IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

INSILICO T-CELL EPITOPE BASED VACCINE DESIGN AGAINST HEPATITIS VIRUS

[1]Deepala. Vasantha Vidya Lakshmi,

[1]Lecturer

[1]Department of Biotechnology,

[1]Sai Sudhir Degree & PG College, ECIL x roads, Hyderabad,India

Abstract : Hepatitis is a serious health risk now-a-days and needs advancement in its treatment against increasing problems of the drug efficacy and the existence of the resistant virus. Hepatitis refers to the inflammation of the liver. It is commonly caused by a virus. Viral infections of the liver include Hepatitis A, B, C, D, E. A different type of virus is responsible for causing various types of Hepatitis disease. Various vaccines are available for treating the Hepatitis. In the present study, the problems of designing an efficient drug were solved using various computer aided tools. With the help of tools like NetCTL and IEDB(Immune Epitope Database and Analysis Resource), various epitopes have been identified, that can be used as ligands for the construction of a drug. Three-dimensional modelling of the selected epitope sequences was done using PEPFOLD and PHYRE softwares.

I. Introduction

Hepatitis refers to an inflammatory condition of the liver. It's commonly caused by a viral infection, but there are other possible causes of hepatitis. These include autoimmune hepatitis and hepatitis that occurs as a secondary result of medications, drugs, toxins, and alcohol. Autoimmune hepatitis is a disease that occurs when your body makes antibodies against your liver tissue.

Your liver is located in the right upper area of your abdomen. It performs many critical functions that affect metabolism throughout your body, including:

- bile production, which is essential to digestion
- filtering of toxins from your body
- excretion of bilirubin (a product of broken-down red blood cells), cholesterol, hormones, and drugs
- breakdown of carbohydrates, fats, and proteins
- activation of enzymes, which are specialized proteins essential to body functions
- storage of glycogen (a form of sugar), minerals, and vitamins (A, D, E, and K)
- synthesis of blood proteins, such as albumin
- synthesis of clotting factors

According to the Centers for Disease Control and Prevention (CDC)Trusted Source, approximately 4.4 million Americans are currently living withchronic hepatitis B and C. Many more people don't even know that they have hepatitis. Treatment options vary depending on which type of hepatitis you have. You can prevent some forms of hepatitis through immunizations and lifestyle precautions.

Types of viral hepatitis

Viral infections of the liver that are classified as hepatitis include hepatitis A, B, C, D, and E. A different virus is responsible for each type of virally transmitted hepatitis.

Hepatitis A is always an acute, short-term disease, while hepatitis B, C, and D are most likely to become ongoing and chronic. Hepatitis E is usually acute but can be particularly dangerous in pregnant women.

Hepatitis A

HepatitisA or infectious jaundice is caused by hepatitis A virus (HAV) {Ryan KJ,et.al(2004), Sherris Medical Microbiology (4th ed.).}, a picornavirus transmitted by the fecal-oral route often associated with ingestion of contaminated food. It causes an acute form of hepatitis and does not have a chronic stage. The patient's immune system makes antibodies against HAV that confer immunity against future infection. People with hepatitis A are advised to rest, stay hydrated and avoid alcohol. A vaccine is available that will prevent HAV infection for up to 10 years. Hepatitis A can be spread through personal contact, consumption of raw sea food, or drinking contaminated water. This occurs primarily in third world countries. Strict personal hygiene and the avoidance of raw and unpeeled foods can help prevent an infection. Infected people excrete HAV with their feces two weeks before and one week after the appearance of jaundice. The time between the infection and the start of the illness averages 28 days (ranging from 15 to 50 days), and most recover fully within

2 months, although approximately 15% of sufferers may experience continuous or relapsing symptoms from six months to a year following initial diagnosis.

