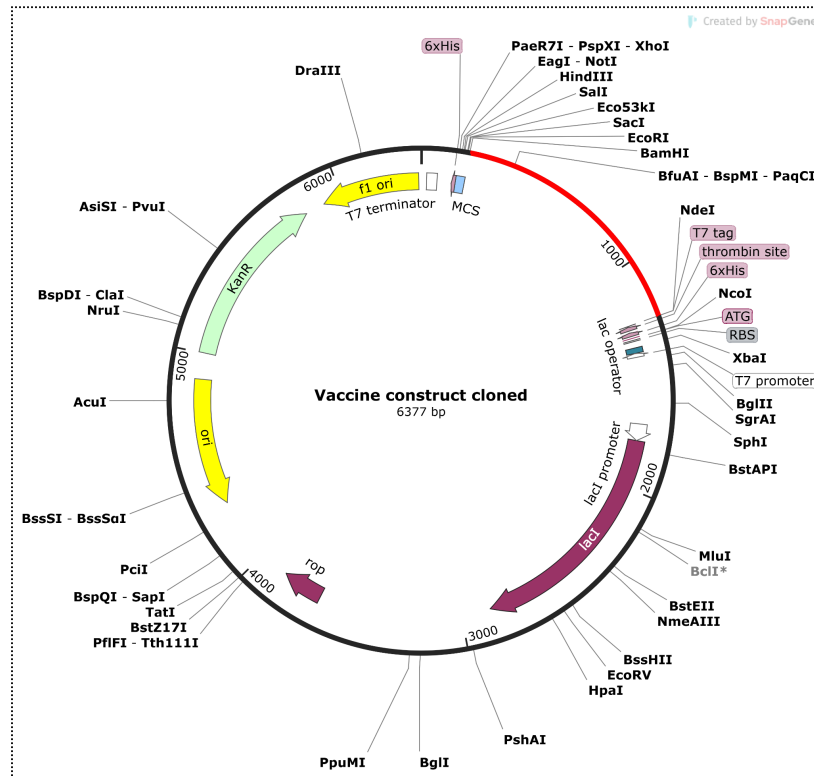
easier deformation. The covariance map and elastic network suggests that the pair of residues that experience correlated motions have stiffer spring interactions. Together, all these results confirm that our novel senovaccine and the anti-uPAR antibody form a stable complex (Figure 4.7).



**Figure 4.7: Results from the NMA molecular dynamics simulation conducted on iMODs. A. Deformability plot:** The peaks reflect the locations of the residues with high deformability values. **B. B factor per residue plot:** Root mean square deviation. **C. Eigenvalues Plot:** Describes relative modal stiffness. **D. Variance:** Describes relative contribution of modes to equilibrium motion. **E. Covariance map:** Describes relative motion of the residues. Red: correlated motion, White: uncorrelated motion, Blue: Anti-correlated motion. **F. Elastic network:** Linking matrix that describes the pair of atoms that are connected by springs. Stiffer springs are represented in darker gray color.

#### 4.7 Codon adaptation and in silico cloning

The gene infinity server performed back translation of the multi-epitope senovaccine construct to the most likely DNA sequence. GeneScript tool was used to assess its expression potential based on properties of Codon Adaption Index (CAI) and GC content. The actual CAI value of the sequence coincided with the ideal value of 1, and the GC content was calculated to be 65.87% (ideal range 30%-70%). The gene infinity output was analyzed in NEBcutter and restriction sites BamHI and NdeI were included in the DNA sequence in accordance with the multiple cloning site of the selected expression vector pET28a(+). In silico clone was prepared on the SnapGene 6.2.2. software, as shown in Figure 4.8.



**Figure 4.8:** In silico cloning map of the B cell multi-epitope senovaccine sequence inserted into the pET28a(+) vector. The red highlighted area shows the placement of the insert using restriction enzymes BamHI and NdeI.

## **CHAPTER 5: DISCUSSION**

The life expectancy continues to outpace health span with each passing year, resulting in an increased proportion of aging individuals suffering from debilitating ailments that depreciate their quality of life. Cellular senescence, which is one of the key players in expedited aging, requires highly specific interventions to potentially delay or reverse the hallmarks of aging and restore vitality. A therapeutic approach deployed is the use of senolytic agents that selectively eliminate SCs and diminish their pathogenic effects within the body. Despite the success of senolytics in preclinical trials and in vitro studies, questions regarding their toxicity and off-target effects remain unanswered. There are concerns that need further investigation regarding the pragmatism of this therapy, including (1) its translation to clinical trials, (2) safety of the therapy in aged individuals and (3) effectiveness in its ability to resolve variable types and stages of age-related pathologies. To sustain the senolytic response without magnifying the toxicity, prophylactics like senovaccines can be used to generate an adaptive immune response against SCs. Cellular markers that are expressed on the surface of SCs can be used as potential immunogens for fabricating senovaccines.

Unlike traditional vaccinology approaches that begin with identification, isolation and purification of an antigen, followed by its sequencing, computational vaccine designing pipelines skip all of these tedious and time intensive stages.

Immunoinformatics prediction tools accelerate the process of antigen identification and epitope prediction by cross referencing the properties of a potential vaccine candidate with the large repertoire of experimentally validated immunogens. They allow for the prompt discovery of novel, structurally and functionally uncharacterized immunogenic epitopes from a protein which can be further developed into efficacious prophylactics such as peptide subunit vaccines or DNA/RNA vaccines.

In our study, we adopted an immunoinformatics pipeline to discover potentially immunogenic epitopes of senoantigens uPAR and GPNMB and construct a multi-epitope vaccine that selectively eliminates SCs and diminish their pathogenic, inflammatory, or protumorigenic impacts within the body.

Immune-mediated clearance is extremely specific and is generally supported by two arms of the immune system, B cells, and T cells. Hence, two kinds of vaccines may be created through an informatics approach, a B-cell epitope vaccine, which would elicit a humoral response in the body, and a T-cell epitope vaccine, which elicits a cytotoxic immune response. SCs are essentially aging self-cells. Since peripheral tolerance is more robust and initiates “anergy” within self-reactive T-cell clones, SCs can easily escape T cell-mediated immune responses by expressing self-antigens and immunosuppressive molecules. To overcome the aforementioned challenge we identified unique senoepitopes that are independent of T-cell activation and can directly stimulate the self-reactive B cells. While standard vaccines aim at producing a humoral as well as cytotoxic response, herein, the senovaccine construct created, solely aims at inducing antibody production.

Since surface interaction is a vital component of an immune response, this research study used the extracellular domains of the uPAR and GPNMB antigens for epitope retrieval. Through our immunoinformatics pipeline, we identified nine epitope sequences from the senoantigens uPAR and GPNMB, that showed excellent antigenicity, surface accessibility, hydrophilicity, non-toxicity, and non-allergenicity. Out of these nine, five epitopes with antigenicities  $>1.0$  were used for fabricating the novel B-cell multiepitope vaccine. The final vaccine construct consisted of repetitive units of the five highly antigenic B-cell epitopes to trigger B-cell receptor clustering. These epitopes were joined together by GPGPG and linked to the adjuvant beta-defensin using (EAAAK)<sub>2</sub> linkers. Beta-defensin is a charged antimicrobial peptide that is known for its immunopotentiator activity, wherein it can efficiently stimulate B-cells, macrophages, and dendritic cells [61]. We believe that this built-in adjuvant would supplement the interaction of the conjugated senoepitopes with the B cell clones and trigger a potent humoral response.

EAAAK and GPGPG linkers are used due to their well-regarded stability, which provides functional flexibility to the tertiary structure of the protein [62].

The results of our study suggest that the proposed senovaccine candidate has a high antigenicity of 0.8402 and has a greater propensity of binding and forming stable complexes with the Fab region of the anti-uPAR antibody ATN-658. The weighted lowest energy score of -337.0 Kcal/mol indicates a productive protein-protein interaction, which can primarily be linked to the intermolecular interactions between the residues of the senoepitopes and the antibody CDRs. Additionally, the optimal physicochemical properties of the vaccine construct, such as its molecular weight of 34 kDa, relative thermal stability, abundance of non-polar and polar amino acid residues along with its hydrophilic nature favor its overall stability in the antibody complex.

The aforementioned data supports our vaccine construct as a promising immunogen that can promote selective immune clearance of accumulating senescent cells via antibody effector functions like opsonization, and complement fixation while also potentially evoking a lasting B cell memory pool. Furthermore, the ideal physicochemical properties offer a production advantage wherein our vaccine candidate can be easily purified and used harmoniously as an active agent in a vaccine concoction that consists of water as a main ingredient, a built-in adjuvant beta-defensin, and preservatives.

However, as this senovaccine will be targeting self cells extreme caution must be taken as potential dangers like autoimmunity and hypersensitivity reactions can arise in the long term. Keeping that in mind, the dosage should be kept as low as possible.

This *in-silico* research acts as a proof of concept for devising future senovaccines that can impede chronic disease manifestation like cancer, Alzheimers, arthritis etc. and potentially extend the health span of an aging individual. The next step would be to determine and validate the efficacy and safety of this conceptualized B cell multiepitope senovaccine through *in vitro* and *in vivo* studies.



The vaccine construct can be further optimized to improve its stability by addition of different linkers like AYY. It may also be packaged with other carrier immunogens/adjuvants like keyhole limpet hemocyanin, aluminum, freunds complex or be conjugated with immunopotentiators like CpG /(Macrophage activating lipopeptide-2 (MALP-2) and delivered via suitable lipid vesicles or nanoparticles.

## **CHAPTER 6: CONCLUSION**

Aging is a biological process wherein there is gradual decline and impairment of phenotype and physiological functions due to accumulation of deleterious factors. This study presents a novel, one-of-a-kind B-cell multi-epitope senovaccine that has been derived from the senescent cell surface antigens uPAR and GPNMB. The effectiveness and safety of the vaccine were confirmed computationally by testing its antigenicity, allergenicity, toxicity, solubility, and stability. The vaccine model showed stable productive interaction with the Fab region of an anti-uPAR antibody, attesting to its ability to generate an effective humoral response within in vivo models. The results of our in silico experiment serve as proof of concept for using uPAR-GPNMB based B cell epitope senovaccine, that not only resolves the vices of senolytics but may also be used as a prophylactic that may potentially tackle age-related pathologies and enhance the quality of life of the aging population. With sufficient in vitro and in vivo research, this vaccine may prove to be a revolutionary prophylactic in reversing aging and addressing various age-associated pathologies.

