



IDENTIFICATION OF T-CELL EPITOPE FOR PROTECTIVE ANTIGEN AND LETHAL FACTOR OF *Bacillus anthracis* - AN IMMUNOINFORMATICS APPROACH

ANURAJ NAYARISSERI^{1*}, MUKESH YADAV¹, SHEAZA AHMED¹, JYOTI SAHU², PRIYANKA GUPTA³, RACHNA CHOURASIA² AND DEEPTI MITTAL⁴

¹In silico Research Laboratory, Eminent Biosciences, Vijaynagar, Indore-452010, MP, India

²Dept of Biotechnology, Softvision College, Vijaynagar, Indore-452010, MP, India.

³Kalinga School of Biotechnology, Patia, Bhubaneswar-751031, Odisha, India.

⁴Dept. of Bioinformatics, Dr. D.Y. Patil Biotechnology & Bioinformatics Institute, Tathawade, Pune-411033, MS, India

*Corresponding Author: Email- anuraj@eminentbio.com

Received: June 11, 2012; Accepted: June 21, 2012

Abstract- Anthrax is an infectious disease caused by spores of the bacterium, *Bacillus anthracis*. The *B. anthracis* spores are highly resistant to inactivation and may be present in the soil, for example, for decades, occasionally infecting grazing animals that ingest the spores. Goats, sheep and cattle are examples of animals that may become infected. Human infection may occur by three routes of exposure to anthrax spores: cutaneous (through the skin), gastrointestinal (by ingestion) and pulmonary (inhalation). A major factor in the virulence of *Bacillus anthracis* is its secretion of three binary toxins, protective antigen, lethal toxin and edema toxin. These toxins possess a common cell receptor-binding (B) component but have distinct biochemically active (A) components. In traditional vaccine approaches made researchers to produce vaccine very easy. The current approach was based upon the *insilico* identification of T-Cell Epitopes and its binding affinity with MHC Class II. The T-Cell Epitope was identified using the HLA Pred which identifies the peptide class that binds with both MHC Class I and II. Further the Epitope was analyzed for peptide characteristics such as antigenic nature, lipophilicity and solvent assessable. Out of all peptides the peptide which shows all the character for further analyzed for protein peptide interaction to find the high affinity binding peptide with low energy.

Key words- T-Cell Epitope, *Bacillus anthracis*, Protective Antigen, Lethal factor.

Citation: Anuraj Nayarissari, et al (2012) Identification of T-Cell Epitope for Protective Antigen and Lethal factor of *Bacillus anthracis* - An Immunoinformatics Approach. International Journal of Drug Discovery, ISSN: 0975-4423 & E-ISSN: 0975-914X, Volume 4, Issue 1, pp.-137-144.

Copyright: Copyright©2012 Anuraj Nayarissari, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Introduction

Bacillus anthracis is a Gram-positive bacteria, was the first bacterium conclusively demonstrated to cause disease, by Robert Koch in 1877 [17]. An endospore forming bacterium, *B. anthracis* is a natural soil-dwelling organism, as well as the causative agent of anthrax. Under conditions of environmental stress, *B. anthracis* bacteria naturally produce endospores which rest in the soil and can survive for decades in this state. When ingested by cattle, sheep, or other herbivores, the bacteria begin to reproduce inside the animal and eventually kill it and then continue to reproduce in its carcass. Once the nutrients are exhausted, new endospores are produced and the cycle repeats. *B. anthracis* has at least 89 known strains, ranging from highly virulent strains with biological warfare and bioterrorism applications (Ames and Vollum) to be-

nign strains used for inoculations (Sterne). The strains differ in presence and activity of various genes, determining their virulence and production of antigens and toxins. The form associated with the 2001 anthrax attacks produced both toxin (consisting of three proteins: the protective antigen, the edema factor and the lethal factor) and a capsule (consisting of a polymer of glutamic acid) [17]. The species name *anthracis* is from the Greek word *anthrakis*, meaning coal and referring to the most common form of the disease. The bacterium can be cultivated in ordinary nutrient medium under aerobic or anaerobic conditions.

Infections with *B. anthracis* can be treated with β -lactam antibiotics such as penicillin and others which are active against Gram-positive bacteria. The Gram-positive bacterium *Bacillus anthracis* infects through intra-dermal inoculation, ingestion, or inhalation of

spores. Spores are dormant forms of *B. anthracis* and are extremely resistant to environmental stress. In the current model of respiratory infections, the spores are first taken up by macrophages where they germinate and become vegetative bacteria [20]. Macrophages then transport the bacteria to the regional lymph nodes. In these organs, the bacteria escape from the macrophage and spread through the lymphatics and blood stream causing massive septicemia. Vegetative *B. anthracis* express two essential virulence factors: the tripartite anthrax toxin and the poly- γ -D glutamic acid capsule. The toxin-protective antigen binds to its specific receptor and translocates the edema and lethal factors into the cytosol. Edema factor is an adenylate cyclase that causes tissue edema, whereas lethal factor is a metalloprotease that inactivates mitogen-activated protein kinase-kinase and provokes cell death [6].

Neutrophils are a vital component of the acute inflammatory response and play a key role in the resolution of microbial infections. They are terminally differentiated cells, incapable of cell division and synthesize very low levels of RNA and protein. Neutrophils engulf microbes into a phagosome that fuses with intracellular granules to form a phagolysosome [9]. In the phagolysosome the bacteria are killed through the interaction of reactive oxygen species (ROS) and oxygen-independent mediators such as enzymes and antimicrobial peptides. Antimicrobial peptides are predominantly cationic and are thought to permeabilize the bacterial membrane and lyse microbes.