Hepatitis A

Marker	**Detection Time**	**Description**	**Significance**
Faecal HAV	2–4 weeks or 28 days	–	Early detection
Ig M anti HAV	4–12 weeks	Enzyme immunoassay for antibodies	During acute Illness
Ig G anti HAV	5 weeks– persistent	Enzyme immunoassay for antibodies	Old infection or reinfection

Table 1: Existing Markers for detection of Viral Hepatitis

Hepatitis B

Hepatitis B (HB) is an infectious disease caused by the hepatitis B virus (HBV) that affects the liver. [Logan, et. al (1987), Logan's Medical and Scientific Abbreviations. J.B.Lippincott and Company.p.232.]. HBV, a heap DNA virus that can cause both acute and chronic hepatitis. Chronichepatitis develops in the 15% of adults who are unable to eliminate the virus after an initial infection. Identified methods of transmission include contactwith blood, bloodtransfusion (nowrare),unsanitary tattoos, sex (throu gh sexual intercourse or contact with bodily fluids), or mother-to-child by breast feeding; there is minimal evidence of transplacental crossing. However, in about half of cases the source of infection cannot be determined. Blood contact can occur by sharing syringes in intravenous drug use, shaving accessories such as razor blades, or touching wounds on infected persons. Needle-exchange programmes have been created in many countries as a form of prevention.

Patients with chronic hepatitis B have antibodies against the virus, but not enough to clear the infected liver cells. The continued production of virus and countervailing antibodies is a likely cause of the immune complex disease seen in these patients. A vaccine is available to prevent infection for life. Hepatitis B infections result in 500,000 to 1,200,000 deaths per year worldwide due to the complications of chronic hepatitis, cirrhosis, and hepatocellular carcinoma. Hepatitis B is endemic in a number of (mainly South-East Asian) countries, making cirrhosis and hepatocellular carcinoma big killers. There are eight treatment options approved by the U.S. Food and Drug Administration (FDA) available for persons with a chronichepatitisBinfection: alpha-interferon, pegylated

interferon, adefovir, entecavir, telbivudine, lamivudine, tenofovir

disoproxil and tenofovir alafenamide with a 65% rate of sustained response.

Hepatitis C

Hepatitis C is an infectious disease caused by the hepatitis C virus (HCV) that primarily affects the liver.[Ryan KJ, Ray CG, eds. (2004). Sherris Medical Microbiology (4th ed.). McGraw Hill. pp. 551–2.]. Hepatitis C (originally "non- A non-B hepatitis") is caused by hepatitis C virus (HCV), an RNA virus of the family Flaviviridae. HCV can be transmitted through contact with blood including through sexual contact if the two parties' blood is mixed) and can also cross the placenta. Hepatitis C usually leads to chronic hepatitis, culminating in cirrhosis in some people. It usually remains asymptomatic for decades. Patients with hepatitis C are susceptible to severe hepatitis if they contract either hepatitis A or B, so all persons with hepatitis C should be immunized against hepatitis A and hepatitis B if they are not already immune, and avoid alcohol. HCV viral levels can be reduced to undetectable levels by a combination of interferon and the antiviral drug ribavirin. The genotype of the virus is the primary determinant of the rate of response to this treatment regimen, with genotype 1 being the most resistant.

Hepatitis C is the most common chronic blood-borne infection in the United States.

Hepatitis D

Hepatitis D (hepatitis delta) is a disease caused by the hepatitis delta virus (*HDV*), a small spherical enveloped virusoid[Magnius,L; T, et.al, "ICTV Virus Taxonomy Profile: Deltavirus". The Journal of General Virology. 99 (12): 1565–1566]. The *Hepatitis D virus* (HDV) or hepatitis delta agent belongs to the genus Deltavirus and causes Type D Hepatitis. It is similar to a viroid as it can only propagate in the presence of the hepatitis B virus, depending on the helper function of HBV for its replication and expression. It has no independent life cycle, but can survive and replicate as long as HBV infection persists in the host body. It can only cause infection when encapsulated by hepatitis B virus surface antigens.

Hepatitis E

Hepatitis E is inflammation of the liver caused by infection with the hepatitis E virus (HEV)[Kamar, et.al (2014), Clinical Microbiology Reviews. 27 (1): 116–138.].The *Hepatitis E virus* (HEV), from the family Hepeviridae, produces symptoms similar to hepatitis A, although it can take a fulminant course in some patients, particularly pregnant women (mortality rate about 20%); chronic infections may occur in immune-compromised patients. It is more prevalent in the Indian subcontinent. The virus is feco- orally transmitted and usually is self-limited.

