2022
9(2), 295

Chemical Biology
LETTERS

Reverse vaccinology based *in silico* analysis of Epitope prediction in *cya*, *lef* and *pagA* genes from *Bacillus anthracis* against Anthrax infected species: An Immunoinformatics approach

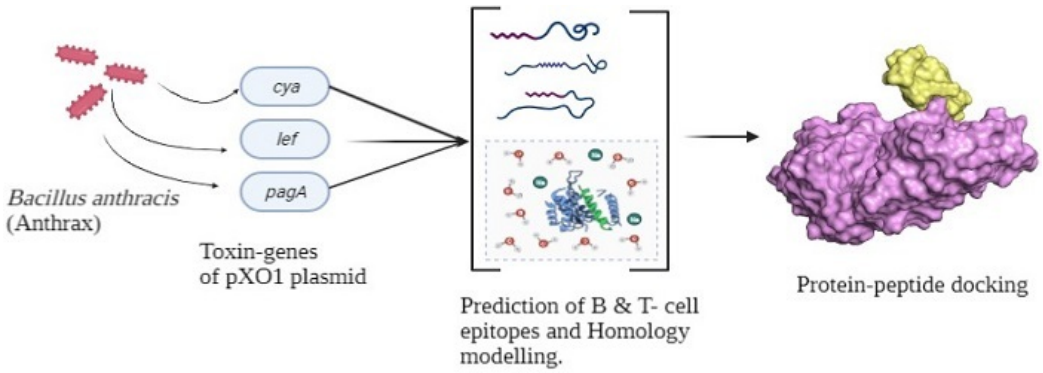
Uma Bharathi Indrabalan,[1] Kuralayanapalya Puttahonnappa Suresh,[1] Mallikarjun S Beelagi,[1] Sharanagouda S Patil,[1] Chandan Shivamallu,[2] Mohan Pappana,[3] Raghavendra G Amachawadi[4]*

[1]ICAR-National Institute of Veterinary Epidemiology and Disease Informatics (NIVEDI), Yelahanka, Bengaluru-560064, India. [2]Department of Biotechnology and Bioinformatics, Faculty of Life Sciences, JSS Academy of Higher Education & Research, Mysuru 570015, India. [3]Huck Institutes of the Life Sciences, 205 Wartik Laboratory, The Pennsylvania State University, University Park, State College, PA 16802, USA. [4]Department of Clinical Sciences, Kansas State University, Manhattan, KS 66506-5606, USA.

Submitted on: 12-Nov-2021, Accepted and Published on: 27-Feb-2022

Article

ABSTRACT



Bacillus anthracis is a Gram-positive spore-forming bacterium that causes the zoonotic disease: anthrax, an abrupt illness that disproportionately impacts grazing livestock and wild ruminants. The anthrax's geographical reach despite years of research on anthrax epizootic and epidemics behaviour, till date remains to be elucidated. Existing therapeutics, however, are ineffective in combating this infectious disease, necessitating the development of a better vaccine to halt the pandemic using immunoinformatics approaches, this study intended to predict an efficient epitope for vaccine against the anthrax in animals and humans of the toxin genes such as cya, lef and pagA of B.anthracis against anthrax. The B-cell and T-cell epitopes were predicted utilizing various bioinformatics tools/software and docking analysis was performed. Consequently, it was found that the evaluated epitopes had no allegenicity, no toxicity and had high antigenicity that provides an effectual and most rapid technique to estimate peptide synthetic vaccines to impede the anthrax.

Keywords: Google Anthrax, cya, lef, pag, B-cell and T-cell epitopes.

INTRODUCTION

In the first decade of the twenty-first century, bioterrorism was a contentious issue. Biological agents are appealing weapons for bioterrorism because they are easy to get, relatively inexpensive to implement and cause broad fear and uncertainty instead of severe physical harm.[1] Anthrax is a zoonotic and epizootic illness disseminated by spore transfer through ingestion, inhalation, or an open skin wound in domestic animals. Humans who make contact with diseased livestock or their contaminated soil are also at risk.[2–4] Grazing in vegetated regions where prior anthrax outbreaks have occurred is the major source of infection in domestic livestock. Ingestion of spores found in soils, plants, or water causes germination, which is followed by rapid replication of vegetative cells and the production of exotoxins, resulting in septicemia and mortality[5–7]

Genes located on two major plasmids comprise the primary virulence components of *Bacillus anthracis*. Genes identified on the 184.5kbp plasmid, pXO1, code for toxin production, while genes identified on the 95.3kbp plasmid, pXO2, control the development of a poly—D-glutamic acid capsules.[8,9] Plasmid pXO1 has genes that code for toxins, while plasmid pXO2 contains genes that code for capsules. pXO1 contains three genes encoding: *pagA* (PA), *lef* (lethal factor), and *cya* (edema factor). The geneatxA, which is expressed on plasmid pXO1, affects the expression of genes encoded on pXO1 and pXO2.[1,4,9] These three

*Corresponding Author: Raghavendra G Amachawadi, Assistant Professor, Food Animal Therapeutics, 1800 Denison Avenue, Q221 Mosier Hall, Department of Clinical Sciences, College of Veterinary Medicine , Kansas State University, Manhattan, Kansas 66506-5606
Tel: 785-532-4356
Email: agraghav@vet.k-state.edu

URN:NBN:sciencein.cbl.2022.v9.295
© ScienceIn Publishing ISSN: 2347–9825
https://pubs.thesciencein.org/cbl

genes are innocuous on their own, but when they combine in the blood or even on the membrane of host cells, they create the anthrax toxin, which belongs to the AB toxin family of toxins.[10,11]

The existing authorised human vaccines, anthrax vaccine precipitated and anthrax vaccine adsorbed or BioThrax, both are made up of screened culture supernatant from cell lines of pXO1 and pXO2 *B. anthracis* isolates that contain varying levels of anthrax toxins.[12] The development of a new vaccine for this new emerging strain employing preventive and therapeutic techniques can indeed be quickly implemented to save lives. Due to developments in designing, stabilization, and distribution, using peptides or epitopes as treatments is a valuable approach.[13,14]