Inhalation of *B. anthracis* infections result in sepsis and death, while cutaneous anthrax almost always remains localized. Interestingly, in untreated cutaneous cases, neutrophils surround the necrotic, bacteria-containing tissue, whereas neutrophil infiltration is rarely seen in the lung during inhalation anthrax. Although there are reports of pulmonary infiltration in *B. anthracis* infections, they were thought to be due to preexisting lesions. Pigs and dogs also develop a cutaneous form when infected subcutaneously [21]. This form is reminiscent of human cutaneous anthrax and is accompanied by massive neutrophil infiltrations. Also, neutrophil recruitment is found in the lungs of pulmonary infected dogs and pigs that survive high infection doses. Therefore, we hypothesized that neutrophils can kill *B. anthracis* [21].

Anthrax toxin is a key virulence factor for *Bacillus anthracis*, the causative agent of anthrax. Here we discuss what is known about the anthrax toxin receptor (ATR), the cellular receptor for anthrax toxin and how this information is being used to develop treatments for anthrax as well as to understand aspects of cancer. ATR was identified recently as a type I transmembrane protein with unknown function that contains an extracellular integrin-like inserted (I) domain. The ATR I domain contains the toxin binding site and a soluble form of this domain was shown to serve as an effective antitoxin to protect cultured cells from toxin action. ATR is encoded by the tumor endothelial marker 8 (TEM8) gene, which is selectively up-regulated during blood vessel formation and in tumor vasculature, raising the possibility that this protein normally functions in angiogenesis [2]. Therefore, identification of the cellular receptor for anthrax toxin has made possible new avenues of research in the areas of anthrax pathogenesis, antitoxin development and cancer biology.

Protective Antigen

Protective antigen (PA) is one of the three proteins composing the anthrax toxin, the agent which infects many mammalian species and that may cause death. PA binds to a receptor (ATR) in sensitive eukaryotic cells, thereby facilitating the translocation of the enzymatic toxin components, edema factor and lethal factor, across the target T-cell membrane. PA associated with LF causes death when injected, PA associated with EF produces edema. PA induces immunity to infection with anthrax. Anthrax toxins are composed of three distinct proteins, a protective antigen (PA), a lethal factor (LF) and an edema factor (EF). None of these is toxic by itself. PA+LF forms the lethal toxin (LeTx); PA+EF forms the edema toxin (EdTx). PA-63 forms heptamers and this oligomerization is required for LF or EF binding. Once activated, at low pH, the heptamer undergoes conformational changes [11].

Lethal Factor

The lethal toxin of *Bacillus anthracis* consists of two proteins, PA and LF, which together induce lethal effects in some animal species and cause macrophage lysis. LF is a zinc-binding protein with metalloprotease activity [8]. With a two-hybrid system approach we identified MAP kinase kinases (MAPKKs) Mek1 and Mek2 as proteins interacting with LF. LF was shown to cleave Mek1 and Mek2 and an additional MAPKK family member MKK3, within their N-terminal region. We examined macrophage cell lines and primary peritoneal cells with different sensitivities to LF but did not find a direct correlation between MAPKKs cleavage and cell death. On the other hand, sublytic doses of LF cleave MAPKKs and cause a reduction in the LPS/IFN γ -induced production of pro-inflammatory mediators [8]. These findings are discussed with respect to the possible role of LF in the initial phase of infection

Edema Factor

One of three protein components of anthrax toxin produced by the pathogenic bacterium, *Bacillus anthracis*. Central moiety that mediates entry of lethal factor (LF) and edema factor (EF) into the target T-cell. Protective antigen (also known as PA or PA63) binds to the cell surface via a type I membrane receptor with a von Willebrand factor A domain, known as anthrax toxin receptor. It is proteolytically activated in vivo by a furin-like protease to produce a 63 kDa protein (PA63). PA63 self-associates to form a heptameric, cation-selective, voltage-gated, membrane channel that subsequently binds to LF and/or EF. Cleavage of the 20 kDa fragment from PA63 is necessary for binding of protective antigen to LF and EF. Highly specific protease component that cleaves members of the mitogen-activated protein kinase kinase (MAPKK) family. LF is comprised of four domains: domain I that binds to protective antigen (necessary for entry into the cellular target) and domains II, III and IV which form a long groove to hold and cleave the target proteins. Requires passage through an acidic vesicle prior to translocation to the cytoplasm of the cell. Cleavage of target proteins leads to the inhibition of multiple signaling pathways and cell death [13].

Cytotoxic T-cell

A cytotoxic T-cell (also known as T_C, CTL or killer T-cell) belongs to a sub-group of T lymphocytes (a type of white blood cell) which are capable of inducing the death of infected somatic or tumor

cells; they kill cells that are infected with viruses (or other pathogens), or are otherwise damaged or dysfunctional. Most cytotoxic T-cells express T-cell receptors (TcRs) that can recognise a specific antigenic peptide bound to Class I MHC molecules, present on all nucleated cells and a glycoprotein called CD8, which is attracted to non-variable portions of the Class I MHC molecule. The affinity between CD8 and the MHC molecule keeps the T_C cell and the target T-cell bound closely together during antigen-specific activation. CD8⁺ T-cells are recognized as T_C cells once they become activated and are generally classified as having a pre-defined cytotoxic role within the immune system [3].

Hematopoietic stem cells in the bone marrow migrate into the thymus, where they undergo VDJ recombination of their beta-chain TcR DNA to form a developmental form of the TcR protein, known as pre-TcR. If that rearrangement is successful, the cells then rearrange their alpha-chain TcR DNA to create a functional alpha-beta TcR complex. This highly-variable genetic rearrangement product in the TcR genes helps create millions of different T-cells with different TcRs, helping the body's immune system respond to virtually any protein of an invader. The vast majority of T-cells express alpha-beta TcRs ($\alpha\beta$ T-cells), but some T-cells in epithelial tissues (like the gut) express gamma-delta TcRs ($\gamma\delta$ T-cells), which recognize non-protein antigens [3].