Type of Hepatitis	**Gene ID**	**Protein ID**
A	MK829707	QCO31664.1
B	MK075117	QCS40651.1
C	MK527509	QCG74145.1
D	MH844625	QCC89118.1
E	LC436450	BBH51390.1

Table 2: Gene ID and Protein ID of Hepatitis Virus

Life Cycle

The life cycle of *Hepatitis B virus* is complex. Hepatitis B is one of a few known non-retroviral viruses which use reverse transcription as a part of its replication process.

Attachment

The virus gains entry into the cell by binding to receptors on the surface of the cell and entering it by endocytosis mediated by either clathrin or caveolin-1.[Zhang Z, et.al (July 2016), "Visualization of hepatitis B virus entry - novel tools and approaches to directly follow virus entry into hepatocytes".]. HBV initially binds to heparin sulfate proteoglycan. The pre-S1 segment of the HBV L protein then binds tightly to the cell surface receptor sodium taurocolatecotransporting polypeptide (NTCP), encoded by the SLC10A1gene [Yan H,et.al(September 2015) Antiviral Research. 121: 24–30.]. NTCP is mostly found in the sinusoidal membrane of liver cells. The presence of NTCP in liver cells correlates with the tissue specificity of HBV infection.

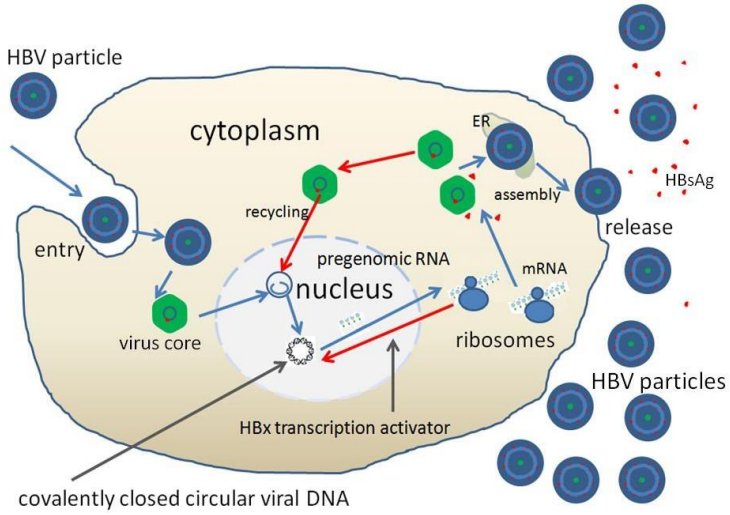


Figure1: Life Cycle of Hepatitis Virus

Penetration

Following endocytosis, the virus membrane fuses with the host cell's membrane, releasing the nucleocapsid into the cytoplasm. [Watashi K, Wakita T (August 2015), Cold Spring Harbor Perspectives in Medicine. 5].

Uncoating

Because the virus multiplies via RNA made by a host enzyme, the viral genomic DNA has to be transferred to the cell nucleus. It is thought the capsid is transported on the microtubules to the nuclear pore. The core proteins dissociate from the partially double stranded viral DNA, which is then made fully double stranded (by host DNA polymerases) and transformed into covalently closed circular DNA (cccDNA) that serves as a template for transcription of four viral mRNAs.

Replication

The largest mRNA, (which is longer than the viral genome), is used to make the new copies of the genome and to make the capsid core protein and the viral RNA-dependent-DNA-polymerase.

Assembly

These four viral transcripts undergo additional processing and go on to form progeny virions which are released from the cell or returned to the nucleus and re-cycled to produce even more copies. [Beck J, Nassal M (January 2007), World Journal of Gastroenterology. 13 (1): 48–64, Bruss V

(January 2007).]

Release

The long mRNA is then transported back to the cytoplasm where the virion P protein synthesizes DNA via its reverse transcriptase activity.