## APPENDIX 1

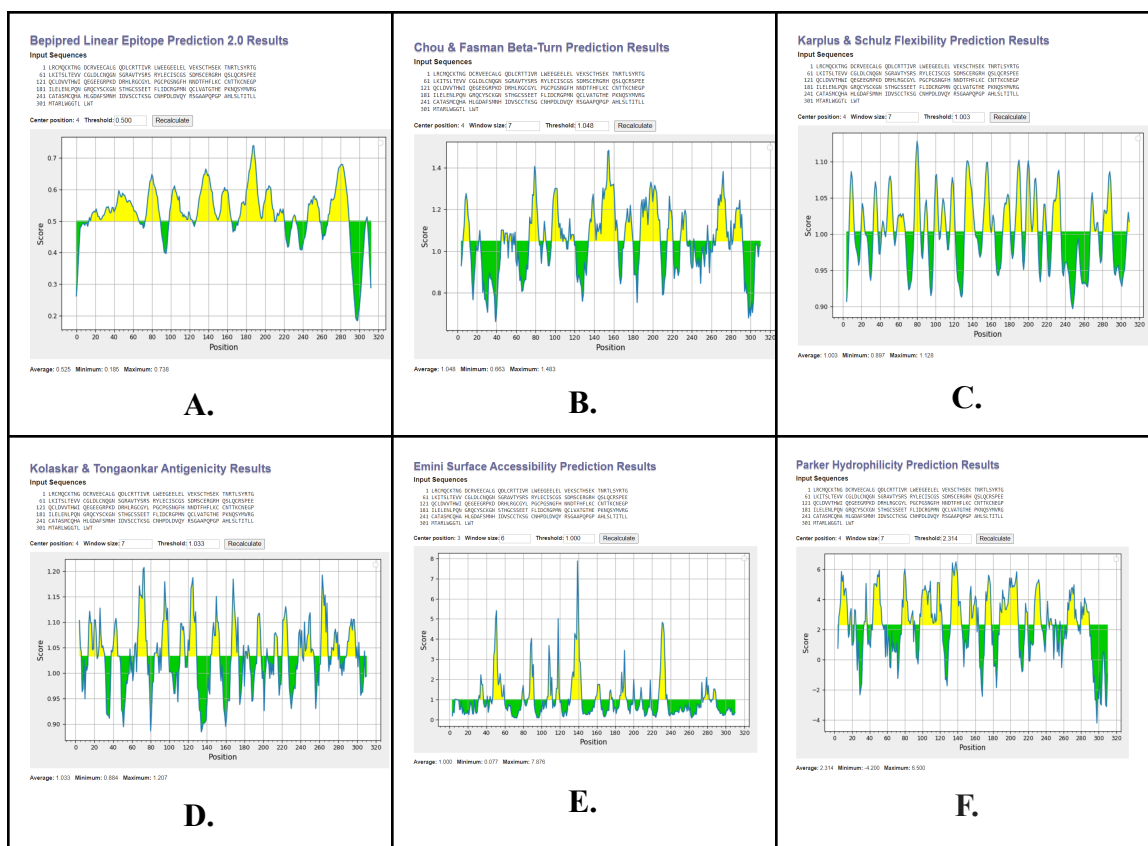
### Linear B-cell epitope prediction tools

1. BepiPred 2.0 (sequential B-cell epitope prediction, threshold : 0.500)  
Operation used: This server uses a Random Forest algorithm to derive epitope sequence stretches from a protein using its crystal structures.
2. Chou and Fasman (beta-turn prediction, threshold: 1.048)  
Operation used: Conceptually derives from the turn scale model for predicting the location of antigenic sites in a protein, this method uses the secondary structure of the input sequence and their beta-turns to predict potential antigenic sites.
3. Emini (surface accessibility, threshold: 1.000)  
Operation used: This surface accessibility scale is a formula-based prediction technique.  
$$\{Formula\ used: S_n (n+4+i) (0.37)^{-6} \}$$

S<sub>n</sub>=surface probability (SB)  
d<sub>n</sub>= fractional SB  
i= (1 → 6)
4. Karplus and Schulz (flexibility, threshold : 1.003)  
Operation used: This technique uses certain known protein and their x-ray structures and B-factors to determine the mobility of a section in the protein.
5. Parker (hydrophilicity, threshold: 2.314)  
Operation used: This method is based on the retention time of a protein/peptide during HPLC.
6. Kolaskar and Tongaonkar (antigenicity, threshold : 1.033)  
Operation used: This tool derives knowledge from experimentally known data and predictable physicochemical proteins of the A.A. residues in a protein. Accuracy rate is 75%.

## APPENDIX 2

### Results from the epitope prediction tool for the uPAR antigen



**Figure A 2.1: Graphs obtained from IEDB B-cell epitope prediction tools for uPAR antigen**

- Bepipred 2.0 epitope prediction (threshold : 0.500); Peptide range: (80-280) amino acids.
- Chou and Fasman beta-turn prediction (threshold value: 1.048); Peptide range: (30-75 ; (8-26), (42-122), (132-223) amino acids.
- Karplus & Schulz Flexibility (Threshold value : 1.003)Peptide range: (9-216) amino acids
- Kolaskar & Tongakar antigenicity scale. (threshold value : 1.033); Peptide range: (11-226) amino acids.
- Emini surface accessibility prediction (threshold value : 1.000); Peptide range: (5-10), (30-221) amino acids.
- Parker Hydrophilicity prediction (threshold value : 2.314); Peptide range: ((4-91),(100-223)) amino acids.



## APPENDIX 4

### Predicted epitopes for the uPAR antigen

**Table A 4.1 : Predicted linear B-cell epitope peptides using BepiPred 2.0 prediction tool**

Threshold value : 0.500

Bepipred 2.0 (0.500)		
Starting	Ending residue	Peptide
15	67	EECALGQDLCRTTIVRLWEEGEELELVEKSCPTHSEKTNRTLSYRTGLKITSLT
74	90	DLCNQGNSGRAVTYSRS
100	124	SSDMSCERGRHQSLQCRSPEEQCLD
128	165	HWIQEGEEGRPKDDRHLRGCGYLPGPCPSNGFHNNDTF
173	210	TTKCNEGPILELENLPQNGRQCYSCKGNSTHGCSEET
213	221	IDCRGPMNQ

**Table A 4.2 : Predicted B-cell epitopes using Parker, Emini, and Kolaskar & Tongaokar prediction tools.**

**Parker Hydrophilicity** : threshold value 2.314

**Emini surface accessibility** : threshold value 1.000

**Kolaskar Antigencity Scale** : threshold value 1.033

Parker (2.314)				Emini (1.000)				Kolaskar and tongaonkar (1.033)			
Starting	Ending residue	Peptide	Score	Starting	Ending residue	Peptide	Score	Starting	Ending residue	Peptide	Score
4	10	MQCKTNG	3.829	5	10	QCKTNG	1.015	11	17	DCRVEEC	1.093
5	11	QCKTNGD	5.857	30	35	RLWEEG	1.2	12	18	CRVEECA	1.121
6	12	CKTNGDC	5.2	31	36	LWEEGE	1.061	13	19	RVEECAL	1.098
7	13	KTNGDCR	5.6	32	37	WEEGEE	2.228	14	20	VEECALG	1.098
8	14	TNGDCRV	4.257	33	38	EEGEEL	1.748	15	21	EECALGQ	1.045
9	15	NGDCRVE	4.629	34	39	EGEELE	1.748	16	22	ECALGQD	1.047
10	16	GDCRVEE	4.743	39	44	ELVEKS	1.171	17	23	CALGQDL	1.104
11	17	DCRVEEC	4.129	42	47	EKSCTH	1.163	18	24	ALGQDLC	1.104
12	18	CRVEECA	3	45	50	CTHSEK	1.163	19	25	LGQDLCR	1.077
15	21	EECALGQ	3.086	46	51	THSEKT	3.131	22	28	DLCRTTI	1.053
16	22	ECALGQD	3.4	47	52	HSEKTN	3.489	23	29	LCRTTIV	1.127
20	26	GQDLCRT	3.329	48	53	SEKTNR	5.022	24	30	CRTTIVR	1.073
21	27	QDLCRTT	3.257	49	54	EKTNRT	5.409	25	31	RTTIVRL	1.05
31	37	LWEEGEE	2.529	50	55	KTNRTL	2.576	26	32	TTIVRLW	1.053
32	38	WEEGEEL	2.529	51	56	TNRTLS	1.726	27	33	TIVRLWE	1.044

**Table A 4.2 (continued)**

33	39	EEGEELE	5.071	52	57	NRTLSTY	1.874	28	34	IVRLWEE	1.036
34	40	EGEELEL	2.643	53	58	RTLSYR	2.282	35	41	GEELELV	1.044
39	45	ELVEKSC	2.329	54	59	TLSYRT	1.682	36	42	EELELVE	1.041
41	47	VEKSCTH	3.571	55	60	LSYRTG	1.153	37	43	ELELVEK	1.052
42	48	EKSCTHS	5.029	56	61	SYRTGL	1.153	38	44	LELVEKS	1.075
43	49	KSCTHSE	5.029	57	62	YRTGLK	1.721	39	45	ELVEKSC	1.098
44	50	SCTHSEK	5.029	77	82	NQGNSG	1.399	40	46	LVEKSCT	1.107
45	51	CTHSEKT	4.843	78	83	QGNSGR	1.704	41	47	VEKSCTH	1.086
46	52	THSEKTN	5.643	83	88	RAVTYS	1.059	55	61	LSYRTGL	1.047
47	53	HSEKTNR	5.5	104	109	SCERGR	1.124	60	66	GLKITSL	1.054
48	54	SEKTNRT	5.943	105	110	CERGRH	1.142	61	67	LKITSLT	1.059
49	55	EKTNRTL	3.7	106	111	ERGRHQ	3.688	63	69	ITSLTEV	1.067
50	56	KTNRTLS	3.514	107	112	RGRHQ	2.854	64	70	TSLTEVV	1.1
51	57	TNRTLSTY	2.429	108	113	GRHQSL	1.202	65	71	SLTEVVC	1.171
56	62	SYRTGLK	2.314	109	114	RHQSLQ	2.103	66	72	LTEVVCG	1.152
74	80	DLCNQGN	3.986	114	119	QCRSPE	1.553	67	73	TEVVCGL	1.152
75	81	LCNQGNS	3.486	115	120	CRSPEE	1.553	68	74	EVVCGLD	1.146
76	82	CNQGNSG	5.614	116	121	RSPEEQ	5.018	69	75	VVCGLDL	1.203
77	83	NQGNSGR	6.014	117	122	SPEEQC	1.373	70	76	VCGLDLC	1.207
78	84	QGNSGRA	5.314	127	132	THWIQE	1.033	71	77	CGLDLCN	1.12
79	85	GNSGRAV	3.929	130	135	IQEGEE	1.485	72	78	GLDLCNQ	1.063
80	86	NSGRAVT	3.857	131	136	QEGEEG	2.097	73	79	LDLCNQG	1.063
81	87	SGRAVTY	2.586	132	137	EGEEGR	2.372	81	87	SGRAVTY	1.039
82	88	GRAVTYS	2.586	133	138	GEEGRP	2.118	94	100	ECISCGS	1.104
83	89	RAVTYSR	2.371	134	139	EEGRPK	4.279	95	101	CISCGSS	1.127
84	90	AVTYSRS	2.7	135	140	EGRPKD	4.127	96	102	ISCGSSD	1.049
85	91	VTYSRSR	3	136	141	GRPKDD	3.979	109	115	RHQSLQC	1.097
				137	142	RPKDDR	7.876	110	116	HQSLQCR	1.097
100	106	SSDMSCE	4.929	138	143	PKDDRH	5.471	111	117	QSLQCRS	1.084
101	107	SDMSCER	4.6	139	144	KDDRHL	2.918	112	118	SLQCRSP	1.091
102	108	DMSCERG	4.486	140	145	DDRHLR	2.858	113	119	LQCRSPE	1.068
103	109	MSCERGR	3.657	141	146	DRHLRG	1.694	117	123	SPEEQCL	1.065
104	110	SCERGRH	4.557	157	162	NGFHNN	1.154	118	124	PEEQCLD	1.044
105	111	CERGRHQ	4.486	158	163	GFHNND	1.199	119	125	EEQCLDV	1.09
106	112	ERGRHQ	5.214	159	164	FHNNDT	1.748	120	126	EQCLDVV	1.166
107	113	RGRHQSL	2.786	160	165	HNNDTF	1.748	121	127	QCLDVVT	1.174
108	114	GRHQSLQ	3.043	161	166	NNDTFH	1.748	122	128	CLDVVTH	1.187
109	115	RHQSLQC	2.429	172	177	NTTKCN	1.375	123	129	LDVVTHW	1.113