T-cells with functionally stable TcRs express both the CD4 and CD8 co-receptors and are therefore termed "double-positive" (DP) T-cells (CD4⁺CD8⁺). The double-positive T-cells are exposed to a wide variety of self-antigens in the thymus and undergo two selection criteria: (1) Negative selection, in which those double-positive T-cells that bind too strongly to MHC-presented self antigens undergo apoptosis because their propensity to become autoreactive could lead to autoimmunity; and (2) Positive selection, in which those double-positive T-cells that bind too weakly to MHC-presented self antigens undergo apoptosis because of their inability to recognize MHC-protein complexes. Only those T-cells that bind to the MHC-self-antigen complexes weakly are positively selected [19]. Those cells that survive positive and negative selection differentiate into single-positive T-cells (either CD4⁺ or CD8⁺) dependent on whether their TcR recognizes an MHC class I presented antigen (CD8) or an MHC class II presented antigen (CD4). It is the CD8⁺ T-cells that will mature and go on to become cytotoxic T-cells following their activation with a class I restricted antigen.

Cytotoxic T-Cell Activation

With the exception of some cell types such as non-nucleated cells (including erythrocytes), Class I MHC is expressed by all host T-cells. When these cells are infected with a virus (or another intracellular pathogen), the cells "break down" foreign proteins via antigen processing. These result in peptide fragments, some of which are presented by MHC Class I to the T-cell antigen receptor (TcR) on CD8⁺ T-cells.

The activation of cytotoxic T-cells is dependent on several simultaneous interactions between molecules expressed on the surface of the T-cell and molecules on the surface of the antigen presenting cell (APC). For instance, consider the two signal model for T_C cell activation. The first signal occurs when the TcR of the T_C cell strongly interacts with a peptide-bound MHC class I molecule on the APC. There is a second interaction between the CD8

coreceptor and the class I MHC molecule to stabilize this signal. The second signal comes from an interaction between the CD28 molecule on the T-cell and either CD80 or CD86 (also called B7-1 and B7-2) on the surface of the APC. CD80 and CD86 are known as costimulators for T-cell activation. This second signal can be assisted (or replaced) by stimulating the T_C cell with cytokines released from helper T-cells [7]. Once activated, the T_C cell undergoes clonal expansion with the help of a cytokine called Interleukin-2 (IL-2) that is a growth and differentiation factor for T-cells. This increases the number of cells specific for the target antigen that can then travel throughout the body in search of antigen-positive somatic cells.

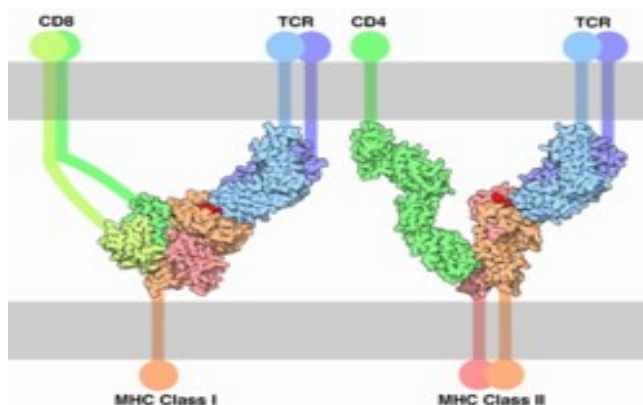


Fig. 1- Molecular association of CD8⁺ T-cells with MHC class I and CD4⁺ T-cells with MHC class II

Several different subsets of T-cells have been described, each with a distinct function.

Cytotoxic T-cells (T_C cells) destroy virally infected cells and tumor cells and are also implicated in transplant rejection. These cells are also known as CD8⁺ T-cells, since they express the CD8 glycoprotein at their surface. Helper T-cells, (T_H cells) are the "middlemen" of the adaptive immune system. Once activated, they divide rapidly and secrete small proteins called cytokines that regulate or "help" the immune response. These cells (also called CD4⁺ T-cells) are a target of HIV infection; the virus infects the cell by using the CD4 protein to gain entry. The loss of T_H cells as a result of HIV infection leads to the symptoms of AIDS. Memory T-cells are a subset of antigen-specific T-cells that persist long-term after an infection has resolved. They quickly expand to large numbers of effector T-cells upon re-exposure to their cognate antigen, thus providing the immune system with "memory" against past infections. Memory T-cells comprise two subtypes: central memory T-cells (T_{CM} cells) and effector memory T-cells (T_{EM} cells). Memory cells may be either CD4⁺ or CD8⁺. Regulatory T-cells (T_{reg} cells), formerly known as suppressor T-cells, are crucial for the maintenance of immunological tolerance. Their major role is to shut down T-cell mediated immunity towards the end of an immune reaction and to suppress auto-reactive T-cells that escaped the process of negative selection in the thymus. Two major classes of regulatory T-cells have been described, including the naturally occurring T_{reg} cells and the adaptive T_{reg} cells. Naturally occurring T_{reg} cells (also known as CD4⁺CD25⁺FoxP3⁺ T_{reg} cells) arise in the thymus, whereas the adaptive T_{reg} cells (also known as Tr1 cells or Th3 cells) may originate during a normal immune