Symptoms and Complications

If you have infectious forms of hepatitis that are chronic, like hepatitis B and C, you may not have symptoms in the beginning. Symptoms may not occur until the damage affects liver function.

Signs and symptoms of acute hepatitis appear quickly. They include:

- fatigue
- flu-like symptoms
- dark urine
- pale stool
- abdominal pain
- loss of appetite
- unexplained weight loss
- yellow skin and eyes, which may be signs of jaundice

Chronic hepatitis develops slowly, so these signs and symptoms may betoo subtle to notice.

Treatment

Treatment options are determined by which type of hepatitis you have and whether the infection is acute or chronic.

Hepatitis A

Hepatitis A usually doesn't require treatment because it's a short-term illness. Bed rest may be recommended if symptoms cause a great deal of discomfort. If you experience vomiting or diarrhea, follow your doctor's orders for hydration and nutrition.

The hepatitis A vaccine is available to prevent this infection. Most children begin vaccination between ages 12 and 18 months. It's a series of two vaccines. Vaccination for hepatitis A is also available for adults and can be combined with the hepatitis B vaccine.

Hepatitis B

Acute hepatitis B infection does not usually require treatment and most adults clear the infection spontaneously. [Hollinger FB, Lau DT (December 2006).]

Chronic hepatitis B is treated with antiviral medications. [Lai CL, Yuen MF (July 2007)]. This form of treatment can be costly because it must be continued for several months or years. Treatment for chronic hepatitis B also requires regular medical evaluations and monitoring to determine if the virus is responding to treatment.

Hepatitis B can be prevented with vaccination. The CDC Trusted Source recommends hepatitis B vaccinations for all newborns. The series of three vaccines is typically completed over the first six months of childhood. The vaccine is also recommended for all healthcare and medical personnel.

Hepatitis C

Antiviral medications are used to treat both acute and chronic forms of hepatitis C. People who develop chronic hepatitis C are typically treated with a combination of antiviral drug therapies.Those with chronic hepatitis C are advised to avoid alcohol and medications toxic to the liver. [Wilkins, et.al (2006)]. They may also need further testing to determine the best form of treatment.

People who develop cirrhosis (scarring of the liver) or liver disease as a result of chronic hepatitis C may be candidates for a liver transplant.

Currently, there is no vaccination for hepatitis C.

Hepatitis D

No antiviral medications exist for the treatment of hepatitis D at this time. According to a 2013 studyTrusted Source, a drug called alpha interferon can be used to treat hepatitis D, but it only shows improvement in about 25 to 30 percent of people. [Yurdaydin C, Idilman R (August 2015)].

Hepatitis D can be prevented by getting the vaccination for hepatitis B, as infection with hepatitis B is necessary for hepatitis D to develop.

Hepatitis E

Currently, no specific medical therapies are available to treat hepatitis E. Because the infection is often acute, it typically resolves on its own. People with this type of infection are often advised to get adequate rest, drink plenty of fluids, get enough nutrients, and avoid alcohol. However, pregnant women who develop this infection require close monitoring and care.

Reviews of existing small studies suggest that ribavirin can be considered effective in immunocompromised people who have developed chronic infection.[Dalton, Harry R.; Kamar, Nassim (2016)].

2. *Vaccines - Introduction*

A **vaccine** is a biological preparation that provides active acquired immunity to a particular disease. A vaccine typically contains an agent that resembles a disease-causing microorganism and is often made from weakened or killed forms of the microbe, its toxins, or one of its surface proteins. The agent stimulates the body's immune system to recognize the agent as a threat, destroy it, and to further recognize and destroy any of the microorganisms associated with that agent that it may encounter in the future. Vaccines can be prophylactic (example: to prevent or ameliorate the effects of a future infection by a natural or "wild" pathogen), or therapeutic (e.g., vaccines against cancer are being investigated) [Melief CJ, et.al (September 2015)].