**Table A 4.2 (continued)**

110	116	HQSLQCR	2.429	173	178	TTKCNE	1.48	124	130	DVVTHWI	1.099
111	117	QSLQCRS	3.057	174	179	TKCNEG	1.015	125	131	VVTHWIQ	1.12
112	118	SLQCRSP	2.5	175	180	KCNEGP	1.088	126	132	VTHWIQE	1.044
113	119	LQCRSPE	2.686	183	188	ELENLP	1.207	141	147	DRHLRGC	1.036
114	120	QCRSPEE	5.114	184	189	LENLPQ	1.207	142	148	RHLRGC	1.037
115	121	CRSPEEQ	5.114	185	190	ENLPQN	2.355	143	149	HLRGCY	1.078
116	122	RSPEEQC	5.114	186	191	NLPQNG	1.345	144	150	LRGCGYL	1.099
117	123	SPEEQCL	3.2	187	192	LPQNGR	1.639	145	151	RGCGYLP	1.073
118	124	PEEQCLD	3.7	188	193	PQNGRQ	3.441	146	152	GCGYLP	1.073
119	125	EEQCLDV	2.871	189	194	QNGRQC	1.193	147	153	CGYLP	1.15
129	135	WIQEGEE	2.443	190	195	NGRQCY	1.079	148	154	GYPGCP	1.1
130	136	IQEGEEG	4.686	198	203	KGNSH	1.994	149	155	YLPGCP	1.1
131	137	QEGEEGR	6.429	206	211	SSEETF	1.602	150	156	LPGCP	1.079
132	138	EGEEGRP	5.871	216	221	RGPMN Q	1.966	163	169	DTFHFLK	1.035
133	139	GEEGRPK	5.571					164	170	TFHFLKC	1.113
134	140	EEGRPKD	6.186					165	171	FHFLKCC	1.184
135	141	EGRPKDD	6.5					166	172	HFLKCCN	1.139
136	142	GRPKDDR	5.986					167	173	FLKCCNT	1.111
137	143	RPKDDRH	5.471					168	174	LKCCNTT	1.085
138	144	PKDDRHL	3.557					169	175	KCCNTTK	1.04
139	145	KDDRHLR	3.857					170	176	CCNTTKC	1.109
140	146	DDRHLRG	3.857					176	182	CNEGPIL	1.054
141	147	DRHLRGC	2.629					178	184	EGPILEL	1.042
151	157	PGCPGSN	4.357					179	185	GPILELE	1.042
152	158	GCPGSNG	4.871					181	187	ILELENL	1.054
153	159	CPGSNGF	2.743					182	188	LELENLP	1.042
154	160	PGSNGFH	2.843					191	197	GRQCYSC	1.108
155	161	GSNGFHN	3.543					192	198	RQCYSC	1.116
156	162	SNGFHNN	3.729					193	199	QCYSCKG	1.117
157	163	NGFHND	4.229					194	200	CYSCKGN	1.082
158	164	GFHNNDT	3.971					201	207	STHGCS	1.048
160	166	HNNDTFH	3.457					209	215	ETFLIDC	1.076
168	174	LKCCNTT	2.386					210	216	TFLIDCR	1.079
169	175	KCCNTTK	4.514					211	217	FLIDCRG	1.074
170	176	CCNTTKC	3.9					212	218	LIDCRGP	1.07
171	177	CNTTKCN	4.7					218	224	PMNQCLV	1.104
172	178	NTTKCNE	5.614					219	225	MNQCLVA	1.104



<b>Table A 4.2 (continued)</b>											
173	179	TTKNEG	5.429					220	226	NQCLVAT	1.116
174	180	TKCNEGP	4.986								
175	181	KCNEGPI	3.1								
185	191	ENLPQNG	3.771								
186	192	NLPQNGR	3.257								
187	193	LPQNGRQ	3.114								
188	194	PQNGRQC	4.629								
189	195	QNGRQCY	4.057								
190	196	NGRQCYS	4.129								
191	197	GRQCYS	3.329								
192	198	RQCYSCK	3.329								
193	199	QCYSCKG	3.543								
194	200	CYSCKGN	3.686								
195	201	YSCKGNS	4.414								
196	202	SCKGNST	5.429								
197	203	CKGNSTH	4.8								
198	204	KGNSTHG	5.414								
199	205	GNSTHGC	4.8								
200	206	NSTHGCS	4.914								
201	207	STHGCS	4.843								
202	208	THGCSSE	5.029								
203	209	HGCSSEE	5.4								
204	210	GCSSEET	5.843								
205	211	CSSEETF	3.714								
214	220	DCRGPMN	3.743								
215	221	CRGPMNQ	3.171								
216	222	RGPMNQC	3.171								
224	230	VATGTHE	3.486								
225	231	ATGTHEP	4.314								
226	232	TGTHEPK	4.829								
227	233	GTHEPKN	5.086								

**Table A 4.3 : Predicted B-cell epitopes using Karpluz & Schulz flexibility and Chou & Fasman beta turns**

**Karpluz & Schulz flexibility** : threshold value 1.003

**Chou & Fasman beta turns** : 1.048

Karplus & Schulz (1.003)				Chou & Fasman (1.048)			
Starting	Ending residue	Peptide	Score	Starting	Ending residue	Peptide	Score
9	15	NGDCRVE	1.013	8	14	TNGDCRV	1.169
17	23	CALGQDL	1.022	9	15	NGDCRVE	1.137
18	24	ALGQDLC	1.042	20	26	GQDLCRT	1.099
19	25	LGQDLCR	1.035	42	48	EKSCTHS	1.101
20	26	GQDLCRT	1.015	43	49	KSCTHSE	1.101
30	36	RLWEEGE	1.024	44	50	SCTHSEK	1.101
31	37	LWEEGEE	1.058	46	52	THSEKTN	1.087
32	38	WEEGEEL	1.072	47	53	HSEKTNR	1.086
33	39	EEGEELE	1.063	48	54	SEKTNRT	1.087
34	40	EGEELEL	1.033	50	56	KTNRTLS	1.066
39	45	ELVEKSC	1.007	51	57	TNRTLSY	1.084
40	46	LVEKSCT	1.018	52	58	NRTLSYR	1.083
41	47	VEKSCTH	1.013	54	60	TLSYRTG	1.084
44	50	SCTHSEK	1.008	56	62	SYRTGLK	1.091
45	51	CTHSEKT	1.032	71	77	CGLDLCN	1.163
46	52	THSEKTN	1.06	72	78	GLDLCNQ	1.133
47	53	HSEKTNR	1.076	73	79	LDLCNQG	1.133
48	54	SEKTNRT	1.08	74	80	DLCNQGN	1.271
49	55	EKTNRTL	1.072	75	81	LCNQGNS	1.267
50	56	KTNRTLS	1.047	76	82	CNQGNQSG	1.406
51	57	TNRTLSY	1.021	77	83	NQGNQSGR	1.371
55	61	LSYRTGL	1.005	78	84	QGNQSGRA	1.243
56	62	<b>SYRTGLK</b>	1.021	79	85	GNSGRAV	1.174
57	63	YRTGLKI	1.026	80	86	NSGRAVT	1.089
58	64	RTGLKIT	1.025	85	91	VTYSRSR	1.051
59	65	TGLKITS	1.028	94	100	ECISCGS	1.144
60	66	GLKITSL	1.023	95	101	CISCGSS	1.243
61	67	LKITSLT	1.026	96	102	ISCGSSD	1.281
62	68	KITSLTE	1.028	97	103	SCGSSDM	1.3
63	69	ITSLTEV	1.017	98	104	CGSSDMS	1.3
64	70	TSLTEVV	1.008	99	105	GSSDMSC	1.3
74	80	DLCNQGN	1.026	100	106	SSDMSCE	1.183
75	81	LCNQGNS	1.077	101	107	SDMSCER	1.114
76	82	CNQGNQSG	1.111	102	108	DMSCERG	1.133
77	83	NQGNQSGR	1.128	103	109	MSCERGR	1.06
78	84	QGNQSGRA	1.119	104	110	SCERGRH	1.11

**Table A 4.3 (continued)**

79	85	GNSGRAV	1.085	106	112	ERGRHQS	1.08
80	86	NSGRAVT	1.039	107	113	RGRHQSL	1.059
85	91	VTYSRSR	1.022	108	114	GRHQSLQ	1.063
96	102	ISCGSSD	1.045	111	117	QSLQCRS	1.079
97	103	SCGSSDM	1.08	112	118	SLQCRSP	1.156
98	104	CGSSDMS	1.082	113	119	LQCRSPE	1.057
99	105	GSSDMSC	1.056	114	120	QCRSPEE	1.079
100	106	SSDMSCE	1.018	115	121	CRSPEEQ	1.079
103	109	MSCERGR	1.015	116	122	RSPEEQC	1.079
104	110	SCERGRH	1.035	132	138	EGEEGRP	1.116
105	111	CERGRHQ	1.049	133	139	GEEGRPK	1.154
106	112	ERGRHQS	1.044	134	140	EEGRPKD	1.14
107	113	RGRHQSL	1.03	135	141	EGRPKDD	1.243
108	114	GRHQSLQ	1.014	136	142	GRPKDDR	1.273
113	119	LQCRSPE	1.027	137	143	RPKDDRH	1.186
114	120	QCRSPEE	1.063	138	144	PKDDRHL	1.134
115	121	CRSPEEQ	1.078	139	145	KDDRHLR	1.053
116	122	RSPEEQC	1.07	140	146	DDRHLRG	1.131
117	123	SPEEQCL	1.046	141	147	DRHLRGC	1.093
118	124	PEEQCLD	1.008	142	148	RHLRGCG	1.107
129	135	WIQEGEE	1.047	143	149	HLRGCGY	1.134
130	136	IQEGEEG	1.088	144	150	LRGCGYL	1.083
131	137	QEGEEGR	1.101	145	151	RGCGYLP	1.216
132	138	EGEEGRP	1.1	146	152	GCGYLPG	1.303
133	139	GEEGRPK	1.094	147	153	CGYLPGC	1.25
134	140	EEGRPKD	1.078	148	154	GYLPGCP	1.297
135	141	EGRPKDD	1.068	149	155	YLPGCPG	1.297
136	142	GRPKDDR	1.059	150	156	LPGCPGS	1.339
137	143	RPKDDRH		151	157	PGCPGSN	1.477
138	144	PKDDRHL	1.026	152	158	GCPGSNG	1.483
139	145	KDDRHLR	1.011	153	159	CPGSNGF	1.346
148	154	GYLPGCP	1.009	154	160	PGSNGFH	1.311
149	155	YLPGCPG	1.029	155	161	GSNGFHN	1.317
150	156	LPGCPGS	1.05	156	162	SNGFHNN	1.317
151	157	PGCPGSN	1.079	157	163	NGFHND	1.321
152	158	GCPGSNG	1.098	158	164	GFHNNDT	1.236
153	159	CPGSNGF	1.099	159	165	FHNNDTF	1.099
154	160	PGSNGFH	1.078	160	166	HNNDTFH	1.149

**Table A 4.3 (continued)**

155	161	GSNGFHN	1.04	161	167	NNDTFHF	1.099
156	162	SNGFHNN	1.01	168	174	LKCCNTT	1.066
158	164	GFHNNDT	1.013	169	175	KCCNTTK	1.126
159	165	FHNNDTF	1.026	170	176	CCNTTKC	1.151
160	166	HNNDTFH	1.02	171	177	CNTTKCN	1.204
170	176	CCNTTKC	1.018	172	178	NTTKCNE	1.14
171	177	CNTTKCN	1.036	173	179	TTKCNEG	1.14
172	178	NTTKCNE	1.043	174	180	TKCNEGP	1.22
173	179	TTKCNEG	1.041	175	181	KCNEGPI	1.15
174	180	TKCNEGP	1.05	176	182	CNEGPIL	1.09
175	181	KCNEGPI	1.055	184	190	LENLPQN	1.077
176	182	CNEGPIL	1.051	185	191	ENLPQNG	1.216
177	183	NEGPIL	1.031	186	192	NLPQNGR	1.246
183	189	ELENLPQ	1.015	187	193	LPQNGRQ	1.163
184	190	LENLPQN	1.039	188	194	PQNGRQC	1.249
185	191	ENLPQNG	1.073	189	195	QNGRQCY	1.194
186	192	NLPQNGR	1.095	190	196	NGRQCYS	1.259
187	193	LPQNGRQ	1.102	191	197	GRQCYS	1.206
188	194	PQNGRQC	1.086	192	198	RQCYSCK	1.127
189	195	QNGRQC Y	1.047	193	199	QCYSCKG	1.214
195	201	YSCKGNS	1.051	194	200	CYSCKGN	1.297
196	202	SCKGNST	1.087	195	201	YSCKGNS	1.331
197	203	CKGNSTH	1.101	196	202	SCKGNST	1.306
198	204	KGNSHGH	1.09	197	203	CKGNSTH	1.237
199	205	GNSTHGC	1.052	198	204	KGNSHGH	1.29
200	206	NSTHGC	1.017	199	205	GNSTHGC	1.316
202	208	THGCSSE	1.012	200	206	NSTHGC	1.297
203	209	HGCSSEE	1.043	201	207	STHGCS	1.279
204	210	GCSSEET	1.074	202	208	THGCSSE	1.18
205	211	CSSEETF	1.077	203	209	HGCSSEE	1.149
206	212	SSEETFL	1.059	204	210	GCSSEET	1.15
207	213	SEETFLI	1.02	212	218	LIDCRGP	1.106
213	219	IDCRGPM	1.018	213	219	IDCRGPM	1.107
214	220	DCRGPM N	1.041	214	220	DCRGPMN	1.263
215	221	CRGPMN Q	1.041	215	221	CRGPMNQ	1.194
216	222	RGPMNQ C	1.019	216	222	RGPMNQC	1.194