response. Naturally occurring Treg cells can be distinguished from other T-cells by the presence of an intracellular molecule called FoxP3. Mutations of the FOXP3 gene can prevent regulatory T-cell development, causing the fatal autoimmune disease IPEX [22]. Natural Killer T-cells (NKT-cells) are a special kind of lymphocyte that bridges the adaptive immune system with the innate immune system. Unlike conventional T-cells that recognize peptide antigen presented by major histocompatibility complex (MHC) molecules, NKT-cells recognize glycolipid antigen presented by a molecule called CD1d. Once activated, these cells can perform functions ascribed to both T_H and T_C cells (i.e. cytokine production and release of cytolytic/cell killing molecules). $\gamma\delta$ T-cells represent a small subset of T-cells that possess a distinct TCR on their surface. A majority of T-cells have a TCR composed of two glycoprotein chains called α - and β - TCR chains. However, in $\gamma\delta$ T-cells, the TCR is made up of one γ -chain and one δ -chain. This group of T-cells is much less common (5% of total T-cells) than the $\alpha\beta$ T-cells, but are found at their highest abundance in the gut mucosa, within a population of lymphocytes known as intraepithelial lymphocytes (IELs). The antigenic molecules that activate $\gamma\delta$ T-cells are still widely unknown. However, $\gamma\delta$ T-cells are not MHC restricted and seem to be able to recognise whole proteins rather than requiring peptides to be presented by MHC molecules on antigen presenting cells. Some recognize MHC class IB molecules though. Human V γ 9/V δ 2 T-cells, which constitute the major $\gamma\delta$ T-cell population in peripheral blood, are unique in that they specifically and rapidly respond to a small non-peptidic microbial metabolite, HMB-PP, an isopentenyl pyrophosphate precursor [22].

B-Cell

B lymphocytes are the cells of the immune system that make antibodies to invading pathogens like viruses. They form memory cells that remember the same pathogen for faster antibody production in future infections. B-cells are lymphocytes that play a large role in the humoral immune response as opposed to the cell-mediated immune response that is governed by T-cells [10]. The principal function of B-cells is to make antibodies against soluble antigens. B-cells are an essential component of the adaptive immune system.

Development of B-Cells

B-cells are produced in the bone marrow of most mammals. Rabbits are an exception; their B-cells develop in the appendix-sacculus rotundus. B-cell development occurs through several stages, each stage representing a change in the genome content at the antibody loci. An antibody is composed of two light (L) and two heavy (H) chains and the genes specifying them are found in the 'H' chain locus and the 'L' chain locus. In the H chain loci there are three regions, V, D and J, which recombine randomly, in a process called VDJ recombination, to produce a unique variable domain in the immunoglobulin of each individual B-cell. Similar rearrangements occur for L chain locus except there are only two regions, namely V and J. The process of immunoglobulin formation at the different stages of B-cell development are 1). Progenitor B-cells - Contains Germline H genes, Germline L genes. 2). Early Pro-B-cells - undergoes D-J rearrangement on the H chains. 3). Late Pro-B-cells - undergoes V-DJ rearrangement on the H

chains.4). Large Pre-B-cells - the H chain is VDJ rearranged, Germline L genes. 5) Small Pre-B-cells - undergoes V-J rearrangement on the L chains, VDJ rearranged on H chains. There is start of expression of IgM receptors. 7) Mature B-cells - There is start of expression of IgD When the B-cell fails in any step of the maturation process, it will die by a mechanism called apoptosis [10]. If it recognizes self-antigen during the maturation process, the B-cell will become suppressed (known as anergy) or undergo apoptosis. B-cells are continuously produced in the bone marrow, but only a small portion of newly made B-cells survive to participate in the long-lived peripheral B-cell pool.

Functions

The human body produce millions of different types of B-cells each day that circulate in the blood and lymph performing the role of immune surveillance. They do not make antibodies until they become fully activated. Each B-cell has a unique receptor protein (referred to as the B-cell receptor (BCR)) on its surface that will bind to one particular antigen. The BCR is a membrane-bound immunoglobulin and it is this molecule that allows the distinction of B-cells from other types of lymphocyte, as well as being the main protein involved in B-cell activation. Once a B-cell encounters its cognate antigen and receives an additional signal from a helper T-cell, it can further differentiate into one of the two types of B-cells listed below [1]. The B-cell may either become one of these cell types directly or it may undergo an intermediate differentiation step, the germinal center reaction, where the B-cell will hypermutate the variable region of its immunoglobulin gene and possibly undergo class switching.

T-Cell Independent Activation

Many antigens are T-independent, meaning they can deliver both of the signals to the B-cell. Mice without a thymus (nude or athymic mice that do not produce any T-cells) can respond to T-independent antigens. Many bacteria have repeating carbohydrate epitopes that stimulate B-cells, through so called pattern recognition receptors, to respond with IgM synthesis in the absence of T-cell help. There are two types of T-cell independent activation; Type 1 T-cell-independent (polyclonal) activation and type 2 T-cell-independent activation (in which macrophages present several of the same antigen in a way that causes cross-linking of antibodies on the surface of B-cells).

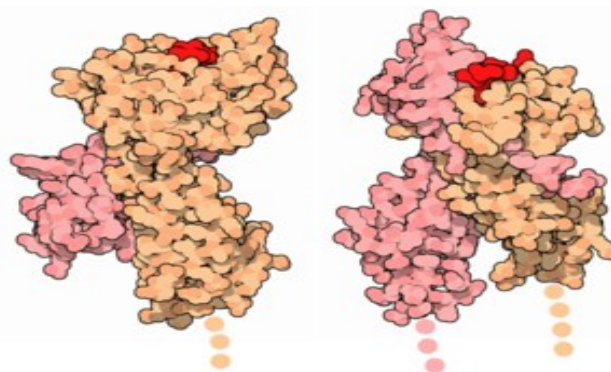


Fig. 2- Protein images comparing the MHC I (1hsa) and MHC II (1dlh) molecules.

The major histocompatibility complex (MHC) is a large genomic region or gene family found in most vertebrates [12]. It is the most gene-dense region of the mammalian genome and plays an important role in the immune system, autoimmunity and reproductive success. The proteins encoded by the MHC are expressed on the surface of cells in all jawed vertebrates and display fragments of molecules from invading microbes or dysfunctional cells (e.g. tumor cells) to a particular type of white blood cell called a T-cell that has the capacity to kill or co-ordinate the killing of the microbe, infected cell or malfunctioning cell.