Epitopes are antigenic determinants that are categorized into B-cell i.e., continuous and discontinuous and T-cell i.e., major histocompatibility complex I (MHC-I) and major histocompatibility complex II (MHC-II). Antigen B- and T-cell epitopes can be identified and predicted using computational methods in order to build recombinant vaccines that are significant in antibody stimulation.[13–16]

Vaccination is an important technique for improving public health and providing an effective way to reduce the spread of illnesses. Plants operate as bioreactors in nature and have been exploited to produce effective vaccine antigens against viral, bacterial, and protozoan diseases. Furthermore, we know that one of the critical processes in vaccine design is the prediction of antibody epitopes using computational methods.[17,18] The use of computational methods considerably contributes to biology by building in silico vaccines and predicting T-cell epitopes, which further reduces costs as well as the need for experimental observations.[19] Immunoinformatics techniques have enabled significant advances towards the design of vaccine candidates by predicting B-cell and T-cell epitopes.[20–22]

Epitope/peptide based vaccines have several advantages over traditional vaccines, and several are already in clinical trials. These next-generation vaccines development feature a large specificity for eliciting immunological responses, a huge production volume, as well as a high efficiency. Besides such benefits, the capacity to provide high doses of potential immunogen at a low cost is the main advantage of epitope- or peptide-based vaccines. Antigenicity and pathogenicity are the main factors that must be present in the viral protein that could be considered as a vaccine candidate. Furthermore, epitope-based peptides are simple to produce, purify, preserve and manipulate. Epitope-based vaccines are commonly considered to be better than regular vaccines .[16,23]

Epitopic or peptide vaccines against HIV, malaria, and tuberculosis produced promising results and preserved the developed vaccine candidates' defensive and therapeutic potential. Immunoinformatics serves an important role in the production of antibodies and antigen detection agents, as well as vaccine design. Vaccines as well as other drug carrier research in the initial stages was solely based on immunological trials compared to the earlier procedures developed were time-consuming and expensive.[24,25]

Subunit vaccines are made up of pathogen protein or glycoprotein constituents capable of activating a protective immune response and can be made using biochemical or recombinant DNA techniques. Recombinant subunit vaccines have distinct advantages over live attenuated and inactivated vaccines in that they are more effective at activating humoral and cell-mediated immune responses, and they eliminate the risks related with pathogen. A subunit vaccine is a cost-efficient and efficient way to avoid health issues. They have minimal side effects than live-attenuated vaccinations, but they still require additives to maximise their effectiveness.[26,27]

Though a study has already been conducted on anthrax epitope prediction of *B. anthracis* but had focused only on a single gene lethal factor (*lef)* [28] and protective antigen (*pag*).[29] These studies had concentrated considering only humans as a host and in these studies docking analysis was employed to check the interactions between the human receptor and the obtained epitopes. The purpose of this study is to use the immunoinformatics approach to find the most vital epitopes for *cya*, *lef* and *pagA* for humans and bovine against anthrax. The advanced bioinformatics software/tools will be used to explore in silico drug design and immunoinformatics methodologies.[30,31] With the use of immunoinformatic techniques, B-cell and T-cell epitopes as potential vaccine candidates will be attempted to identify for both humans and animals in the current study. Molecular modelling and docking methods have been used to explore the post-docking interactions of peptide-MHC complexes in order to identify viable candidates for peptide vaccine development.

MATERIALS AND METHODS

Protein Sequence Retrieval:

The protein sequences of *B. anthracis*: *cya*, *lef* and pagA in pXO1 were retrieved from the Genbank database (https://www.ncbi.nlm.nih.gov/protein/) individually for both human and bovine. Antigenicity prediction was used for these sequences in order to evaluate the availability of antigens. The protein sequences with VaxiJen v.2.0.[32] values greater than the cutoff level 0.4 were being used for further study. The ExPASy server's ProtParam tool (https://web.expasy.org/protparam/) was being used to evaluate the physical and chemical characteristics of the selected proteins.[33]

Linear B-Cell Epitope Prediction:

B-lymphocyte cells differentiate into memory cells and antibody-secreting plasma cells after interacting with antigens such as B-cell epitopes. B-cell epitopes are required for inducing a humoral immune response, which activates B cells for immunogenicity and plays a vital role in vaccine design[34] B-cell epitopes from cya, lef, and pagA genes of humans and animals had been identified utilizing BCPred tool[35] (http://ailabprojects1.ist.psu.edu:8080/bcpred/predict.html).

Epitope prediction for Cytotoxic T Cells:

The prediction of Cytotoxic T Lymphocyte (CTL) epitopes is essential for the manufacturing of the subunit vaccinations[36] Further, for the identifying of promiscuity T cell epitopes, numerous alternative prediction approaches are being used.

MHC-I epitopes are those that are identified by more than one MHC allele and recognised by much more than a T cell type[37] So, for the prediction of CTL epitopes, the amino acid sequence was examined using the IEDB tool[38] using the prediction method as NetMHCpan EL 4.1 Epitopes associated with distinct MHC-I alleles of higher affinity (IC50) were chosen using NetCTL and SMM-based IEDB MHC-I binding prediction algorithms, MHC-I source species as human alleles (HLA-A*01:01 and HLA-A*02:01) and cow alleles (BoLA-1:00901 and BoLA-2:00501) respectively and the epitope was chosen to be of 14bp in length.

Epitope prediction for Helper T Cells:

The prediction of Helper T Lymphocyte was utilised to analyse peptide binding to MHC-II molecules, and since they detect MHC-II peptides obtained from extracellular protein, the HTL plays an important role in triggering both cellular and humoral immune responses. As a result, HTL epitopes play an important role in the development of immunotherapy vaccines.[39] The NetCTL.1.2 (http://www.cbs.dtu.dk/services/NetCTL/) server was utilized for prediction with MHC-II source species as human alleles (DRB1*0101, DRB1*0401 and DRB1*1501) and bovine alleles (BoLA-DRB3*0101 and BoLA-DRB3*0201) respectively.