The administration of vaccines is called vaccination. Vaccination is the most effective method of preventing infectious diseases widespread immunity due to vaccination is largely responsible for the world wide eradication of smallpox and

the restriction of diseases such as polio, measles and tetanus from much of the world. The effectiveness of vaccination has been widely studied and verified; for example, vaccines that have proven effective include the influenza vaccine the HPV vaccine [Chang Y, et.al (July 2009)] and chicken pox vaccine [Liesegang TJ (August 2009)].

The terms *vaccine* and *vaccination* are from *Variolaevaccinae* (smallpox of the cow), the term devised by Edward Jenner to denote cowpox. He used it in 1798 in the long title of his *Inquiry into the Variolaevaccinae known as the Cow Pox*, in which he described the protective effect of cowpox against smallpox [Baxby D (January 1999)]. In 1881, to honor Jenner, Louis

Pasteur proposed that the terms should be extended to cover the new protective inoculations then being developed.

History

Prior to the introduction of vaccination with material from cases of cowpox (heterotypic immunization), smallpox could be prevented by deliberate inoculation of smallpox virus, later referred to as variolation to distinguish it from smallpox vaccination [Needham, Joseph. (2000)]. The earliest hints of the practice of inoculation for smallpox in China come during the 10th century. The Chinese also practiced the oldest documented use of variolation, dating back to the fifteenth century. They implemented a method of "nasal insufflation" administered by blowing powdered smallpox material, usually scabs, up the nostrils. Various insufflation techniques have been recorded throughout the sixteenth and seventeenth centuries within China. Two reports on the Chinese practice of inoculation were received by the Royal Society in London in 1700; one by Dr. Martin Lister who received a report by an employee of the East India Company stationed in China and another by Clopton Havers.

Sometime during the late 1760s whilst serving his apprenticeship as a surgeon/apothecary **Edward Jenner** learned of the story, common in rural areas, that dairy workers would never have the often-fatal or disfiguring disease smallpox, because they had already had cowpox, which has a very mild effect in humans. In 1796, Jenner took pus from the hand of a milkmaid with cowpox, scratched it into the arm of an 8-year-old boy, James Phipps, and six weeks later inoculated (variolated) the boy with smallpox, afterwards observing that he did not catch smallpox. Jenner extended his studies and in1798 reported that his vaccine was safe in children and adults and could be transferred from arm-to-arm reducing reliance on uncertain supplies from infected cows [Needham, Joseph. (2000)] Since vaccination with cowpox was much safer than smallpox inoculation, the latter, though still widely practised in England, was banned in 1840.

The second generation of vaccines was introduced in the 1880s by Louis Pasteur who developed vaccines for chicken cholera and anthrax [Pasteur L (1881)], and from the late nineteenth century vaccines were considered a matter of national prestige, and compulsory vaccination laws were passed.

The twentieth century saw the introduction of several successful vaccines, including those against diphtheria, measles, mumps, and rubella. Major achievements included the development of the polio vaccine in the 1950s and the eradication of smallpox during the 1960s and 1970s. Maurice Hilleman was the most prolific of the developers of the vaccines in the twentieth century. As vaccines became more common, many people began taking them for granted. However, vaccines remain elusive for many important diseases, including herpes simplex, malaria, gonorrhea, and HIV[Stern AM, Markel H (2005)].

Types of Vaccines

Vaccines are dead or inactivated organisms or purified products derived from them. There are several types of vaccines in use. These represent different strategies used to try to reduce the risk of illness while retaining the ability to induce a beneficial immune response.

Inactivated

Some vaccines contain inactivated, but previously virulent, micro-organisms that have been destroyed with chemicals, heat, or radiation.Examples include the polio vaccine, hepatitis A vaccine, rabies vaccine and some influenza vaccines.