				217	223	GPMNQCL	1.143
--	--	--	--	-----	-----	---------	-------

## APPENDIX 5

### Predicted epitopes for the GPNMB antigen

**Table A 5.1 : Predicted linear B-cell epitope peptides using BepiPred 2.0 prediction tool**

Threshold value : 0.500

BepiPred 2.0 (0.500)(3 peaks selected)			
Starting	Ending residue	Peptide	Score
28	75	VLGNERPSAYMREHNQLNGWSSDENDWNEKLYPVWKRGD MRWKN SWKG	48
117	157	EKNCRNEAGLSADPYVYNWTAWSESDGENG TGQSHHNVFP	41
323	372	CPPPPPPRPSKPTPSLATT LKSYDSNTPGPAGDNPLELSRIPDENCQIN	50
400	408	MPVPWPRESS	9
184	204	FQKLGRC SVRVS VNTANVT LG	21
216	224	HGRAYVPIA	9
216	224	HGRAYVPIA	9

**Table A 5.2: Predicted B-cell epitopes using Parker and Emini prediction tools.**

**Parker Hydrophilicity** : threshold value 2.314

**Emini surface accessibility** : threshold value 1.000

Parker (2.314)				Emini (1.000)			
Starting	Ending residue	Peptide	Score	Starting	Ending residue	Peptide	Score
18	24	PLDAAKR	2.429	19	24	LDAAKR	1.253
20	26	DAAKRFH	2.429	20	25	DAAKRF	1.316
21	27	AAKRFHD	2.429	21	26	AAKRFH	1.072
26	32	HDVLGNE	2.814	22	27	AKRFHD	1.772
27	33	DVLGNER	3.114	23	28	KRFHDV	1.302
29	35	LGNERPS	3.443	29	34	LGNERP	1.567
30	36	GNERPSA	5.057	30	35	GNERPS	2.546
31	37	NERPSAY	3.971	31	36	NERPSA	2.599
32	38	ERPSAYM	2.371	32	37	ERPSAY	2.532
34	40	PSAYMRE	2.371	33	38	RPSAYM	1.447
35	41	SAYMREH	2.371	34	39	PSAYMR	1.447
36	42	AYMREHN	2.443	35	40	SAYMRE	1.62
37	43	YMREHNQ	3	36	41	AYMREH	1.645
39	45	REHNQLN	3.557	37	42	YMREHN	2.619
40	46	EHNQLNG	3.771	38	43	MREHNQ	2.895
44	50	LNGWSSD	2.357	39	44	REHNQL	2.412
45	51	NGWSSDE	4.786	40	45	EHNQLN	1.981

**Table A 5.2 (continued)**

46	52	GWSSDEN	4.786	41	46	HNQLNG	1.132
47	53	WSSDEND	5.4	45	50	NGWSSD	1.142
48	54	SSDENDW	5.4	46	51	GWSSDE	1.23
49	55	SDENDWN	5.471	47	52	WSSDEN	1.999
50	56	DENDWNE	5.657	48	53	SSDEND	3.174
51	57	ENDWNEK	5.043	49	54	SDENDW	2.491
52	58	NDWNEKL	2.614	50	55	DENDWN	2.989
65	71	GDMRWKN	2.629	51	56	ENDWNE	3.099
66	72	DMRWKNS	2.743	52	57	NDWNEK	3.579
70	76	KNSWKGG	3.757	53	58	DWNEKL	1.835
71	77	NSWKGGR	3.543	54	59	WNEKLY	1.722
74	80	KGGRVQA	3.671	55	60	NEKLYP	2.533
75	81	GGRVQAV	2.329	56	61	EKLYPV	1.169
79	85	QAVLTSD	2.414	59	64	YPVWKR	1.685
80	86	AVLTSDS	2.486	60	65	PVWKRK	1.065
81	87	VLTSDSP	2.486	61	66	VWKRGD	1.15
82	88	LTSDSPA	3.314	62	67	WKRGD	1.533
83	89	TSDSPAL	3.314	63	68	KRGDMR	2.855
102	108	FPRCQKE	2.571	64	69	RGDMRW	1.501
103	109	PRCQKED	5.314	65	70	GDMRWK	1.533
104	110	RCQKEDA	5.314	66	71	DMRWKN	2.491
105	111	CQKEDAN	5.714	67	72	MRWKNS	1.999
106	112	QKEDANG	6.329	68	73	RWKNSW	2.124
107	113	KEDANGN	6.471	69	74	WKNSWK	2.169
108	114	EDANGNI	4.514	70	75	KNSWKG	2.041
109	115	DANGNIV	2.871	71	76	NSWKGG	1.01
115	121	VYEKNCR	2.929	72	77	SWKGGR	1.23
116	122	YEKNCRN	4.457	74	79	KGGRVQ	1.122
117	123	EKNCRNE	5.843	82	87	LTSDSP	1.256
118	124	KNCRNEA	5.029	83	88	TSDSPA	1.539
119	125	NCRNEAG	5.029	102	107	FPRCQK	1.108
120	126	CRNEAGL	2.714	103	108	PRCQKE	2.216
121	127	RNEAGLS	3.443	104	109	RCQKED	2.393
122	128	NEAGLSA	3.143	105	110	CQKEDA	1.234
123	129	EAGLSAD	3.571	106	111	QKEDAN	3.703
124	130	AGLSADP	2.757	107	112	KEDANG	2.116
136	142	TAWSEDS	4.014	108	113	EDANGN	1.702
137	143	AWSESDS	4.7	113	118	NIVYEK	1.033

**Table A 5.2 (continued)**

138	144	WSESDSG	5.214	114	119	IVYEKN	1.033
139	145	SESDGE	7.757	116	121	YEKNCR	2.085
140	146	EDSDGEN	7.829	117	122	EKNCRN	2.14
141	147	DSDGENG	7.529	118	123	KNCRNE	2.14
142	148	SDGENGT	6.843	119	124	NCRNEA	1.081
143	149	DGENGTG	6.729	121	126	RNEAGL	1.023
144	150	GENGTGQ	6.157	126	131	LSADPY	1.028
145	151	ENGTGQS	6.271	128	133	ADPYVY	1.082
146	152	NGTGQSH	5.457	129	134	DPYVYN	1.722
147	153	GTGQSHH	4.757	130	135	PYVYNW	1.084
148	154	TGQSHHN	4.943	131	136	YVYNWT	1.012
149	155	GQSHHNV	3.671	136	141	TAWSED	1.352
154	160	NVFPDGK	2.514	137	142	AWSEDS	1.256
157	163	PDGKPPF	2.643	138	143	WSESDS	2.075
158	164	DGKPPPH	2.643	139	144	SESDSG	1.953
185	191	QKLGRC	2.9	140	145	EDSDGE	2.524
193	199	RVSNTA	2.514	141	146	DSDGEN	2.344
194	200	VSVNTAN	2.914	142	147	SDGENG	1.389
195	201	SVNTANV	2.914	143	148	DGENGT	1.496
196	202	VNTANVT	2.729	145	150	ENGTGQ	1.551
213	219	YRRHGRA	2.943	146	151	NGTGQS	1.2
214	220	RRHGRAY	2.943	147	152	GTGQSH	1.016
241	247	TMFQKND	2.929	148	153	TGQSHH	1.397
242	248	MFQKNDR	2.786	149	154	GQSHHN	1.556
243	249	FQKNDRN	4.386	150	155	QSHHNV	1.167
244	250	QKNDRNS	6.629	156	161	FPDGKP	1.557
245	251	KNDRNSS	6.7	157	162	PDGKPF	1.557
246	252	NDRNSSD	7.314	158	163	DGKPPF	1.557
247	253	DRNSSDE	7.429	159	164	GKPPPH	1.269
248	254	RNSSDET	6.743	160	165	KPPPHH	1.745
249	255	NSSDETF	4.829	161	166	PPPHHP	1.349
250	256	SSDETF	2.514	163	168	PHHPGW	1.048
251	257	SDETF	2.4	164	169	HPPGWR	1.328
252	258	DETF	2.9	165	170	HPGWRR	1.911
284	290	SFGDNTG	4.414	166	171	PGWRRW	1.477
293	299	VSTNHTV	2.657	167	172	GWRRWN	1.536
294	300	STNHTVN	4.186	168	173	WRRWNF	1.344
295	301	TNHTVNH	3.557	181	186	GQYFQK	1.833



**Table A 5.2 (continued)**

296	302	NHTVNHT	3.557	182	187	QYFQKL	1.527
316	322	KAAAPGP	3.129	184	189	FQKLGR	1.091
317	323	AAAPGPC	2.514	210	215	VTVYRR	1.088
318	324	AAPGPCP	2.514	211	216	TVYRRH	1.994
319	325	APGPCPP	2.514	212	217	VYRRHG	1.367
320	326	PGPCPPP	2.514	213	218	YRRHGR	3.608
321	327	GPCPPPP	2.514	214	219	RRHGRA	2.326
325	331	PPPPPR	2.4	215	220	RHGRAY	1.861
326	332	PPPPRP	2.4	225	230	QVKDVY	1.136
327	333	PPPRPS	3.029	241	246	TMFQKN	1.568
328	334	PPRPSK	3.543	242	247	MFQKND	1.814
329	335	PPRPSKP	3.543	243	248	FQKNDR	3.59
330	336	PRPSKPT	3.986	244	249	QKNDRN	6.667
331	337	RPSKPTP	3.986	245	250	KNDRNS	5.159
332	338	PSKPTPS	4.314	246	251	NDRNSS	3.457
333	339	SKPTPSL	2.7	247	252	DRNSSD	3.59
336	342	TPSLATT	2.443	248	253	RNSSDE	3.723
341	347	TTLKSYD	3.071	249	254	NSSDET	2.743
342	348	TLKSYDS	3.257	250	255	SSDETF	1.477
343	349	LKSYDSN	3.514	252	257	DETFLK	1.356
344	350	KSYDSNT	5.571	253	258	ETFLKD	1.356
345	351	SYDSNTP	5.057	267	272	IHDPSH	1.022
346	352	YDSNTPG	4.943	268	273	HDPSHF	1.263
347	353	DSNTPGP	5.514	275	280	NYSTIN	1.25
348	354	SNTPGPA	4.386	276	281	YSTINY	1.218
349	355	NTPGPAG	4.271	277	282	STINYK	1.555
350	356	TPGPAGD	4.7	278	283	TINYKW	1.22
351	357	PGPAGDN	4.957	279	284	INYKWS	1.133
352	358	GPAGDNP	4.957	280	285	NYKWSF	1.399
353	359	PAGDNPL	2.829	284	289	SFGDNT	1.013
354	360	AGDNPLE	3.643	293	298	VSTNHT	1.032
362	368	SRIPDEN	4.229	294	299	STNHTV	1.032
363	369	RIPDENC	3.5	295	300	TNHTVN	1.238
364	370	IPDENCQ	3.757	296	301	NHTVNH	1.167
365	371	PDENCQI	3.757	297	302	HTVNHT	1.048
366	372	DENCQIN	4.457	298	303	TVNHTY	1.206
367	373	ENCQINR	3.629	322	327	PCPPPP	1.078
413	419	VVTCQGS	2.486	323	328	CPPPPP	1.078