Prediction of allergenicity, antigenicity and toxicity:

The webserver AllerTop v.2.0 (https://www.ddg-pharmfac.net/AllerTOP/index.html) was used to predict allergic and nonallergenic behaviours of the obtained epitopes. The predicting of allergenicity of known peptide sequences is based on similarity. The webserver VaxiJen v.2.0. (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html)to determine whether the epitopes are antigenic or not, and the ToxinPred server to predict the toxicity of the obtained epitopes.(https://webs.iiitd.edu.in/raghava/toxinpred/protein.php)

Modelling of Epitopes and Molecular Docking:

To assess the binding effectiveness of the obtained epitopes that are bound to the Human Leukocyte Antigen (HLA), molecular docking studies were carried out and also to predict the interaction between the epitope and antigen, hence an efficient vaccination must interact effectively with the immunological receptors of the host to elicit better immune system responses. To summarise, chosen HTL and CTL epitopes was presented to the PEP-FOLD v.3.0 server (https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/)[40] The better peptide models were then docked to the selected class MHC-I and MHC-II molecules with their respective alleles obtained from RCSB PDB server using the PatchDock docking server (https://bioinfo3d.cs.tau.ac.il/PatchDock/). The PatchDock server's algorithm uses molecular geometry to find docking modifications with high molecular shape similarity.[41,42] The structures that resulted were refined using the FireDock server (https://bioinfo3d.cs.tau.ac.il/FireDock/).[43] Interaction analysis and molecular dynamics simulations were performed on high energy compounds. The energy minimization was carried out with the UCSF Chimera tool,[44] the interactions were determined with Discovery studio.[45] The editing and visualization were obtained with the PyMol tool.[46]

RESULTS

Protein Sequence Retrieval:

A total of 6 protein sequences of *B. anthracis*: *cya*, *lef* and *pagA* from humans and animals were downloaded in FASTA format. The sequence retrieval accession numbers and other details are provided with the accession numbers for the retrieved sequences as shown in Table.1. All the obtained sequences were subjected to antigenicity, allergenicity and toxicity prediction analysis. Also, these sequences were determined for their chemical and physical properties (Table S1).

Table 1: List of the protein accessions from Genbank -NCBI.

Genes	Accession No. Bovine	Accession No. Human	length
cya	WP_000197748.1	AJG68144.1	800
lef	WP_001022097.1	AJG68004.1	806
pagA	WP_000746486.1	AJG68118.1	760

Linear B-Cell Epitope Prediction:

The linear B-cell epitopes from *cya*, *lef*, and *pagA* were predicted using BCPred with requirements specified to have 75% specificity, epitopes of length 20mer and also, the included epitopes were nonoverlapping. Higher scores indicate that peptides are more easily recognised by B cells, implying that they have been more likely to be epitopes. A total of 17, 15 and 15 epitopes were predicted from *cya*, *lef* and *pagA* sequences respectively. The top ten predicted B-cell epitopes of *cya*, *lef* and *pagA* are shown in Table 2.

Epitope prediction for Cytotoxic T Cells:

To determine the T-cell epitope indicated interacting with various types of MHC class I alleles, the envelope protein from *cya*, *lef* and *pagA* for human alleles (HLA-A*01:01 and HLA-A*02:01) and cow alleles (BoLA-1:00901 and BoLA-2:00501) respectively were examined using the IEDB MHC-I interaction prediction tool with an epitope of length 14mer (Table S2). A total of 1577 epitopes in cya and lef, 1505 epitopes in pagA were predicted using the software. The scores having higher values were considered for further analysis.

Epitope prediction for Helper T Cells:

The HTL epitopes of 14mer were determined based on their IC50 value of 50 nM, minimum percentile rank value, and best prediction score, indicating high epitope binding affinity with various DRB1 and BoLA DRB3 allele MHC-II variants. The list of all alleles investigated during the investigation is included in the HTL recognition that majorly focused and predicted MHC-II source of human alleles (DRB1*0101, DRB1*0401 and DRB1*1501) and bovine alleles (BoLA-DRB3*0101 and BoLA-DRB3*0201) respectively (Table S3).

Prediction of allergenicity, antigenicity and toxicity analysis:

The obtained B-cell and T-cell epitopes were subjected to the allergenicity, antigenicity and toxicity analysis. The Best scored

epitopes were tested and the epitope for each gene for human and bovine was identified and listed (Table 3, 4 and 5). The listed epitopes have higher antigenicity and were also analyzed as non-allergens and non-toxic epitopes, which is absolute for designing/developing vaccines.

Table 2: the predicted B-cell epitopes of the three genes from the BCpred tool

cya		
Position	**Epitope**	**Score**
44	KRNHKTEKNKTEKEKFKDSI	0.997
559	RLNEAVKYTGYTGGDVVNHG	0.994
641	AYIEWTDPITKAKINTIPTS	0.988
537	GIERKPDSTKGTLSNWQKQM	0.985
708	ANHIFSQEKKRKISIFRGIQ	0.976
444	FRISDENNEVQYKTKEGKIT	0.975
414	KENGIILKGKKEIDNGKKYY	0.975
507	QIPQKEWDKVVNTPNSLEKQ	0.974
675	NVGVYKDSGDKDEFAKKESV	0.972
346	KGLNVHGKSSDWGPVAGYIP	0.971
lef		
Position	**Epitope**	**Score**
23	SGPVFIPLVQGAGGHGDVGM	1
52	DENKRKDEERNKTQEEHLKE	0.999
342	HSLSQEEKELLKRIQIDSSD	0.999
74	KHIVKIEVKGEEAVKKEAAE	0.996
372	KKLQIDIRDSLSEEEKELLN	0.995
754	EGSNLTSYGRTNEAEFFAEA	0.961
579	DAKVVPKSKIDTKIQEAQLN	0.96
307	IKQHYQHWSDSLSEEGRGLL	0.951
778	HSTDHAERLKVQKNAPKTFQ	0.94
393	IQVDSSNPLSEKEKEFLKKL	0.934
pagA		
Position	**Epitope**	**Score**
731	TKENTIINPSENGDTSTNGI	0.998
56	LNFQAPMVVTSSTTGDLSIP	0.997
196	RSTSAGPTVPDRDNDGIPDS	0.995
240	HEKKGLTKYKSSPEKWSTAS	0.995
535	AAVNPSDPLETTKPDMTLKE	0.981
317	TISKNTSTSRTHTSEVHGNA	0.971
28	QAEVKQENRLLNESESSSQG	0.961
141	RLYQIKIQYQRENPTEKGLD	0.948
640	VVKEAHREVINSSTEGLLLN	0.931
93	IWSGFIKVKKSDEYTFATSA	0.925