HLA Genes

The best-known genes in the MHC region are the subset that encodes cell-surface antigen-presenting proteins. In humans, these genes are referred to as human leukocyte antigen (HLA) genes, although people often use the abbreviation MHC to refer to HLA gene products. To clarify the usage, some of the biomedical literature uses HLA to refer specifically to the HLA protein molecules and reserves MHC for the region of the genome that encodes for this molecule; however this convention is not consistently adhered to. The most intensely-studied HLA genes are the nine so-called classical MHC genes: HLA-A, HLA-B, HLA-C, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1, HLA-DRA and HLA-DRB1. In humans, the MHC is divided into three regions: Class I, II and III. The A, B and C genes belong to MHC class I, whereas the six D genes belong to class II [5]. Besides being scrutinized by immunologists for its pivotal role in the immune system, the MHC has also attracted the attention of many evolutionary biologists, due to the high levels of allelic diversity found within many of its genes. Indeed, much theory has been devoted to explaining why this particular region of the genome harbors so much diversity, especially in light of its immunological importance [3].

MHC Class I

MHC class I molecules are found on almost every nucleated cell of the body. MHC class I molecules are heterodimers, consisting of a single transmembrane polypeptide chain (the α -chain) and a β_2 microglobulin (which is encoded elsewhere, not in the MHC). The α chain has two polymorphic domains, α_1 , α_2 , which binds peptides derived from cytosolic proteins. Because MHC class I molecules present peptides derived from cytosolic proteins, the pathway of MHC class I presentation is often called the cytosolic or endogenous pathway. The peptides are mainly generated in the cytosol by the proteasome [16]. The proteasome is a macromolecule that consists of 28 subunits, of which half of them contain proteolytic activity. The proteasome degrades intracellular proteins into small peptides that are then released into the cytosol. The peptides have to be translocated from the cytosol into the endoplasmic reticulum (ER) to meet the MHC class I molecule, whose peptide-binding site is in the lumen of the ER. The peptide translocation from the cytosol into the lumen of the ER is accomplished by the transporter associated with antigen processing (TAP). TAP is a member of the ABC transporter family and is a heterodimeric multimembrane-spanning polypeptide consisting of TAP1 and TAP2 [16]. The two subunits form a peptide binding site and two ATP binding sites that face the lumen of the cytosol. TAP binds peptides on the cytoplasmic site and translocates them under ATP consumption into to the lumen of the ER. The MHC

class I molecule is then in turn loaded with peptides in the lumen of the ER. The peptide-loading process involves several other molecules that form a large multimeric complex consisting of TAP, tapasin, calreticulin, calnexin and ERP57.

Once the peptide is loaded onto the MHC class I molecule, it leaves the ER through the secretory pathway to reach the cell surface. The transport of the MHC class I molecules through the secretory pathway involves several posttranslational modifications of the MHC molecule. Some of the posttranslational modifications occur in the ER and involve change to the N-glycan regions of the protein, followed by extensive changes to the N-glycans in the Golgi apparatus. The N-glycans mature fully before they reach the cell surface. Peptides that fail to bind MHC class I molecules in the lumen of the endoplasmic reticulum are removed from the ER via the sec61 channel into the cytosol, where they might undergo further trimming in size and might be translocated by TAP back into ER for binding to an MHC class I molecule [18]. MHC class I molecules are loaded with proteins generated in the cytosol. As viruses infect a cell by entering its cytoplasm, this cytosolic, MHC class I-dependent pathway of antigen presentation is the primary way for a virus-infected cell to signal T-cells. MHC class I molecules generally interact exclusively with CD8⁺ ("cytotoxic") T-cells (CTLs). The fate of the virus-infected cell is almost always apoptosis initiated by the CTL, effectively reducing the risk of infecting neighboring cells.

MHC Class II

MHC Class II molecules are found only on a few specialized cell types, including macrophages, dendritic cells, activated T-cells and B-cells, all of which are professional antigen-presenting cells (APCs). Like MHC class I molecules, class II molecules are also heterodimers, but in this case consist of two homologous peptides, an α and β chain, both of which are encoded in the MHC [23]. The peptides presented by class II molecules are derived from extracellular proteins (not cytosolic as in class I); hence, the MHC class II-dependent pathway of antigen presentation is called the endocytic or exogenous pathway. Loading of class II molecules must still occur inside the cell; extracellular proteins are endocytosed, digested in lysosomes and bound by the class II MHC molecule prior to the molecule's migration to the plasma membrane. Because the peptide-binding groove of MHC class II molecules is open at both ends while the corresponding groove on class I molecules is closed at each end, the peptides presented by MHC class II molecules are longer, generally between 15 and 24 amino acid residues long.

Because class II MHC is loaded with extracellular proteins, it is mainly concerned with presentation of extracellular pathogens (for example, bacteria that might be infecting a wound or the blood). Class II molecules interact exclusively with CD4⁺ ("helper") T-cells (T_HS). The helper T-cells then help to trigger an appropriate immune response which may include localized inflammation and swelling due to recruitment of phagocytes.

Materials and Methods

Entrez Gene

Entrez Gene is NCBI's database for gene-specific information. Entrez Gene includes records from genomes that have been completely sequenced, that have an active research community to

contribute gene-specific information or that are scheduled for intense sequence analysis [4]. The content of Entrez Gene represents the result of both curation and automated integration of data from NCBI's Reference Sequence project (RefSeq), from collaborating model organism databases and from other databases within NCBI Records.

Antigenic Peptide Prediction

Antibodies (Abs) find multiple applications in a variety of areas including biotechnology, medicine and diagnosis and indeed they are one of the most powerful tools for life science research. Abs directed against protein antigens can recognize either linear or native three-dimensional (3D) epitopes. The abstention of Abs that recognize 3D epitopes require the use of whole native protein as immunogens. Unfortunately, this not always a choice due to various technical reasons: for example the native protein is just not available, or the protein is toxic. In such cases, immunization with peptides is the alternative. Of course, Abs generated in this manner will recognize linear epitopes and they might or might not recognize the source native protein, but yet they will be useful for standard laboratory applications such as western blots. EMBOSS, The European Molecular Biology Open Software Suite. An open source project started by the EMBnet community in order to replace proprietary systems like GCG.