**Table A 5.2 (continued)**

415	421	TCQGSIP	2.7	324	329	PPPPPP	3.111
416	422	CQGSIPT	2.7	325	330	PPPPPP	3.111
417	423	QGSIPTE	3.614	326	331	PPPPPR	3.94
428	434	ISDPTCE	3.571	327	332	PPPPRP	3.94
429	435	SDPTCEI	3.571	328	333	PPPRPS	3.415
430	436	DPTCEIT	3.386	329	334	PPRPSK	4.416
431	437	PTCEITQ	2.814	330	335	PRPSKP	4.416
432	438	TCEITQN	3.514	331	336	RPSKPT	4.122
433	439	CEITQNT	3.514	332	337	PSKPTP	3.254
434	440	EITQNTV	2.786	333	338	SKPTPS	2.82
436	442	TQNTVCS	3.943	334	339	KPTPSL	1.736
437	443	QNTVCSP	3.5	340	345	ATTLKS	1.058
439	445	TVCSPVD	2.543	341	346	TTLKSY	1.641
441	447	CSPVDVD	3.229	342	347	TLKSYD	1.899
442	448	SPVDVDE	4.143	343	348	LKSYDS	1.764
443	449	PVDVDEM	2.614	344	349	KSYDSN	3.439
444	450	VDVDEMC	2.514	345	350	SYDSNT	2.482
455	461	RRTFNGS	3.371	346	351	YDSNTP	2.864
456	462	RTFNGSG	3.586	347	352	DSNTPG	1.809
457	463	TFNGSGT	3.729	348	353	SNTPGP	1.675
458	464	FNGSGTY	2.714	349	354	NTPGPA	1.262
459	465	NGSGTYC	4.229	353	358	PAGDNP	1.461
460	466	GSGTYCV	2.7	355	360	GDNPLE	1.336
461	467	SGTYCVN	2.886	356	361	DNPLEL	1.113
467	473	NLTGDD	2.786	358	363	PLELSR	1.088
468	474	LTLGDDT	2.529	362	367	SRIPDE	1.872
469	475	TLGDDTS	4.771	363	368	RIPDEN	2.247
470	476	LGDDTSL	2.714	365	370	PDENCQ	1.519
471	477	GDDTSLA	4.329	370	375	QINRYG	1.349
483	489	ISVPDRD	3.014	371	376	INRYGH	1.06
484	490	SVPDRDP	4.457	372	377	NRYGHF	1.31
485	491	VPDRDPA	3.829	373	378	RYGHFQ	1.41
486	492	PDRDPAS	5.286	401	406	PVPWPE	1.137
487	493	DRDPASP	5.286	403	408	PWPESS	1.779
488	494	RDPASPL	2.543	429	434	SDPTCE	1.055
489	495	DPASPLR	2.543	434	439	EITQNT	1.603
				452	457	LTVRRT	1.113
				453	458	TVRRTF	1.169

				454	459	VRRTFN	1.302
				455	460	RRTFNG	1.736
				456	461	RTFNGS	1.188
				459	464	NGSGTY	1.086
				469	474	TLGDDT	1.079
				470	475	LGDDTS	1.002
				471	476	GDDTSL	1.002
				472	477	DDTSLA	1.023
				484	489	SVPDRD	1.912
				485	490	VPDRDP	2.206
				486	491	PDRDPA	3.003
				487	492	DRDPAS	2.602
				488	493	RDPASP	2.409
				489	494	DPASPL	1.015
				490	495	PASPLR	1.19

**Table A 5.3 : Predicted B-cell epitopes using Karpluz & Schulz flexibility and Chou & Fasman beta turns**

**Karpluz & Schulz flexibility** : threshold value 1.003

**Chou & Fasman beta turns** : threshold value 1.048

**Kolaskar and Tongaokar antigenicity scale** : threshold value 1.033

Karplus & Schulz (1.003)				Chou & Fasman (1.048)				Kolaskar and tongaonkar (1.033)			
Starting	Ending residue	Peptide	Score	Starting	Ending residue	Peptide	Score	Starting	Ending residue	Peptide	Score
27	33	DVLGNER	1.026	26	32	HDVLGNE	1.051	1	7	MECLYYF	1.107
28	34	VLGNERP	1.054	27	33	DVLGNER	1.051	2	8	ECLYYFL	1.168
29	35	LGNERPS	1.065	28	34	VLGNERP	1.06	3	9	CLYYFLG	1.171
30	36	GNERPSA	1.057	29	35	LGNERPS	1.193	4	10	LYYFLGF	1.125
31	37	NERPSAY	1.032	30	36	GNERPSA	1.203	5	11	YYFLGFL	1.125
39	45	REHNQLN	1.006	31	37	NERPSAY	1.143	6	12	YFLGFLL	1.138
40	46	EHNQLNG	1.005	40	46	EHNQLNG	1.134	7	13	FLGFLLL	1.151
43	49	QLNGWSS	1.012	41	47	HNQLNGW	1.166	8	14	LGFLLLA	1.147
44	50	LNGWSSD	1.02	42	48	NQLNGWS	1.234	9	15	GFLLLAA	1.12
45	51	NGWSSDE	1.045	43	49	QLNGWSS	1.216	10	16	FLLLAAR	1.12
46	52	GWSSDEN	1.07	44	50	LNGWSSD	1.284	11	17	LLLAARL	1.143
47	53	WSSDEND	1.079	45	51	NGWSSDE	1.306	12	18	LLAARLP	1.116

**Table A 5.3 (continued)**

48	54	SSDENDW	1.082	46	52	GWSSDEN	1.306	13	19	LAARLPL	1.116
49	55	SDENDWN	1.069	47	53	WSSDEND	1.291	14	20	AARLPLD	1.062
50	56	DENDWNE	1.051	48	54	SSDENDW	1.291	15	21	ARLPLDA	1.062
51	57	ENDWNEK	1.036	49	55	SDENDWN	1.31	16	22	RLPLDAA	1.062
52	58	NDWNEKL	1.031	50	56	DENDWNE	1.211	17	23	LPLDAAK	1.07
53	59	DWNEKLY	1.027	51	57	ENDWNEK	1.147	22	28	AKRFHDV	1.045
54	60	WNEKLYP	1.014	52	58	NDWNEKL	1.126	23	29	KRFHDVL	1.071
60	66	PVWKRGD	1.022	53	59	DWNEKLY	1.066	24	30	RFHDVLG	1.063
61	67	VWKRGD	1.044	54	60	WNEKLYP	1.074	25	31	FHDVLGN	1.049
62	68	WKRGD	1.052	59	65	YPVWKR	1.091	55	61	NEKLYPV	1.059
63	69	KRGDMRW	1.032	60	66	PVWKRGD	1.137	56	62	EKLYPVW	1.076
67	73	MRWKNSW	1.014	62	68	WKRGD	1.07	57	63	KLYPVWK	1.087
68	74	RWKNSWK	1.034	63	69	KRGDMRW	1.07	58	64	LYPVWKR	1.079
69	75	WKNSWKG	1.045	64	70	RGDMRWK	1.07	75	81	GGRVQAV	1.067
70	76	KNSWKGG	1.052	65	71	GDMRWKN	1.157	76	82	GRVQAVL	1.12
71	77	NSWKGGR	1.066	66	72	DMRWKNS	1.139	77	83	RVQAVLT	1.125
72	78	SWKGGRV	1.073	67	73	MRWKNSW	1.067	78	84	VQAVLTS	1.145
73	79	WKGGRVQ	1.064	68	74	RWKNSWK	1.126	79	85	QAVLTSD	1.071
74	80	KGGRVQA	1.035	69	75	WKNSWKG	1.213	80	86	AVLTSDS	1.071
80	86	AVLTSDS	1.038	70	76	KNSWKGG	1.299	81	87	VLTSDSP	1.071
81	87	VLTSDSP	1.074	71	77	NSWKGGR	1.29	84	90	SDSPALV	1.093
82	88	LTS DSPA	1.083	72	78	SWKGGRV	1.139	85	91	DSPALVG	1.073
83	89	TSDSPAL	1.077	73	79	WKGGRVQ	1.074	86	92	SPALVGS	1.094
84	90	SDSPALV	1.043	81	87	VLTSDSP	1.127	87	93	PALVGSN	1.06
85	91	DSPALVG	1.004	82	88	LTS DSPA	1.15	88	94	ALVGSNI	1.073
88	94	ALVGSNI	1.016	83	89	TSDSPAL	1.15	89	95	LVGSNIT	1.051
89	95	LVGSNIT	1.033	84	90	SDSPALV	1.084	92	98	SNITFAV	1.055
90	96	VGSNITF	1.021	85	91	DSPALVG	1.103	94	100	ITFAVNL	1.089
101	107	IFPRCQK	1.006	86	92	SPALVGS	1.099	95	101	TFAVNLI	1.089
102	108	FPRCQKE	1.023	87	93	PALVGSN	1.117	96	102	FAVNLI	1.115
103	109	PRCQKED	1.043	103	109	PRCQKED	1.121	97	103	AVNLIFP	1.111
104	110	RCQKEDA	1.057	105	111	CQKEDAN	1.086	98	104	VNLIFPR	1.084
105	111	CQKEDAN	1.064	106	112	QKEDANG	1.139	99	105	NLIFPRC	1.088
106	112	QKEDANG	1.071	107	113	KEDANGN	1.221	100	106	LIFPRCQ	1.122
107	113	KEDANGN	1.064	108	114	EDANGNI	1.144	101	107	IFPRCQK	1.077
108	114	EDANGNI	1.047	109	115	DANGNIV	1.11	102	108	FPRCQKE	1.034
109	115	DANGNIV	1.029	110	116	ANGNIVY	1.064	114	120	IVYEKNC	1.095
115	121	VYEKNCR	1.014	111	117	NGNIVYE	1.076	115	121	VYEKNCR	1.055

**Table A 5.3 (continued)**

116	122	YEKNCRN	1.022	116	122	YEKNCRN	1.164	125	131	GLSADPY	1.042
117	123	EKNCRNE	1.027	117	123	EKNCRNE	1.107	126	132	LSADPYV	1.114
118	124	KNCRNEA	1.035	118	124	KNCRNEA	1.096	127	133	SADPYVY	1.102
119	125	NCRNEAG	1.037	119	125	NCRNEAG	1.174	128	134	ADPYVYN	1.068
120	126	CRNEAGL	1.028	121	127	RNEAGLS	1.07	129	135	DPYVYNW	1.043
121	127	RNEAGLS	1.01	124	130	AGLSADP	1.126	130	136	PYVYNWT	1.05
125	131	GLSADPY	1.01	125	131	GLSADPY	1.194	131	137	YVYNWTA	1.05
126	132	LSADPYV	1.02	127	133	SADPYVY	1.121	149	155	GQSHHNV	1.039
127	133	SADPYVY	1.01	128	134	ADPYVYN	1.14	150	156	QSHHNVF	1.07
136	142	TAWSEDS	1.018	129	135	DPYVYNW	1.183	151	157	SHHNVFP	1.077
137	143	AWSESDS	1.048	130	136	PYVYNWT	1.111	152	158	HHNVFPD	1.056
138	144	WSESDG	1.071	133	139	YNWTAWS	1.096	155	161	VFPDGKP	1.039
139	145	SESDGGE	1.093	136	142	TAWSEDS	1.091	159	165	GKPFPHH	1.033
140	146	EDSDGEN	1.096	137	143	AWSESDS	1.163	160	166	KPFPHHP	1.06
141	147	DSDGENG	1.106	138	144	WSESDG	1.291	161	167	PFPHHPG	1.052
142	148	SDGENGT	1.109	139	145	SESDGGE	1.26	170	176	RWNFIYV	1.047
143	149	DGENGTG	1.111	140	146	EDSDGEN	1.279	171	177	WNFIYVF	1.078
144	150	GENGTGQ	1.12	141	147	DSDGENG	1.396	172	178	NFIYVFH	1.108
145	151	ENGTGQS	1.116	142	148	SDGENGT	1.324	173	179	FIYVFHT	1.127
146	152	NGTGQSH	1.108	143	149	DGENGTG	1.343	174	180	IYVFHTL	1.15
147	153	GTGQSHH	1.086	144	150	GENGTGQ	1.274	175	181	YVFHTLG	1.11
148	154	TGQSHHN	1.046	145	151	ENGTGQS	1.256	176	182	VFHTLGQ	1.09
149	155	GQSHHNV	1.004	146	152	NGTGQSH	1.286	177	183	FHTLGQY	1.058
154	160	NVFPDGK	1.021	147	153	GTGQSHH	1.199	178	184	HTLGQYF	1.058
155	161	VFPDGKP	1.052	148	154	TGQSHHN	1.199	179	185	TLGQYFQ	1.045
156	162	FPDGKPF	1.071	149	155	GQSHHNV	1.133	180	186	LGQYFQK	1.048
157	163	PDGKPPF	1.062	151	157	SHHNVFP	1.073	181	187	GQYFQKL	1.048
158	164	DGKPFPH	1.034	152	158	HHNVFPD	1.077	182	188	QYFQKLG	1.048
164	170	HHPGWRR	1.009	153	159	HNVFPDG	1.164	184	190	FQKLGRC	1.064
183	189	YFQKLGR	1.011	154	160	NVFPDGK	1.173	185	191	QKLGRC	1.052
184	190	FQKLGRC	1.014	155	161	VFPDGKP	1.167	186	192	KLGRCSV	1.105
185	191	QKLGRC	1.012	156	162	FPDGKPF	1.181	187	193	LGRCVSR	1.097
195	201	SVNTANV	1.005	157	163	PDGKPPF	1.313	188	194	GRCSVSRV	1.116
201	207	VTLGPQL	1.006	158	164	DGKPFPH	1.231	189	195	RCSVRSV	1.135
202	208	TLGPQLM	1.005	159	165	GKPFPHH	1.159	190	196	CSVRSVSV	1.208
212	218	VYRRHGR	1.003	160	166	KPFPHHP	1.153	191	197	SVRSVSN	1.117
213	219	YRRHGRA	1.008	161	167	PFPHHPG	1.231	192	198	VRVSVNT	1.103
231	237	VVTDQIP	1.004	162	168	FPHHPGW	1.151	193	199	RVSVNTA	1.057