Table 3: List of B-cell epitopes that are antigenic, non-allergenic and toxic of the three genes

B-cell				
Genes	**start**	**end**	**Epitope**	**score**
cya	44	63	KRNHKTEKNKTEKEKFKDSI	0.997
lef	23	42	SGPVFIPLVQGAGGHGDVGM	1
pagA	196	215	RSTSAGPTVPDRDNDGIPDS	0.995

Table 4: List of T-cell epitopes that are antigenic, non-allergenic and toxic of the three genes in human alleles (HLA-A*01:01 and HLA-A*02:01) and bovine alleles (BoLA-1:00901 and BoLA-2:00501)

MHC-I Bovine					
Genes	**Alleles**	**start**	**end**	**Epitope**	**score**
cya	BoLA-1:00901	20	14	IPNKFSIISFSVLL	2.90E-05
	BoLA-2:00501	475	488	EVMAKNVEGVLKPL	0.004
lef	**BoLA-1:00901**	**674**	**687**	**EQYTHQDEIYEQVH**	**0.158**
	BoLA-2:00501	276	289	YMDKFNEQEINLSL	0.015
pagA	BoLA-1:00901	326	339	RTHTSEVHGNAEVH	0.022
	BoLA-2:00501	106	119	YTFATSADNHVTMW	0.0085
MHC-I Human					
Genes	**Alleles**	**start**	**end**	**Epitope**	**score**
cya	HLA-A*01:01	234	247	LTEFQHAFSLAFSY	0.235
	HLA-A*02:01	18	31	VLLFAISSSQAIEV	0.010
lef	HLA-A*01:01	469	482	YLYENMNINNLTAT	1.50E-05
	HLA-A*02:01	276	289	YMDKFNEQEINLSL	0.096
pagA	**HLA-A*01:01**	**704**	**717**	**FIDFKKYNDKLPLY**	**0.387**
	HLA-A*02:01	594	607	QLAELNVTNIYTVL	0.027

Table 5: List of T-cell epitopes that are antigenic, non-allergenic and toxic of the three genes of human alleles (DRB1*0101, DRB1*0401 and DRB1*1501) and bovine alleles (BoLA-DRB3*0101 and BoLA-DRB3*0201)

MHC-II Bovine					
Genes	**Alleles**	**start**	**end**	**Epitope**	**score**
cya	BoLA-DRB3*0101	323	336	TYILFRPVNKLATNL	0.824
	BoLA-DRB3*0201	148	161	TPKLIINIKDYAINS	0.798
lef	BoLA-DRB3*0101	784	797	RLKVQKNAPKTFQFI	0.840
	BoLA-DRB3*0201	784	797	RLKVQKNAPKTFQFI	0.738
pagA	**BoLA-DRB3*0101**	**612**	**625**	**NAKMNILIRDKRFHY**	**0.887**
	BoLA-DRB3*0201	660	673	DKDIRKILSGYIVEI	0.873
MHC-II Human					
Genes	**Alleles**	**start**	**end**	**Epitope**	**score**
cya	HLA-DRB1*0101	233	246	LTEFQHAFSLAFSYY	0.886
	HLA-DRB1*0401	233	246	LTEFQHAFSLAFSYY	0.642
	HLA-DRB1*1501	236	249	EFQHAFSLAFSYYFA	0.636
lef	**HLA-DRB1*0101**	**530**	**543**	**RLKWRIQLSPDTRAG**	**0.889**
	HLA-DRB1*0401	532	545	LKWRIQLSPDTRAGY	0.663
	HLA-DRB1*1501	110	123	IGGKIYIVDGDITKH	0.381
pagA	HLA-DRB1*0101	713	726	LPLYISNPNYKVNVY	0.842
	HLA-DRB1*0401	390	403	NANIRYVNTGTAPIY	0.345
	HLA-DRB1*1501	390	403	NANIRYVNTGTAPIY	0.455

Modelling and Molecular Docking:

The *cya*, *lef* and *pagA* genes were modelled with the Modeller tool. The obtained B-cell for *cya*, *lef* and *pagA* "KRNHKTEKNKTEKEKFKDSI","SGPVFIPLVQGAGGHGD VGM" and "RSTSAGPTVPDRDNDGIPDS" and the T-cell bovine MHC-I for *cya*, *lef* and *pagA* "EVMAKNVEGVLKPL", "EQYTHQDEIYEQVH" and "RTHTSEVHGNAEVH" and MHC-II "TYILFRPVNKLATNL", "RLKVQKNAPKTFQFI" and "NAKMNILIRDKRFHY" epitopes that had higher protective antigenicity, and that were non-allergen and non-toxic were highlighted and visualized on the *cya*, *lef* and *pagA* genes respectively (Figure 1 and Figure 2).

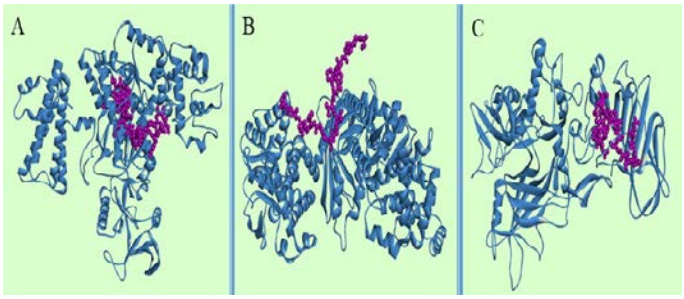


Figure 1: The modelled and predicted B cell epitopes highlighted in pink for *cya* (A), *lef* (B) and *pagA* (C).