MHC Class-I MHC Class-II Prediction

A MHC plays an essential role the immune system. A MHC molecule binds a peptide derived from an antigen and then displays it on a cell surface for recognition for T-cells. Hence, it is important for therapy of autoimmune diseases and cancer to determine which peptides can bind to given a MHC molecule. However it is difficult to predict the peptides bind to MHC class II molecules because of various amino acid lengths of them.

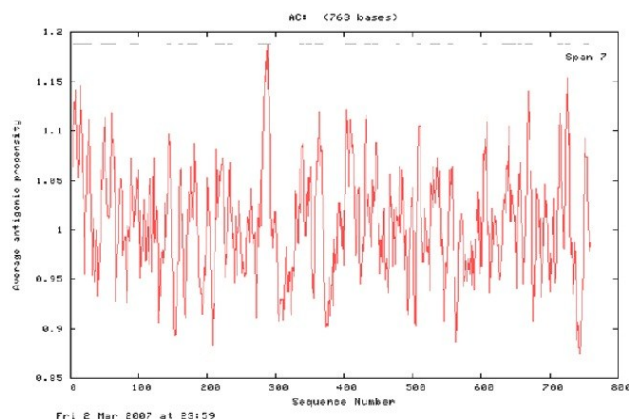
HLAPred

The MHC Class II peptide binding was predicted using HLAPred. This method is used to predict the HLA binding alleles for antigenic sequence. MHC-peptide complexes are potential tools for diagnosis and treatment of pathogens and cancer, as well as for the development of peptide vaccines [17]. Only one in 100 to 200 potential binders actually binds to a certain MHC molecule, therefore a good prediction method for MHC class II binding peptides can reduce the number of candidate binders that need to be synthesized and tested. These predicted HLA-binder's were further compared with Antigenic peptide which was predicted using the Antigenic peptide prediction server.

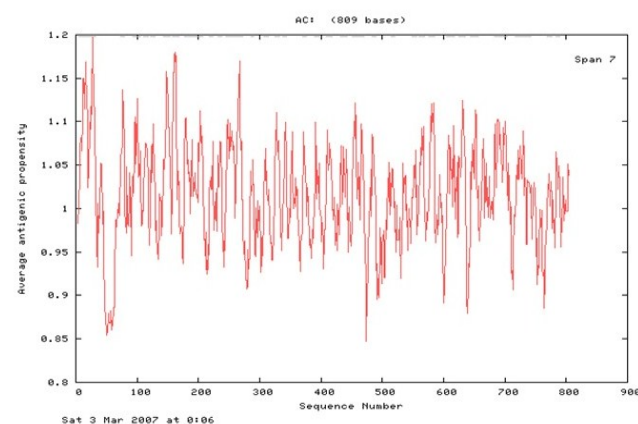
Results and Discussions

The antigenic peptide of Protective antigen and lethal factor were identified using Antigenic peptide prediction tool. This tool predicted the peptide according the basic amino acid which lies on solvent assessable nature with lipophilicity. The Antigenic plot was plotted according to windows size 7 with an average of all amino acids.

27 Antigenic peptides were obtained from protective antigen of *Bacillus anthracis*. The longest antigenic peptide length found 23 amino acids and the shortest antigenic peptide shows the length of 7 amino acids.



Graph 1- Antigenic Plot for Protective antigen



Graph 2- Antigenic Plot for lethal factor

Table 1- Antigenic peptide predicted for Protective antigen

| S. No. | Start Position | Sequence | End Position |
|--------|----------------|------------------------|--------------|
| 1 | 4 | RKVLPLMALSTLVSS | 20 |
| 2 | 22 | GNLEVIQAE | 30 |
| 3 | 44 | SSQGLLGYFSDLNFQAPMVVTS | 66 |
| 4 | 88 | YFQSAIWSGFIKV | 100 |
| 5 | 141 | RLYQIKI | 147 |
| 6 | 159 | LDFKLYW | 165 |
| 7 | 171 | KKEVISSDNLQLPEL | 185 |
| 8 | 212 | IPDSLEVEGYTVDV | 225 |
| 9 | 230 | TFLSPWI | 236 |
| 10 | 276 | VSPEARHPLVAAYPIVHV | 293 |
| 11 | 334 | GNAEVHASF | 342 |
| 12 | 345 | IGGSVSA | 351 |
| 13 | 358 | SSTVAIDHLSL | 369 |
| 14 | 402 | PIYNVLPPTTSLVLGK | 416 |
| 15 | 429 | QLSQILAPNN | 438 |
| 16 | 441 | PSKNLAPIAL | 450 |
| 17 | 475 | KQLRLDTD | 482 |
| 18 | 506 | WSEVLPQ | 512 |
| 19 | 526 | DLNLVERRIAAVNP | 539 |
| 20 | 552 | LKEALKI | 558 |
| 21 | 602 | NIYTVLDK | 609 |
| 22 | 632 | IAVGADESVKEAHREVIN | 650 |
| 23 | 653 | TEGLLL | 659 |
| 24 | 664 | IRKILSGYIVE | 674 |
| 25 | 712 | DKLPLYIS | 719 |
| 26 | 723 | YKVNYYAV | 730 |
| 27 | 750 | IKKILIFS | 757 |

36 Antigenic peptides were obtained from Lethal factor of *Bacillus anthracis*. The longest antigenic peptide length found 29 amino acids and the shortest antigenic peptide shows the length of 7 amino acids.