**Table A 5.3 (continued)**

241	247	TMFQKND	1.003	163	169	PHHPGWR	1.201	194	200	VSVNTAN	1.043
242	248	MFQKNDR	1.037	164	170	HHPGWRR	1.12	195	201	SVNTANV	1.043
243	249	FQKNDRN	1.063	165	171	HPGWRRW	1.121	199	205	ANVTLGP	1.046
244	250	QKNDRNS	1.076	166	172	PGWRRWN	1.209	200	206	NVTLGPQ	1.039
245	251	KNDRNSS	1.09	167	173	GWRRWNF	1.077	201	207	VTLGPQL	1.106
246	252	NDRNSSD	1.103	185	191	QKLGRC	1.101	204	210	GPQLMEV	1.038
247	253	DRNSSDE	1.113	199	205	ANVTLGP	1.05	205	211	PQLMEVT	1.043
248	254	RNSSDET	1.112	200	206	NVTLGPQ	1.096	206	212	QLMEVTV	1.088
249	255	NSSDETF	1.087	243	249	FQKNDRN	1.16	207	213	LMEVTVY	1.109
250	256	SSDETF	1.057	244	250	QKNDRNS	1.279	208	214	MEVTVYR	1.055
251	257	SDETF	1.029	245	251	KNDRNSS	1.343	209	215	EVTVYRR	1.062
252	258	DETFLKD	1.01	246	252	NDRNSSD	1.407	210	216	VTVYRRH	1.098
253	259	ETFLKDL	1.01	247	253	DRNSSDE	1.29	215	221	RHGRAYV	1.048
254	260	TFLKDLP	1.014	248	254	RNSSDET	1.219	216	222	HGRAYVP	1.075
255	261	FLKDLPI	1.006	249	255	NSSDETF	1.169	217	223	GRAYVPI	1.082
266	272	LIHDPSH	1.006	266	272	LIHDPSH	1.053	218	224	RAYVPIA	1.109
267	273	IHDPSHF	1.014	267	273	IHDPSHF	1.054	219	225	AYVPIAQ	1.129
268	274	HDPHF	1.008	268	274	HDPHF	1.071	220	226	YVPIAQV	1.175
283	289	WSFGDNT	1.008	269	275	DPSHFLN	1.159	221	227	VPIAQVK	1.142
284	290	SFGDNTG	1.036	270	276	PSHFLNY	1.113	222	228	PIAQVKD	1.068
285	291	FGDNTGL	1.053	271	277	SHFLNYS	1.1	223	229	IAQVKDV	1.113
286	292	GDNTGLF	1.045	274	280	LNYS	1.101	224	230	AQVKDVY	1.115
287	293	DNTGLFV	1.018	275	281	NYSTINY	1.18	225	231	QVKDVYV	1.16
292	298	FVSTNHT	1.008	276	282	YSTINYK	1.101	226	232	VKDVYVV	1.213
293	299	VSTNHTV	1.004	277	283	STINYKW	1.076	227	233	KDVYVVT	1.145
303	309	YVLNGTF	1.008	278	284	TINYKWS	1.076	228	234	DVYVVT	1.136
304	310	VLNGTFS	1.022	280	286	NYKWSFG	1.18	229	235	VYVVT	1.157
305	311	LNGTFS	1.013	281	287	YKWSFGD	1.166	230	236	YVVT	1.124
317	323	AAAPGPC	1.029	282	288	KWSFGDN	1.226	231	237	VVT	1.11
318	324	AAPGPCP	1.054	283	289	WSFGDNT	1.219	232	238	VTDQIPV	1.11
319	325	APGPCPP	1.062	284	290	SFGDNTG	1.304	233	239	TDQIPVF	1.069
320	326	PGPCPPP	1.055	285	291	FGDNTGL	1.184	234	240	DQIPVFV	1.136
321	327	GPCPPPP	1.054	286	292	GDNTGLF	1.184	235	241	QIPVFVT	1.142
322	328	PCPPPPP	1.051	294	300	STNHTVN	1.131	236	242	IPVFVTM	1.115
323	329	CPPPPPP	1.053	295	301	TNHTVNH	1.063	237	243	PVFVTMF	1.107
324	330	PPPPPPP	1.057	296	302	NHTVNHT	1.063	238	244	VFVTMFQ	1.1
325	331	PPPPPPR	1.054	306	312	NGTFS	1.18	239	245	FVTMFQK	1.035
326	332	PPPPPRP	1.053	316	322	KAAAPGP	1.084	254	260	TFLKDLP	1.051

**Table A 5.3 (continued)**

327	333	PPPPRPS	1.059	317	323	AAAPGPC	1.11	255	261	FLKDLPI	1.086
328	334	PPPRPSK	1.067	318	324	AAPGPCP	1.233	256	262	LKDLPIIM	1.048
329	335	PPRPSKP	1.078	319	325	APGPCPP	1.356	259	265	LPIMFDV	1.09
330	336	PRPSKPT	1.09	320	326	PGPCPPP	1.479	260	266	PIMFDVL	1.09
331	337	RPSKPTP	1.087	321	327	GPCPPPP	1.479	261	267	IMFDVLI	1.103
332	338	PSKPTPS	1.079	322	328	PCPPPPP	1.473	262	268	MFDVLIH	1.096
333	339	SKPTPSL	1.065	323	329	CPPPPPP	1.473	263	269	FDVLIHD	1.102
334	340	KPTPSLA	1.041	324	330	PPPPPPP	1.52	264	270	DVLIHDP	1.098
335	341	PTPSLAT	1.024	325	331	PPPPPPR	1.439	265	271	VLIHDPS	1.119
336	342	TPSLATT	1.01	326	332	PPPPPRP	1.439	266	272	LIHDPSH	1.079
338	344	SLATTLK	1.011	327	333	PPPPRPS	1.426	267	273	IHDPSHF	1.056
339	345	LATTLKS	1.02	328	334	PPPRPSK	1.353	268	274	HDPSHFL	1.07
340	346	ATTLKSY	1.023	329	335	PPRPSKP	1.353	270	276	PSHFLNY	1.066
341	347	TTLKSYD	1.034	330	336	PRPSKPT	1.273	271	277	SHFLNYS	1.058
342	348	TLKSYDS	1.041	331	337	RPSKPTP	1.273	272	278	HFLNYST	1.043
343	349	LKSYDSN	1.053	332	338	PSKPTPS	1.341	273	279	FLNYSTI	1.05
344	350	KSYDSNT	1.077	333	339	SKPTPSL	1.209	288	294	NTGLFVS	1.042
345	351	SYDSNTP	1.092	334	340	KPTPSLA	1.099	289	295	TGLFVST	1.061
346	352	YDSNTPG	1.099	335	341	PTPSLAT	1.091	290	296	GLFVSTN	1.042
347	353	DSNTPGP	1.1	341	347	TTLKSYD	1.079	291	297	LFVSTNH	1.075
348	354	SNTPGPA	1.092	342	348	TLKSYDS	1.146	293	299	VSTNHTV	1.068
349	355	NTPGPAG	1.086	343	349	LKSYDSN	1.231	297	303	HTVNHTY	1.05
350	356	TPGPAGD	1.074	344	350	KSYDSNT	1.284	298	304	TVNHTYV	1.089
351	357	PGPAGDN	1.067	345	351	SYDSNTP	1.357	299	305	VNHTYVL	1.138
352	358	GPAGDNP	1.064	346	352	YDSNTPG	1.376	300	306	NHTYVLN	1.051
353	359	PAGDNPL	1.057	347	353	DSNTPGP	1.43	301	307	HTYVLNG	1.065
354	360	AGDNPLE	1.051	348	354	SNTPGPA	1.316	302	308	TYVLNGT	1.037
355	361	GDNPLEL	1.033	349	355	NTPGPAG	1.334	303	309	YVLNGTF	1.063
356	362	DNPLELS	1.011	350	356	TPGPAGD	1.32	304	310	VLNGTFS	1.042
359	365	LELSRIP	1.006	351	357	PGPAGDN	1.406	309	315	FSLNLTV	1.096
360	366	ELSRIPD	1.021	352	358	GPAGDNP	1.406	310	316	SLNLTVK	1.073
361	367	LSRIPDE	1.036	353	359	PAGDNPL	1.267	311	317	LNLTVKA	1.08
362	368	SRIPDEN	1.046	354	360	AGDNPLE	1.156	312	318	NLTVKAA	1.054
363	369	RIPDENC	1.042	355	361	GDNPLEL	1.146	313	319	LTVKAAA	1.095
364	370	IPDENCQ	1.029	356	362	DNPLELS	1.127	314	320	TVKAAAP	1.068
365	371	PDENCQI	1.007	357	363	NPLELSR	1.054	315	321	VKAAAPG	1.063
402	408	VPWPESS	1.039	362	368	SRIPDEN	1.161	317	323	AAAPGPC	1.087
403	409	PWPESSL	1.061	363	369	RIPDENC	1.127	318	324	AAPGPCP	1.087