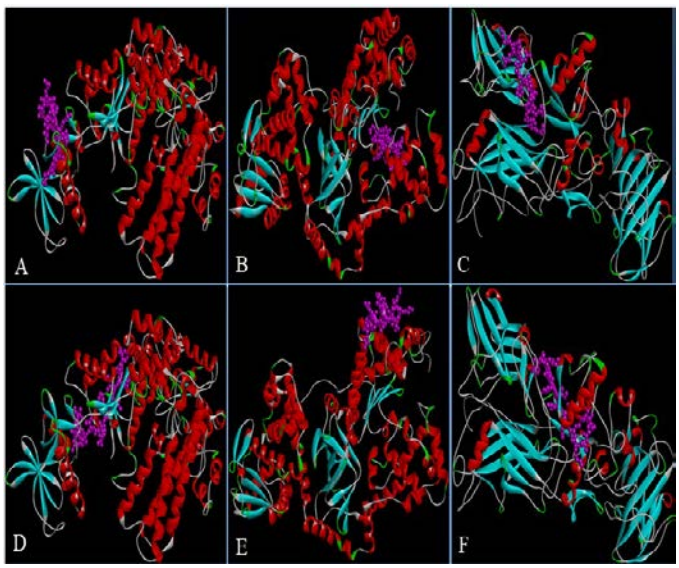


Figure 2: The predicted T cell epitopes highlighted in pink for Bovine Alleles BoLA-1:00901 and BoLA-2:00501 (MHC-I) for *cya* (A), *lef* (B) and *pagA* (C). The predicted epitopes for the alleles BoLA-DRB3*0101 and BoLA-DRB3*0201 (MHC-II) for *cya* (D), *lef* (E) and *pagA* (F) has been highlighted in pink.

Among the predicted epitopes, only those epitopes that had a high score were chosen for modelling using the PEP FOLD 3 web-based tool and the best epitope structure was selected based on conformity. The obtained 3D epitope structures are provided in a supplementary file (S4). The lowest binding energy was chosen based on the binding energy to produce the optimal binding and to forecast genuine CTL and HTL epitope as accurately as feasible and further was used for docking with respective MHC-I (6AT9, 5HHR) and MHC-II (2FSE) alleles were obtained from RCSB PDB databank and were utilized for docking studies. All HLA had considerable binding affinities for the predicted peptide. The CTL epitopes with MHC-I alleles (HLA*0101, HLA*0201) were docked and it was observed that 6AT9-"LTEFQHAFSLAFSY",5HHR- "YMDKFNEQEINLSL" and 6AT9- "FIDFKKYNDKLPLY" had a binding score of -41.5, -26.85 and -43.81 respectively along with ligand interactions (Figure 3).

Similarly, HTL epitopes with MHC-II (HLA-DRB*10101) with 2FSE-"LTEFQHAFSLAFSYY", 2FSE "RLKWRIQLSPDTRAG" and 2FSE- "LPLYISNPNYKVNVY" and their binding energies were shown to be negative -58.08, -45.05 and -66.48 along with ligand interactions respectively (Figure 4).

Discussion

This study sought to recognize highly immunogenic epitopes for B and T cells using the cya, lef and pagA genes as target candidates, the primary molecules of humoral and cell-mediated immunity, as peptide vaccine candidates for anthrax disease. Though the vaccines are already being developed, still anthrax outbreaks and the spread of the disease is continuous. As a result, new candidate vaccines against anthrax are desperately needed, and some are now being tested in clinical trials. The emergence of reverse vaccinology, as well as the availability of genomes and proteomics data, aid in vaccine development. Epitope prediction's fundamental goal is to create a molecule that can take over the role of an antigen in the procedure of inducing a suitable immune response.[44] The designed molecules are preferred for vaccine production because they are cost-effective and non-infectious, as opposed to complete pathogenic organisms, which may pose dangers to researchers or test participants animals and humans.16 Immunogenic antigen identification is a critical stage in vaccine development since it could be used to predict epitopes in silico. Hence, when compared to conventional drug designing, the effective adoption of bioinformatics techniques is advantageous.[31,47]

Therefore, utilizing various bioinformatics tools/software the linear B and T-cell epitopes were predicted from genes (cya, lef and pagA) against anthrax. The epitopes play a critical part in cellular and biological development, they were chosen to predict vaccines. These B-cell and T-cell epitopes might theoretically be exploited to develop anthrax vaccines that are effective at stimulating both humoral and cell-mediated defence.[17] The obtained 20mer B-cell epitopes with best scores, having antigenicity, that are non-allergenic and non-toxic were found for cya- "KRNHKTEKNKTEKEKFKDSI", lef- "SGPVFIPLVQGAGGHGDVGM" and pagA- "RSTSAGPTVPDRDNDGIPDS" for both humans and bovine respectively.

A systematic study of all susceptible HLA and BoLA alleles of MHC class I and MHC class II compounds was done for CTL and HTL target identification. This study examined only HLA individual alleles for epitope prediction analysis and further for