Peptides were constructed using HyperChem. These peptides were further optimized using Molecular mechanics according to the MM+ Method. Further the properties of these peptides were also identified according to conjugate gradient method. Energy minimization of the model has been carried out using HyperChem and its given the energy of -19725 KJ/mol [Fig. 3].

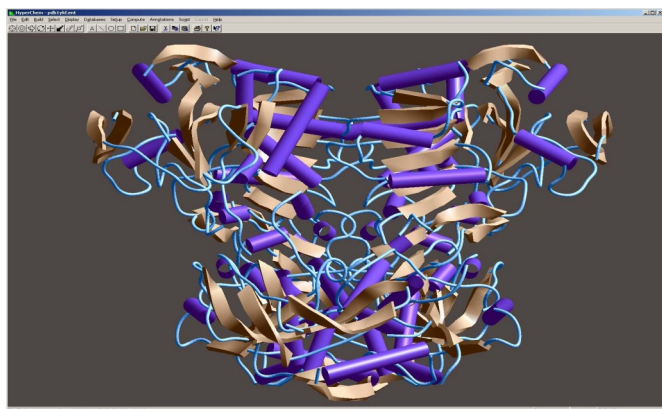


Fig. 3- Energy Minimization of peptides using HyperChem

Table 2- Antigenic Peptide predicted for Lethal factor

| S. No. | Start Position | Sequence | End Position |
|--------|----------------|-------------------------------|--------------|
| 1 | 5 | KEFIKVISMSCLVTAITLSGPVFIPLVQG | 33 |
| 2 | 73 | MKHIVKI | 79 |
| 3 | 92 | AEKLEKVPDVL | 106 |
| 4 | 111 | GGKIYVD | 118 |
| 5 | 122 | TKHISLEA | 129 |
| 6 | 142 | GKDALLHEHYVYA | 154 |
| 7 | 157 | GYEPVLVIQSSE | 168 |
| 8 | 176 | KALNVYIEIGK | 186 |
| 9 | 188 | LSRDILSKINQPYQKFLDVLN | 208 |
| 10 | 218 | GQDLLFT | 224 |
| 11 | 232 | TDFSVEF | 238 |
| 12 | 245 | EVQEVFAKAFAYYIE | 259 |
| 13 | 261 | QHRDVLQLYAPE | 272 |
| 14 | 285 | INLSLEE | 291 |
| 15 | 307 | IKQHYQHWS | 315 |
| 16 | 325 | LLKLLQPI | 333 |
| 17 | 338 | DDIHS | 344 |
| 18 | 370 | FLKLLQI | 376 |
| 19 | 389 | LLNRIQVDS | 397 |
| 20 | 408 | FLKLLKLDIQPY | 419 |
| 21 | 431 | LIDSPSINLDVRK | 443 |
| 22 | 451 | NIDALLHQSIGS | 462 |
| 23 | 464 | LYNKIYL | 470 |
| 24 | 483 | LGADLVD | 489 |
| 25 | 514 | SNYMIVD | 520 |
| 26 | 548 | ENGLILQR | 556 |
| 27 | 560 | LEIKDVQIIK | 569 |
| 28 | 576 | IRIDAKVVPK | 585 |
| 29 | 605 | KALGLPKYTKLITF | 618 |
| 30 | 625 | ASNIVESAYLIL | 636 |
| 31 | 644 | QSDLIKKVTNYLV | 656 |
| 32 | 682 | IYEQVHSGKGLYVPESRSILLHGP | 704 |
| 33 | 717 | FIHEFGHAVDDYAGYLLD | 734 |
| 34 | 737 | QSDLVTN | 743 |
| 35 | 769 | FFAEAFRL | 776 |
| 36 | 784 | ERLKVQK | 790 |

Table 3- Optimized energy for Protective antigen

| Allele | Energy | Gradient |
|------------|------------|-----------|
| HLA-B35 | 5.7946 | 0.091674 |
| HLA-B7 | -55.742439 | 0.082098 |
| HLA-B3701 | 29.024731 | 0.092768 |
| HLA-B44 | -11.830387 | 0.092768 |
| HLA-B27 | -61.650208 | 0.099309 |
| HLA-B2703 | -38.207798 | 0.095461 |
| HLA-2902 | -44.597549 | 0.094285 |
| HLA-B2705 | -47.368973 | 0.095407 |
| HLA-A0202 | -53.444164 | 0.0939926 |
| HLA-A0202A | -32.227783 | 0.0936 |

Table 4- Optimized energy for Lethal factor

| Allele | Energy | Gradient |
|-----------|------------|-----------|
| HLA-B40 | -13.413716 | 0.099231 |
| HLA-B35 | -37.245682 | 0.09999 |
| HLA-A31 | -34.118382 | 0.0940915 |
| HLA-A0203 | -66.91198 | 0.097274 |
| HLA-A0206 | -76.449951 | 0.093897 |

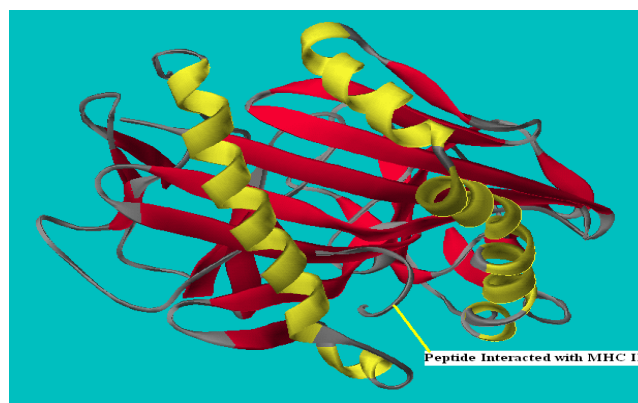


Fig. 4- Peptide docked with MHC Class II



Fig. 5- Structure of MHC Class II

The protein peptide interaction was identified using BiomedCache. The software predicts the interaction according to PMF (Potential of Mean Force) Scoring function. The active of the MHC Protein was predicted between the peptide cleft of α_1 and α_2 chain. All the peptides were interacted with these complex and the peptides which show least energy with high affinity was considered for further studies.