**Table A 5.3 (continued)**

404	410	WPESLI	1.063	364	370	IPDENCQ	1.131	319	325	APGPCPP	1.087
405	411	PESSLID	1.036	365	371	PDENCQI	1.131	320	326	PGPCPPP	1.087
414	420	VTCQGS	1.019	366	372	DENCQIN	1.137	321	327	GPCPPPP	1.087
415	421	TCQGSIP	1.046	367	373	ENCQINR	1.064	322	328	PCPPPPP	1.114
416	422	CQGSIPT	1.055	368	374	NCQINRY	1.121	323	329	CPPPPPP	1.114
417	423	QGSIPTE	1.056	369	375	CQINRYG	1.121	324	330	PPPPPPP	1.064
418	424	GSIPTEV	1.05	370	376	QINRYGH	1.087	325	331	PPPPPPR	1.037
419	425	SIPTEVC	1.027	372	378	NRYGHFQ	1.106	326	332	PPPPPRP	1.037
427	433	IISDPTC	1.011	400	406	MPVPWPE	1.051	333	339	SKPTPSL	1.034
428	434	ISDPTCE	1.02	401	407	PVPWPES	1.17	334	340	KPTPSLA	1.042
429	435	SDPTCEI	1.014	402	408	VPWPES	1.157	335	341	PTPSLAT	1.039
431	437	PTCEITQ	1.003	403	409	PWPESL	1.17	337	343	PSLATTL	1.065
432	438	TCEITQN	1.024	405	411	PESSLID	1.091	338	344	SLATTLK	1.046
433	439	CEITQNT	1.051	415	421	TCQGSIP	1.159	339	345	LATTLKS	1.046
434	440	EITQNTV	1.071	416	422	CQGSIPT	1.159	340	346	ATTLKSY	1.034
435	441	ITQNTVC	1.059	417	423	QGSIPTE	1.094	358	364	PLELSRI	1.065
436	442	TQNTVCS	1.033	425	431	CTIISDP	1.071	359	365	LELSRIP	1.065
437	443	QNTVCSP	1.005	427	433	IISDPTC	1.071	369	375	CQINRYG	1.038
442	448	SPVDVDE	1.003	428	434	ISDPTCE	1.11	376	382	HFQATIT	1.035
453	459	TVRRFTN	1.01	429	435	SDPTCEI	1.11	377	383	FQATITI	1.042
454	460	VRRTFNG	1.019	436	442	TQNTVCS	1.083	378	384	QATITIV	1.083
455	461	RRTFNCS	1.028	437	443	QNTVCSP	1.163	379	385	ATITIVE	1.06
456	462	RTFNCSG	1.054	438	444	NTVCSPV	1.094	380	386	TITIVEG	1.033
457	463	TFNCSGT	1.084	439	445	TVCSPVD	1.08	381	387	ITIVEGI	1.068
458	464	FNCSGTY	1.094	441	447	CSPVDVD	1.151	382	388	TIVEGIL	1.082
459	465	NGSGTYC	1.081	442	448	SPVDVDE	1.087	383	389	IVEGILE	1.073
460	466	GSGTYCV	1.042	455	461	RRTFNCS	1.144	384	390	VEGILEV	1.106
468	474	LTLGDDT	1.009	456	462	RTFNCSG	1.231	386	392	GILEVNI	1.063
469	475	TLGDDTS	1.026	457	463	TFNCSGT	1.233	387	393	ILEVNII	1.102
470	476	LGDDTSL	1.035	458	464	FNCSGTY	1.259	388	394	LEVNIQ	1.083
471	477	GDDTSLA	1.029	459	465	NGSGTYC	1.343	392	398	IIQMTDV	1.043
472	478	DDTSLAL	1.009	460	466	GSGTYCV	1.191	393	399	IQMTDVL	1.057
476	482	LALTSTL	1.025	461	467	SGTYCVN	1.191	396	402	TDVLPVP	1.097
477	483	ALTSTLI	1.038	462	468	GTVCVNL	1.071	397	403	DVLPVPV	1.119
478	484	LTSTLIS	1.023	467	473	NLTGDD	1.169	398	404	VLPVPVP	1.123
483	489	ISVPDRD	1.014	468	474	LTLGDDT	1.083	399	405	LMPVPVP	1.078
484	490	SVPDRDP	1.034	469	475	TLGDDTS	1.203	401	407	PVPWPES	1.047
485	491	VPDRDPA	1.045	470	476	LGDDTSL	1.15	402	408	VPWPES	1.04



**Table A 5.3 (continued)**

486	492	PDRDPAS	1.049	471	477	GDDTSLA	1.16	404	410	WPESLI	1.033
487	493	DRDPASP	1.053	483	489	ISVPDRD	1.113	406	412	ESSLIDF	1.033
488	494	RDPASPL	1.05	484	490	SVPDRDP	1.263	407	413	SSLIDFV	1.109
489	495	DPASPLR	1.038	485	491	VPDRDPA	1.153	408	414	SLIDFVV	1.162
				486	492	PDRDPAS	1.286	409	415	LIDFVVT	1.148
				487	493	DRDPASP	1.286	410	416	IDFVVTTC	1.171
				488	494	RDPASPL	1.161	411	417	DFVVTTCQ	1.151
				489	495	DPASPLR	1.161	412	418	FVVTTCQG	1.152
								413	419	VVTCQGS	1.141
								414	420	VTCQGS	1.108
								415	421	TCQGSIP	1.063
								416	422	CQGSIPT	1.063
								418	424	GSIPTEV	1.035
								419	425	SIPTEVC	1.112
								420	426	IPTEVCT	1.097
								421	427	PTEVCTI	1.097
								422	428	TEVCTII	1.11
								423	429	EVCTIIS	1.124
								424	430	VCTIISD	1.127
								425	431	CTIISDP	1.081
								427	433	IISDPTC	1.081
								428	434	ISDPTCE	1.038
								429	435	SDPTCEI	1.038
								431	437	PTCEITQ	1.045
								435	441	ITQNTVC	1.079
								436	442	TQNTVCS	1.059
								437	443	QNTVCSP	1.082
								438	444	NTVCSPV	1.134
								439	445	TVCSPVD	1.147
								440	446	VCSPVDV	1.215
								441	447	CSPVDVD	1.141
								442	448	SPVDVDE	1.061
								443	449	PVDVDEM	1.034
								444	450	VDVDEMC	1.084
								445	451	DVDEMCL	1.065
								446	452	VDEMCLL	1.12
								447	453	DEMCLLT	1.052
								448	454	EMCLLTV	1.126

**Table A 5.3 (continued)**

								449	455	MCLLTVR	1.129
								450	456	CLLTVRR	1.136
								451	457	LLTVRRT	1.064
								452	458	LTVRRTF	1.041
								460	466	GSGTYCV	1.089
								461	467	SGTYCVN	1.075
								462	468	GTVCVNL	1.109
								463	469	TYCVNLT	1.114
								464	470	YCVNLT	1.163
								465	471	CVNLTG	1.122
								466	472	VNLTG	1.044
								473	479	DTSLALT	1.037
								474	480	TSLALTS	1.058
								475	481	SLALTST	1.058
								476	482	LALTSTL	1.092
								477	483	ALTSTLI	1.078
								478	484	LTSTLIS	1.071
								479	485	TSTLISV	1.09
								480	486	STLISVP	1.112
								481	487	TLISVPD	1.091
								482	488	LISVPDR	1.086

## APPENDIX 6

### Results of Galaxy Refine

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	2.989	4.4	12.6	56.5
MODEL 1	0.8869	0.571	2.054	8.6	0.8	88.4
MODEL 2	0.8818	0.599	1.977	7.2	0.4	88.7

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	2.989	4.4	12.6	56.5
MODEL 1	0.8869	0.571	2.054	8.6	0.8	88.4
MODEL 2	0.8818	0.599	1.977	7.2	0.4	88.7

**Initial** : refers to the I-TASSER predicted tertiary structure (used as input file on Galaxy refine)

**MODEL 2** : refers to the refined model with the most suitable scores. This model was selected for further analysis and simulation studies.

**Figure A 6.1: Results from GalaxyRefine for refinement of the tertiary structure of the senovaccine construct.**

## LIST OF PUBLICATIONS

[1] M.Goja , M.F.H. Shahanshah and A. Das. “Harnessing cellular senescence and oncolytic viruses as unconventional cancer immunotherapeutics”. **(submitted to Elsevier Life Sciences, under process)**.

[2] M.F.H Shahanshah, M.Goja and A. Das. “B cell multiepitope senovaccine for healthy aging and tackling age associated pathologies”. **(under process for publication)**

## REFERENCES

- [1] S. M. Lagoumtzi and N. Chondrogianni, “Senolytics and senomorphics: Natural and synthetic therapeutics in the treatment of aging and chronic diseases,” *Free Radic Biol Med*, vol. 1, pp. 169-190, August 2021.
- [2] M. K. Ruhland and E. Alspach, “Senescence and immunoregulation in the tumor microenvironment,” *Frontiers in Cell and Developmental Biology*, vol. 9, 2021.
- [3] T. Saleh, S. Bloukh, V. J. Carpenter, E. Alwohoush, J. Bakeer, S. Darwish, B. Azab, and D. A. Gewirtz, “Therapy-induced senescence: An ‘old’ friend becomes the enemy,” *Cancers*, vol. 12, no. 4, p. 822, 2020.
- [4] C. Amor, J. Feucht, J. Leibold, Y.-J. Ho, C. Zhu, D. Alonso-Curbelo, J. Mansilla-Soto, J. A. Boyer, X. Li, T. Giavridis, A. Kulick, S. Houlihan, E. Peerschke, S. L. Friedman, V. Ponomarev, A. Piersigilli, M. Sadelain, and S. W. Lowe, “Senolytic car T cells reverse senescence-associated pathologies,” *Nature*, vol. 583, no. 7814, pp. 127–132, 2020.
- [5] “Ageing and health,” *World Health Organization*. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=By%202050%2C%20the%20world's%20population,will%20double%20\(2.1%20billion.](https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=By%202050%2C%20the%20world's%20population,will%20double%20(2.1%20billion.) [Accessed: 08-Mar-2023].
- [6] Schmeer, Kretz, Wengerodt, Stojiljkovic, and Witte, “Dissecting aging and senescence—current concepts and open lessons,” *Cells*, vol. 8, no. 11, p. 1446, 2019. doi:10.3390/cells8111446
- [7] L. Wyld et al., “Senescence and cancer: A review of clinical implications of senescence and Senotherapies,” *Cancers*, vol. 12, no. 8, p. 2134, 2020. doi:10.3390/cancers12082134
- [8] L. Hayflick and P. S. Moorhead, “The serial cultivation of human diploid cell strains,” *Experimental Cell Research*, vol. 25, no. 3, pp. 585–621, 1961. doi:10.1016/0014-4827(61)90192-6
- [9] H. Wang et al., “Senolytics (DQ) mitigates radiation ulcers by removing senescent cells,” *Frontiers in Oncology*, vol. 9, 2020. doi:10.3389/fonc.2019.01576
- [10] B. G. Childs et al., “Senescent intimal foam cells are deleterious at all stages of atherosclerosis,” *Science*, vol. 354, no. 6311, pp. 472–477, 2016. doi:10.1126/science.aaf6659
- [11] F. A. Valentijn, L. L. Falke, T. Q. Nguyen, and R. Goldschmeding, “Cellular senescence in the aging and diseased kidney,” *Journal of Cell Communication and Signaling*, vol. 12, no. 1, pp. 69–82, 2017. doi:10.1007/s12079-017-0434-2

- [12] V. Schirmacher, “From chemotherapy to biological therapy: A review of novel concepts to reduce the side effects of systemic cancer treatment (review),” *International Journal of Oncology*, vol. 54, no. 2, pp. 407–419, 2018. doi:10.3892/ijo.2018.4661
- [13] M. Kritsilis et al., “Ageing, cellular senescence and neurodegenerative disease,” *International Journal of Molecular Sciences*, vol. 19, no. 10, p. 2937, 2018. doi:10.3390/ijms19102937
- [14] L. Wang, L. Lankhorst, and R. Bernards, “Exploiting senescence for the treatment of cancer,” *Nature Reviews Cancer*, vol. 22, no. 6, pp. 340–355, 2022.
- [15] Y. Huang and T. Liu, “Step further towards targeted senolytic therapy: Therapeutic potential of UPAR-car T cells for senescence-related diseases,” *Signal Transduction and Targeted Therapy*, vol. 5, no. 1, 2020.
- [16] A. L. Cardoso *et al.*, “Towards frailty biomarkers: Candidates from genes and pathways regulated in aging and age-related diseases,” *Ageing Research Reviews*, vol. 47, pp. 214–277, 2018. doi:10.1016/j.arr.2018.07.004
- [17] H. Burris *et al.*, “A phase (PH) I/II study of CR011-VcMMAE, an antibody-drug conjugate, in patients (PTS) with locally advanced or metastatic breast cancer (MBC).,” *Cancer Research*, vol. 69, no. 24\_Supplement, pp. 6096–6096, 2009. doi:10.1158/0008-5472.sabcs-09-6096
- [18] A. A. N. Rose *et al.*, “Glycoprotein nonmetastatic B is an independent prognostic indicator of recurrence and a novel therapeutic target in breast cancer,” *Clinical Cancer Research*, vol. 16, no. 7, pp. 2147–2156, 2010. doi:10.1158/1078-0432.ccr-09-1611
- [19] A. A. N. Rose *et al.*, “Osteoactivin promotes breast cancer metastasis to Bone,” *Molecular Cancer Research*, vol. 5, no. 10, pp. 1001–1014, 2007. doi:10.1158/1541-7786.mcr-07-0119
- [20] G. I. Mun, S. J. Lee, S. M. An, I. K. Kim, and Y. C. Boo, “Differential gene expression in young and senescent endothelial cells under static and laminar shear stress conditions,” *Free Radical Biology and Medicine*, vol. 47, no. 3, pp. 291–299, 2009. doi:10.1016/j.freeradbiomed.2009.04.032
- [21] H.-L. Jong, M. R. Mustafa, P. M. Vanhoutte, S. AbuBakar, and P.-F. Wong, “MicroRNA 299-3P modulates replicative senescence in endothelial cells,” *Physiological Genomics*, vol. 45, no. 7, pp. 256–267, 2013. doi:10.1152/physiolgenomics.00071.2012
- [22] M. C. Boonstra, H. W. Verspaget, S. Ganesh, F. J.G.M. Kubben, A. L. Vahrmeijer, C. J.H. van de Velde, P. J.K. Kuppen, P. H.A. Quax, and C. F.M. Sier, “Clinical applications of the urokinase receptor (UPAR) for cancer patients,” *Current Pharmaceutical Design*, vol. 17, no. 19, pp. 1890–1910, 2011.
- [23] M. Suda, I. Shimizu, G. Katsuomi, Y. Yoshida, Y. Hayashi, R. Ikegami, N. Matsumoto, Y. Yoshida, R. Mikawa, A. Katayama, J. Wada, M. Seki, Y. Suzuki, A. Iwama, H. Nakagami, A.