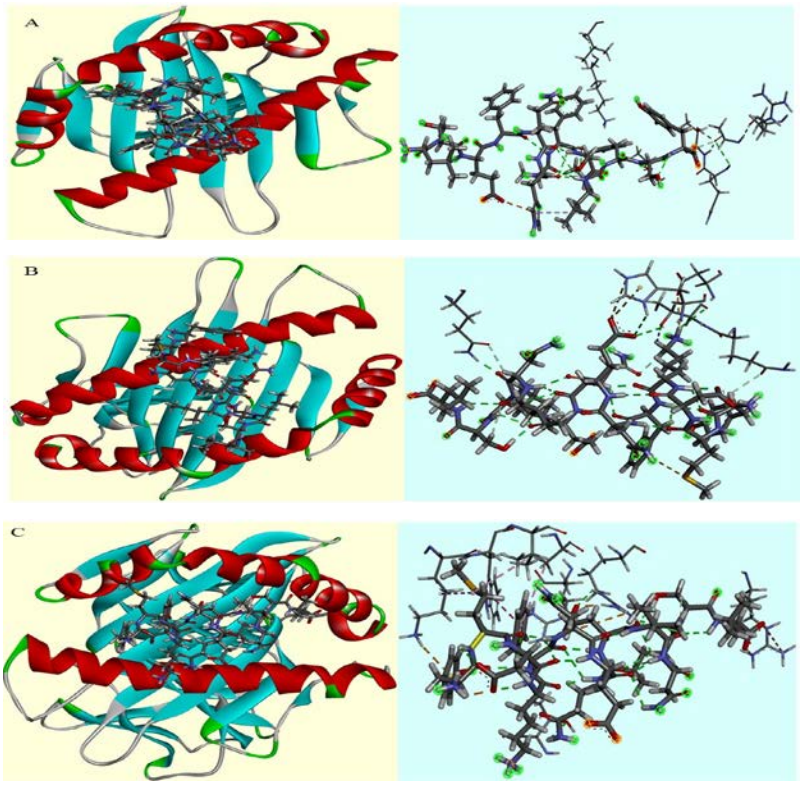


Figure 3: The predicted T cell epitopes docked with HLA*0101, HLA*0201 (MHC-I) for *cya* (A), *lef* (B) & *pagA* (C)- left and the interactions of each peptide with its respective allele –right.

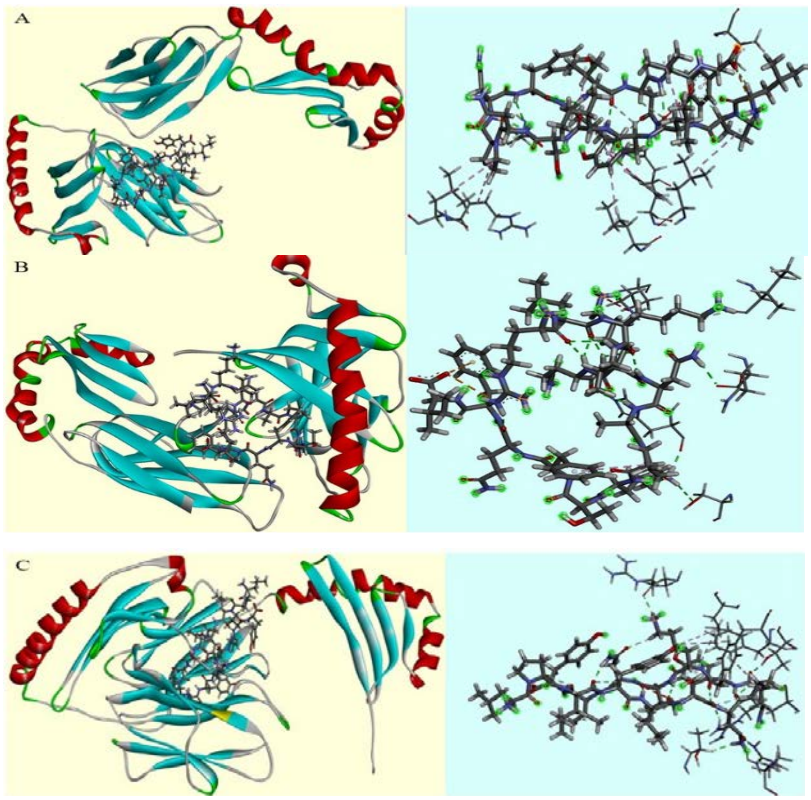


Figure 4: The predicted T cell epitopes docked with HLA-DRB*10101 (MHC-II) for cya (A), lef (B) & pagA (C)- left and the interactions of each peptide epitope with the respective allele –right

docking studies, because of the lack of protein structure for the BoLA alleles, even though the epitopes were predicted for the bovine considering respective alleles. Among the 14mer CTL epitopes for BoLA alleles of MHC-I, the best epitope for cya is "EVMAKNVEGVLKPL", lef "EQYTHQDEIYEQVH" and pagA "RTHTSEVHGNAEVH" in BoLA-1:00901, BoLA-1:00901 and BoLA-2:00501 alleles respectively. Therefore, the epitope peptide "EQYTHQDEIYEQVH" of lef had a higher score, hence being suitable to develop a peptide vaccine. Among the CTL epitopes for HLA-A alleles, the epitopes for cya, lef and pagA were "LTEFQHAFSLAFSY", "YMDKFNEQEINLSL" and "FIDFKKYNDKLPLY" in the HLA-A*01:01, HLA-A*02:01 and HLA-A*01:01 alleles respectively. Among these epitopes "FIDFKKYNDKLPLY" of pagA was found to be an appropriate CTL epitope for MHC-I.

Similarly, among the 14mer HTL epitopes for BoLA alleles of MHC-II, the epitopes for cya, lef and pagA were "TYILFRPVNKLATNL", "RLKVQKNAPKTFQFI" and pagA "NAKMNILIRDKRFHY" in BoLA-DRB3*0101 allele. Also, the epitope peptide "NAKMNILIRDKRFHY" of pagA had a higher score, hence being suitable to develop a peptide vaccine. It was observed that all the three predicted epitopes belong to BoLA-DRB3*0101 allele. Further, the HTL epitopes for HLA-A alleles, the epitopes for cya, lef and pagA were "LTEFQHAFSLAFSYY", "RLKWRIQLSPDTRAG" and "LPLYISNPNYKVNVY" in the allele HLA-DRB1*0101. Among these epitopes "RLKWRIQLSPDTRAG" of pagA was found to be an appropriate CTL epitope for MHC-II.

Further, the selected epitopes were docked with respective MHC alleles that were downloaded from RCSB PDB had negative binding energy. Among the CTL and MHC-I docking, the pagA gene with "FIDFKKYNDKLPLY"- "6AT9" had high negative energy of -43.81. Similarly, in the HTL and MHC-II docking analysis it was observed that the epitope "LPLYISNPNYKVNVY"- "2FSE" had high negative energy of -66.48. *In silico* study reported in other reports,[28,29,47] on prediction of epitopes against anthrax had predicted potential T cell epitopes for MHC class I- HLA-A*30:01 and MHC class II -LA-DRB5*01:01 alleles respectively, but the

epitope score and docking energy was less compared to this study. Since all the obtained epitopes had a higher score and also high negative binding energy compared, these epitopes can be chosen as appropriate peptide epitopes for drug targets for vaccine development.