Table 5- Docking energy score for Protective antigen

| Allele | Docking Score |
|-----------|---------------|
| HLA-0202A | 1032666.5 |
| HLA-0202 | 10470.716 |
| HLA-B2703 | 197899.36 |
| HLA-B35 | 22389.866 |
| HLA-B7 | 4566.798 |
| HLA-B3701 | -19.823 |
| HLA-B27 | 30161.463 |
| HLA-B44 | -50.888 |
| HLA-B2902 | 35862.879 |
| HLA.B2705 | 29548.098 |

Table 6- Docking energy score for Lethal factor

| Allele | Docking Score |
|-----------|---------------|
| HLA-A31 | 16114.568 |
| HLA-A0203 | 37500.426 |
| HLA-A0206 | 67554.273 |
| HLA-B35 | 10512.686 |
| HLA-B40 | 41099.267 |

Conclusion

Epitopes are responsible to elicit antigen-antibody reactions. These Epitopes could be considered as the sequence which produces memory cells for destroying the pathogens. These are considered as vaccines. Using in-silico analysis we identified 10 epitopes for protective antigen and 5 Epitopes from lethal factor. These showed best epitopic regions, when compared with several tools. In-vitro analysis can be done on these sequences to confirm the efficiency. Further studies can be continue with software or tool can be developed to predict population coverage of T-cell epitope based on MHC binding data. Accordingly, epitope-based vaccines or diagnostics can be designed to maximize population coverage.

References

- [1] Abboud N., De Jesus M., Nakouzi A., Cordero R.J., Pujato M., Fiser A., Rivera J., Casadevall A. (2011) *J. Biol. Chem.*, 284 (37), 25077-25086.
- [2] Alexander F. Kintzer, Harry J. Sterling, Iok I. Tang, Evan R. Williams and Bryan A. Krantz (2010) *PLoS One*. 8, 5(11).
- [3] Ballard J.D., Collier R.J., Starnbach M.N. (1996) *Proc. Natl. Acad. Sci. USA.*, 93(22), 12531-12534.
- [4] Benson D.A., Karsch-Mizrachi I., Clark K., Lipman D.J., Ostell J., Sayers E.W. (2012) *Nucleic Acids Research*, D48-53.
- [5] Brittingham K.C., Ruthel G., Panchal R.G., Fuller C.L., Ribot W.J., Hoover T.A., Young H.A., Anderson A.O., Bavari S. (2005) *J. Immunol.*, 174(9), 5545-5552.
- [6] Chen D., Ma L., Kanalas J.J., Gao J., Pawlik J., Jimenez M.E., Walter M.A., Peterson J.W., Gilbertson S.R., Schein C.H. (2012) *Bioorg. Med. Chem.*, 20(1), 368-76.
- [7] Chen G., Driks A., Tawfiq K., Mallozzi M., Patil S. (2010) *Colloids Surf B Biointerfaces*. 76(2), 512-518.
- [8] Chow E.M., Batty S., Mogridge J. (2010) *Cell Microbiol.*, 12(4), 557-568.
- [9] Cui X., Su J., Li Y., Shiloach J., Solomon S., Kaufman J.B., Mani H., Fitz Y., Weng J., Altaweel L., Besch V., Eichacker P.Q. (2010) *Intensive Care Med.* 36(1), 148-56.
- [10] DelVecchio V.G., Sabato M.A., Trichilo J., Dake C., Grewal P., Alefantis T. (2010) *Crit. Rev. Immunol.* 30(3), 239-254.
- [11] Dupuy L.C., Schmaljohn C.S. (2009) *Expert Rev. Vaccines.*, 8 (12), 1739-1754.
- [12] Fowler K., McBride B.W., Turnbull P.C., Baillie L.W. (2009) *J. Appl. Microbiol.*, 87(2), 305.
- [13] Froude J.W., Thullier P., Pelat T. (2011) *Toxins (Basel)*, 3(11), 1433-1452.
- [14] Han Y., Chen F., Li N., Zhu B., Li X. (2010) *Int. J. Mol. Sci.*, 11 (2), 507-20.
- [15] Huang J.M., La Ragione R.M., Cooley W.A., Todryk S., Cutting S.M. (2008) *Vaccine*, 26(48), 6043-6052.
- [16] Leppla S.H., Arora N., Varughese M. (1999) *J. Appl. Microbiol.*, 87(2), 284.
- [17] Li Z., Zhang M., Hu H., Liu S., Lu Z. (2010) *On predicting the T-cell and B-cell epitopes of platelet membrane glycoprotein II b/ III a antibody from human and mice.*, 27(5), 1146-51.
- [18] McEvers K., Elrefaei M., Norris P., Deeks S., Martin J., Lu Y., Cao H. (2005) *Vaccine.*, 23(32), 4128-4135.
- [19] Paccani S.R., Baldari C.T. (2011) *Toxins(Basel)*, 3(6), 660-71.
- [20] Pavan M.E., Pettinari M.J., Cairó F., Pavn E.E., Cataldi A.A. (2011) *Bacillus anthracis: a molecular look at a famous pathogen.*, 43(4), 294-310.
- [21] Popescu R., Pistol A., Miltaru L., Caplan D., Cucuiu R., Popovici F. (2011) *Euro. Surveill.*, 10, 16(45).
- [22] Sinha K., Bhatnagar R. (2010) *Mol. Immunol.* 48(1-3), 264-71.
- [23] Westwood A., Healey G.D., Williamson E.D., Eyles J.E. (2005) *Vaccine*, 23(29), 3857-3863.