- Nagasawa, R. Morishita, M. Sugimoto, S. Okuda, M. Tsuchida, K. Ozaki, M. Nakanishi-Matsui, and T. Minamino, "Senolytic vaccination improves normal and pathological age-related phenotypes and increases lifespan in progeroid mice," *Nature Aging*, vol. 1, no. 12, pp. 1117–1126, 2021.
- [24] "Uniprot: The Universal Protein Knowledgebase in 2021," *Nucleic Acids Research*, vol. 49, no. D1, 2020.
- [25] A. Krogh, B. Larsson, G. von Heijne, and E. L. L. Sonnhammer, "Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes," *Journal of Molecular Biology*, vol. 305, no. 3, pp. 567–580, 2001.
- [26] S. Möller, M. D. Croning, and R. Apweiler, "Evaluation of methods for the prediction of membrane spanning regions," *Bioinformatics*, vol. 17, no. 7, pp. 646–653, 2001.
- [27] "Free epitope database and prediction resource," *IEDB.org: Free epitope database and prediction resource*. [Online]. Available: <https://www.iedb.org/>. [Accessed: 08-Mar-2023].
- [28] M. C. Jespersen, B. Peters, M. Nielsen, and P. Marcatili, "BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes," *Nucleic Acids Research*, vol. 45, no. W1, 2017.
- [29] P. Y. Chou and G. D. Fasman, "Prediction of the secondary structure of proteins from their amino acid sequence," *Advances in Enzymology - and Related Areas of Molecular Biology*, pp. 45–148, 2006.
- [30] E. A. Emini, J. V. Hughes, D. S. Perlow, and J. Boger, "Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide," *Journal of Virology*, vol. 55, no. 3, pp. 836–839, 1985.
- [31] P. A. Karplus and G. E. Schulz, "Prediction of chain flexibility in proteins," *Naturwissenschaften*, vol. 72, no. 4, pp. 212–213, 1985.
- [32] J. M. Parker, D. Guo, and R. S. Hodges, "New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites," *Biochemistry*, vol. 25, no. 19, pp. 5425–5432, 1986.
- [33] A. S. Kolaskar and P. C. Tongaonkar, "A semi-empirical method for prediction of antigenic determinants on protein antigens," *FEBS Letters*, vol. 276, no. 1-2, pp. 172–174, 1990.
- [34] I. A. Doytchinova and D. R. Flower, "VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines," *BMC Bioinformatics*, vol. 8, no. 1, 2007.
- [35] I. A. Doytchinova and D. R. Flower, "Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties," *Vaccine*, vol. 25, no. 5, pp. 856–866, 2007.

- [36] I. Dimitrov, L. Naneva, I. Doytchinova, and I. Bangov, “AllergenFP: Allergenicity prediction by Descriptor Fingerprints,” *Bioinformatics*, vol. 30, no. 6, pp. 846–851, 2013.
- [37] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, and G. P. Raghava, “In silico approach for predicting toxicity of peptides and proteins,” *PLoS ONE*, vol. 8, no. 9, 2013.
- [38] W. L. DeLano, “The PyMOL molecular graphics system,” *PyMOL*. [Online]. Available: <https://pymol.org/2/>. [Accessed: 24-Apr-2023].
- [39] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch, “Protein identification and analysis tools on the expasy server,” *The Proteomics Protocols Handbook*, pp. 571–607, 2005.
- [40] M. Hebditch, M. A. Carballo-Amador, S. Charonis, R. Curtis, and J. Warwicker, “Protein-sol: A web tool for predicting protein solubility from sequence,” *Bioinformatics*, vol. 33, no. 19, pp. 3098–3100, 2017.
- [41] L. J. McGuffin, K. Bryson, and D. T. Jones, “The PSIPRED protein structure prediction server,” *Bioinformatics*, vol. 16, no. 4, pp. 404–405, 2000. doi:10.1093/bioinformatics/16.4.404
- [42] C. Geourjon and G. Deléage, “SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments,” *Bioinformatics*, vol. 11, no. 6, pp. 681–684, 1995. doi:10.1093/bioinformatics/11.6.681
- [43] X. Zhou *et al.*, “I-tasser-MTD: A deep-learning-based platform for multi-domain protein structure and function prediction,” *Nature Protocols*, vol. 17, no. 10, pp. 2326–2353, 2022. doi:10.1038/s41596-022-00728-0
- [44] W. Zheng *et al.*, “Folding non-homologous proteins by coupling deep-learning contact maps with I-Tasser Assembly simulations,” *Cell Reports Methods*, vol. 1, no. 3, p. 100014, 2021. doi:10.1016/j.crmeth.2021.100014
- [45] J. Yang and Y. Zhang, “I-Tasser Server: New development for protein structure and function predictions,” *Nucleic Acids Research*, vol. 43, no. W1, 2015. doi:10.1093/nar/gkv342
- [46] C. Seok *et al.*, “Accurate protein structure prediction: What comes next?,” *BIODESIGN*, vol. 9, no. 3, pp. 47–50, 2021. doi:10.34184/kssb.2021.9.3.47
- [47] J. Ko, H. Park, L. Heo, and C. Seok, “GalaxyWEB server for protein structure prediction and refinement,” *Nucleic Acids Research*, vol. 40, no. W1, 2012. doi:10.1093/nar/gks493
- [48] C. J. Williams *et al.*, “Molprobit: More and better reference data for improved all-atom structure validation,” *Protein Science*, vol. 27, no. 1, pp. 293–315, 2017. doi:10.1002/pro.3330



- [49] M. Wiederstein and M. J. Sippl, “Prosa-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins,” *Nucleic Acids Research*, vol. 35, no. Web Server, 2007. doi:10.1093/nar/gkm290
- [50] M. J. Sippl, “Recognition of errors in three-dimensional structures of proteins,” *Proteins: Structure, Function, and Genetics*, vol. 17, no. 4, pp. 355–362, 1993. doi:10.1002/prot.340170404
- [51] X. Xu, Y. Cai, Y. Wei, F. Donate, J. Juarez, G. Parry, L. Chen, E. J. Meehan, R. W. Ahn, A. Ugolkov, O. Dubrovskiy, T. V. O’Halloran, M. Huang, and A. P. Mazar, “Identification of a new epitope in Upar as a target for the cancer therapeutic monoclonal antibody ATN-658, a structural homolog of the UPAR Binding Integrin CD11B ( $\alpha$ m),” *PLoS ONE*, vol. 9, no. 1, 2014.
- [52] D. Kozakov *et al.*, “The ClusPro web server for protein–protein docking,” *Nature Protocols*, vol. 12, no. 2, pp. 255–278, 2017. doi:10.1038/nprot.2016.169
- [53] Y. Li, G. Parry, L. Chen, J. A. Callahan, D. E. Shaw, E. J. Meehan, A. P. Mazar, and M. Huang, “An anti-urokinase plasminogen activator receptor (upar) antibody: Crystal Structure and binding epitope,” *Journal of Molecular Biology*, vol. 365, no. 4, pp. 1117–1129, 2007.
- [54] R. A. Laskowski, J. Jabłońska, L. Pravda, R. S. Vařeková, and J. M. Thornton, “PDBsum: Structural summaries of PDB entries,” *Protein Science*, vol. 27, no. 1, pp. 129–134, 2017. doi:10.1002/pro.3289
- [55] E. Liberis, P. Veličković, P. Sormanni, M. Vendruscolo, and P. Liò, “Parapred: Antibody paratope prediction using convolutional and recurrent neural networks,” *Bioinformatics*, vol. 34, no. 17, pp. 2944–2950, 2018.
- [56] J. R. López-Blanco, J. I. Garzón, and P. Chacón, “IMOD: Multipurpose normal mode analysis in internal coordinates,” *Bioinformatics*, vol. 27, no. 20, pp. 2843–2850, 2011. doi:10.1093/bioinformatics/btr497
- [57] Back-translation, [http://www.geneinfinity.org/sms/sms\\_backtranslation.html](http://www.geneinfinity.org/sms/sms_backtranslation.html) (accessed May 10, 2023).
- [58] Rare codon analysis tool - genscript, <https://www.genscript.com/tools/rare-codon-analysis> (accessed May 10, 2023).
- [59] Nebcutter 3.0, <https://nc3.neb.com/NEBcutter/> (accessed May 10, 2023).
- [60] “SnapGene,” SnapGene Software, <https://www.snapgene.com/> (accessed May 10, 2023).
- [61] Y. Lei, F. Zhao, J. Shao, Y. Li, S. Li, H. Chang, and Y. Zhang, “Application of built-in adjuvants for epitope-based vaccines,” *PeerJ*, vol. 6, 2019.
- [62] R. C. Sanches, S. Tiwari, L. C. Ferreira, F. M. Oliveira, M. D. Lopes, M. J. Passos, E. H. Maia, A. G. Taranto, R. Kato, V. A. Azevedo, and D. O. Lopes, “Immunoinformatics design

of multi-epitope peptide-based vaccine against schistosoma mansoni using transmembrane proteins as a target,” *Frontiers in Immunology*, vol. 12, 2021.

PAPER NAME

**final.pdf**

---

WORD COUNT

**18407 Words**

CHARACTER COUNT

**87550 Characters**

PAGE COUNT

**74 Pages**

FILE SIZE

**5.7MB**

SUBMISSION DATE

**May 27, 2023 11:32 PM GMT+5:30**

REPORT DATE

**May 27, 2023 11:33 PM GMT+5:30**

---

**● 10% Overall Similarity**

The combined total of all matches, including overlapping sources, for each database.

- 7% Internet database
- Crossref database
- 7% Submitted Works database
- 6% Publications database
- Crossref Posted Content database

**● Excluded from Similarity Report**

- Bibliographic material

DELHI TECHNOLOGICAL UNIVERSITY  
(Formerly Delhi College of Engineering)  
Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Mohd Fardeen Husain Shahanshah, 2K21/MSCBIO/27 of M.Sc. Biotechnology, hereby declare that the project Dissertation titled "B cell multiepitope senovaccine for healthy aging and tackling age associated pathologies" which is submitted by me to the Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Science, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

*Fardeen*

Place: Delhi

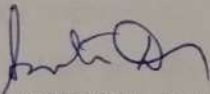
**MOHD FARDEEN HUSAIN SHAHANSHAH**

Date: 30.05.2023

DEPARTMENT OF BIOTECHNOLOGY  
DELHI TECHNOLOGICAL UNIVERSITY  
(Formerly Delhi College of Engineering)  
Bawana Road, Delhi-110042

CERTIFICATE

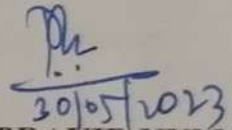
I hereby certify that the Project Dissertation titled "B cell multi epitope senovaccine for healthy aging and tackling age associated pathologies" which is submitted by Mohd Fardeen Husain Shahanshah, 2K21/MSCBIO/27, Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Science, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.



**DR. ASMITA DAS**

**SUPERVISOR**

Department of Biotechnology  
Delhi Technological University



**PROF. PRAVIR KUMAR**

**HEAD OF THE DEPARTMENT**

Department of Biotechnology  
Delhi Technological University