Conclusion

This work aimed to predict *in silico* peptide vaccine against the *cya*, *lef*, and *pagA* genes of *B. anthracis* based on immunoinformatics and docking studies because anthrax consequences have arisen in the form of biological warfare and outbreaks have been reported to date. Further, based on the findings of this study, using final anticipated areas of high immunogenic scores acquired using bioinformatics methods could be a viable substitute to starting actual trials targeting anthrax. The discovery of enigmatic B and T cell epitopes that may be important in the induction of neutralizing antibody levels to the *cya*, *lef*, and *pagA* genes that could enable the cellular and humoral immune systems to provide substantial protection against anthrax disease after immunization. Through bioinformatics, researches have been published to aid peptide design, not all peptides predicted computationally are effectively immunogenic in vivo, therefore testing the predicted peptides in vivo is still required to validate that T cell responses are triggered.

Acknowledgments

The authors are thankful to the institute ICAR-NIVEDI for providing necessary infrastructure, facility and guidance throughout the study.

Author Contributions

UBI conceptualized the methodology for analysis; UBI performed the analysis and drafted the manuscript; MSB performed analysis KPS and SSP supervised the work; UBI, KPS, SSP, CS, MP and RA read the draft and approved the final manuscript.

Supplementary Information

Table S1: List of the physical and chemical properties of each gene.

Table S2: List of T-cell (CTL) epitopes for MHC-I in human alleles (HLA-A*01:01 and HLA-A*02:01) and bovine alleles (BoLA-1:00901 and BoLA-2:00501) of the three genes.

Table S3: List of T-cell (HTL) epitopes for MHC-II of human alleles (DRB1*0101, DRB1*0401 and DRB1*1501) and bovine alleles (BoLA-DRB3*0101 and BoLA-DRB3*0201) of the three genes

Figure S4: The predicted CTL (A, B, C) and HTL (D, E, F) epitopes structures of the three genes obtained from the PEP-FOLD server.

Conflict of Interest: Authors declare no conflict of interest.

References

1. A.K. Goel. Anthrax: A disease of biowarfare and public health importance. *World J. Clin. Cases* **2015**, 3 (1), 20.
2. M. Mock, A. Fouet. Anthrax: Annual Review. **2001**, 55, 647–671.
3. P.L. Goossens, J.N. Tournier. Crossing of the epithelial barriers by Bacillus anthracis: The Known and the Unknown. *Front. Microbiol.* **2015**, 6 (OCT), 1–15.
4. M. Mock, T. Mignot. Anthrax toxins and the host: A story of intimacy. *Cell. Microbiol.* **2003**, 5 (1), 15–23.
5. D. Romero-Alvarez, A.T. Peterson, J.S. Salzer, et al. Potential distributions of Bacillus anthracis and Bacillus cereus biovar anthracis causing anthrax in Africa. *PLoS Negl. Trop. Dis.* **2020**, 14 (3), 1–20.
6. C.J. Carlson, I.T. Kracalik, N. Ross, et al. The global distribution of Bacillus anthracis and associated anthrax risk to humans, livestock and wildlife. *Nat. Microbiol.* **2019**, 4 (8), 1337–1343.
7. K.V. Chanu, D. Thakuria, R. Sood, et al. Identification of immunodominant epitopes in the HA2 subunit of H5N1 haemagglutinin by immunoassay using synthetic peptides as antigens. *Chem. Biol. Lett.* **2017**, 4 (1), 20–26.
8. T. Candela, A. Fouet. Poly-gamma-glutamate in bacteria. *Mol. Microbiol.* **2006**, 60 (5), 1091–1098.
9. N.S. Duesbery, G.F. Vande Woude. Anthrax toxins. *Cell. Mol. Life Sci.* **1999**, 55 (12), 1599–1609.
10. J.G. Bann. Anthrax toxin protective antigen - Insights into molecular switching from prepore to pore. *Protein Sci.* **2012**, 21 (1), 1–12.
11. H. Barth, K. Aktories, M.R. Popoff, B.G. Stiles. Binary Bacterial Toxins: Biochemistry, Biology, and Applications of Common Clostridium and Bacillus Proteins . *Microbiol. Mol. Biol. Rev.* **2004**, 68 (3), 373–402.
12. E.D. Williamson, E.H. Dyson. Anthrax prophylaxis: Recent advances and future directions. *Frontiers in Microbiology*. 2015, pp 1–8.
13. J. Shi, J. Zhang, S. Li, et al. Epitope-based vaccine target screening against highly pathogenic MERS-CoV: An In Silico approach applied to emerging infectious diseases. *PLoS One* **2015**, 10 (12), 1–16.
14. S. Srivastava, M. Kamthania, S. Singh, A.K. Saxena, N. Sharma. Structural basis of development of multi-epitope vaccine against middle east respiratory syndrome using in silico approach. *Infect. Drug Resist.* **2018**, 11, 2377–2391.
15. B.B.M. Gaafar, S.A. Ali, K.A. Abd-elrahman, Y.A. Almofti. Immunoinformatics Approach for Multiepitope Vaccine Prediction from H, M, F, and N Proteins of Peste des Petits Ruminants Virus. *J. Immunol. Res.* **2019**, 2019, 1–18.
16. H. Zahroh, A. Ma'rup, U.S.F. Tambunan, A.A. Parikesit. Immunoinformatics approach in designing epitopebased vaccine against meningitis-inducing bacteria (Streptococcus pneumoniae,Neisseria meningitidis,and Haemophilus influenzae type b). *Drug Target Insights* **2016**, 10, 19–29.
17. K.K. Dubey, G.A. Luke, C. Knox, et al. Vaccine and antibody production in plants: Developments and computational tools. *Brief. Funct. Genomics* **2018**, 17 (5), 295–307.
18. A. Ali, A. Khan, A.C. Kaushik, et al. Immunoinformatic and systems biology approaches to predict and validate peptide vaccines against Epstein–Barr virus (EBV). *Sci. Rep.* **2019**, 9 (1), 1–12.
19. L. Florea, B. Halldórsson, O. Kohlbacher, et al. Epitope prediction algorithms for peptide-based vaccine design. *Proc. 2003 IEEE Bioinforma. Conf. CSB 2003* **2003**, 17–26.
20. M.H.V. Van Regenmortel. Mapping epitope structure and activity: From one-dimensional prediction to four-dimensional description of antigenic specificity. *Methods A Companion to Methods Enzymol.* **1996**, 9 (3), 465–472.
21. G.M.A. Gillespie, M.R. Wills, V. Appay, et al. Functional Heterogeneity and High Frequencies of Cytomegalovirus-Specific CD8 + T Lymphocytes in Healthy Seropositive Donors. *J. Virol.* **2000**, 74 (17), 8140–8150.
22. D.S. Doering, P. Matsudaira. Cysteine scanning mutagenesis at 40 of 76 positions in villin headpiece maps the F-actin binding site and structural features of the domain. *Biochemistry* **1996**, 35 (39), 12677–12685.
23. P. Oyarzún, B. Kobe. Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production. *Hum. Vaccines Immunother.* **2016**, 12 (3), 763–767.
24. A.M. Khan, O. Miotto, A.T. Heiny, et al. A systematic bioinformatics approach for selection of epitope-based vaccine targets. *Cell. Immunol.* **2006**, 244 (2), 141–147.
25. A. Khan, M. Junaid, A.C. Kaushik, et al. Computational identification, characterization and validation of potential antigenic peptide vaccines

from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches. *PLoS One* **2018**, 13 (5), 1–25.
26. R.V. Abinaya, P. Viswanathan. Chapter 2 - Biotechnology-based therapeutics. In *Translational Biotechnology*; Hasija, Y., Ed.; Academic Press, **2021**; pp 27–52.
27. P. Lidder, A. Sonnino. Biotechnologies for the Management of Genetic Resources for Food and Agriculture. In *Advances in Genetics*; Goodwin, S. F., Friedmann, T., Dunlap, J. C., Eds.; Advances in Genetics; Academic Press, **2012**; Vol. 78, pp 1–167.
28. S. Ascough, R.J. Ingram, K.K.Y. Chu, et al. CD4+T cells targeting dominant and cryptic epitopes from Bacillus anthracis lethal factor. *Front. Microbiol.* **2016**, 6 (JAN), 1–12.
29. M. Tahmoorespur, N. Nazifi, Z. Pirkhezranian. In Silico Prediction of B-Cell and T-Cell Epitopes of Protective Antigen of Bacillus anthracis in Development of Vaccines Against Anthrax. *Iran. J. Appl. Anim. Sci.* **2017**, 7 (3), 429–436.
30. M. Shrivastava, P. Sharma, R. Singh. Identification of potential CYP51 inhibiting anti-Aspergillus phytochemicals using molecular docking and ADME/T studies. *Chem. Biol. Lett.* **2021**, 8 (1), 18–21.
31. V. Chandel, M. Srivastava, A. Srivastava, S. Asthana, D. Kumar. In-silico interactions of active Phytochemicals with c-MYC EGFR and ERBB2 oncoproteins. *Chem. Biol. Lett.* **2020**, 7 (1), 47–54.
32. I.A. Doytchinova, D.R. Flower. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* **2007**, 8, 1–7.
33. E. Gasteiger, C. Hoogland, A. Gattiker, et al. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker, J. M., Ed.; Humana Press, Totowa, NJ, **2005**; pp 571–607.
34. D.T. Nair, K. Singh, Z. Siddiqui, et al. Epitope Recognition by Diverse Antibodies Suggests Conformational Convergence in an Antibody Response. *J. Immunol.* **2002**, 168 (5), 2371–2382.
35. Y. El-Manzalawy, D. Dobbs, V. Honavar. Predicting linear B-cell epitopes using string kernels. *J. Mol. Recognit.* **2008**, 21 (4), 243–255.
36. F. Borrego. The First Molecular Basis of the “Missing Self” Hypothesis. *J. Immunol.* **2006**, 177 (9), 5759–5760.
37. A. Jain, P. Tripathi, A. Shrotriya, R. Chaudhary, A. Singh. In silico analysis and modeling of putative T cell epitopes for vaccine design of Toscana virus. *3 Biotech* **2015**, 5 (4), 497–503.
38. R. Vita, J.A. Overton, J.A. Greenbaum, et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res.* **2015**, 43 (D1), D405–D412.
39. S. Pyasi, V. Sharma, K. Dipti, N.A. Jonniya, D. Nayak. Immunoinformatics approach to design multi-epitope-subunit vaccine against bovine ephemeral fever disease. *Vaccines* **2021**, 9 (8), 1–20.
40. A. Lamiable, P. Thévenet, J. Rey, et al. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res.* **2016**, 44 (W1), W449–W454.
41. D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H.J. Wolfson. PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Res.* **2005**, 33 (SUPPL. 2), 363–367.
42. D. Duhovny, R. Nussinov, H.J. Wolfson. Efficient unbound docking of rigid molecules. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; **2002**; Vol. 2452, pp 185–200.
43. N. Andrusier, R. Nussinov, H.J. Wolfson. FireDock: Fast interaction refinement in molecular docking. *Proteins Struct. Funct. Bioinforma.* **2007**, 69 (1), 139–159.
44. E.F. Pettersen, T.D. Goddard, C.C. Huang, et al. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, 25 (13), 1605–1612.
45. U. BIOVIA, Dassault Systèmes (Dassault Systèmes; San Diego, CA. Discovery Studio.
46. S. Yuan, H.C.S. Chan, Z. Hu. Using <scp>PyMOL</scp> as a platform for computational drug design. *WIREs Comput. Mol. Sci.* **2017**, 7 (2), 1298.
47. A. Shamakhi, E. Kordbacheh. Immunoinformatic design of an epitope-based immunogen candidate against Bacillus anthracis. *Informatics Med. Unlocked* **2021**, 24, 100574.