Attenuated

Some vaccines contain live, attenuated microorganisms. Many of these are active viruses that have been cultivated under conditions that disable their virulent properties, or that use closely related but less dangerous organisms to produce a broad immune response. Although most attenuated vaccines are viral, some are bacterial in nature. Examples include the viral diseases yellow fever, measles, mumps, and rubella, and the bacterial disease typhoid. The live *Mycobacterium tuberculosis* vaccine developed by Calmette and Guérin is not made of a contagious strain but contains a virulently modified strain called "BCG" used to elicit an immune response to the vaccine. The live attenuated vaccine containing strain *Yersinia pestis* EV is used for plague immunization. Attenuated vaccines have some advantages and disadvantages. They typically provoke more durable immunological responses and are the preferred type for healthy adults. But they may not be safe for use in immunocompromised individuals, and on rare occasions mutate to a virulent form and cause disease[Sinha JK, Bhattacharya S]

Toxoid

Toxoid vaccines are made from inactivated toxic compounds that cause illness rather than the micro-organism. Examples of toxoid-based vaccines include tetanus and diphtheria. Toxoid vaccines are known for their efficacy. Not all toxoids are for micro-organisms; for example, *Crotalusatrox* toxoid is used to vaccinate dogs against rattlesnake bites.

Subunit

Protein subunit – rather than introducing an inactivated or attenuated micro- organism to an immune system (which would constitute a "whole-agent" vaccine), a fragment of it can create an immune response. Examples include the subunit vaccine against Hepatitis B virus that is composed of only the surface proteins of the virus (previously extracted from the blood serum of chronically infected patients, but now produced by recombination of the viral genes into yeast) or as an edible algae vaccine, the virus-like particle (VLP) vaccine against human papillomavirus (HPV) that is composed of the viral major capsid protein,andthe hemagglutinin and neuraminidase subunits of the influenza virus. Subunit vaccine is being used for plague immunization.

Conjugate

Conjugate – certain bacteria have polysaccharide outer coats that are poorly immunogenic. By linking these outer coats to proteins (e.g., toxins), the immune system can be led to recognize the polysaccharide as if it were a protein antigen. This approach is used in the *Haemophilus influenzae* type B vaccine.

Experimental

A number of innovative vaccines are also in development and in use:

- Dendritic cell vaccines combine dendritic cells with antigens in order to present the antigens to the body's white blood cells, thus stimulating an immune reaction. These vaccines have shown some positive preliminary results for treating brain tumors [Kim W, Liau LM (January 2010)] and are also tested in malignant melanoma.
- Recombinant vector – by combining the physiology of one micro- organism and the DNA of another, immunity can be created against diseases that have complex infection processes. An example is the RVSV-ZEBOV vaccine licensed to Merck that is being used in 2018 to combat ebola in Congo [McKenzie, David (26 May 2018)].

- DNA vaccination – an alternative, experimental approach to vaccination called *DNA vaccination*, created from an infectious agent's DNA, isunder development. The proposed mechanism is the insertion (and expression, enhanced by the use of electroporation, triggering immune system recognition) of viral or bacterial DNA into human or animal cells. Some cells of the immune system that recognize the proteins expressed will mount an attack against these proteins and cells expressing them. Because these cells live for a very long time, if the pathogen that normally expresses these proteins is encountered at a later time, they will be attacked instantly by the immune system. One potential advantage of DNA vaccines is that they are very easy to produce and store. As of 2015, DNA vaccination is still experimental and is not approved for human use.
- T-cell receptor peptide vaccines are under development for several diseases using models of Valley Fever, stomatitis, and atopic dermatitis. These peptides have been shown to modulate cytokine production and improve cell-mediated immunity.
- Targeting of identified bacterial proteins that are involved in complement inhibition would neutralize the key bacterial virulence mechanism [Meri S, Jördens M, (December 2008)].

While most vaccines are created using inactivated or attenuated compounds from micro-organisms, synthetic vaccines are composed mainly or wholly of synthetic peptides, carbohydrates, or antigens.

Effectiveness

There is overwhelming scientific consensus that vaccines are a very safe and effective way to fight and eradicate infectious diseases [Orenstein WA, et.al (1985). "Field evaluation of vaccine efficacy"]. Limitations to their effectiveness, nevertheless, exist. Sometimes, protection fails because the host's immune system simply does not respond adequately or at all. Lack of response commonly results from clinical factors such as diabetes, steroid use, HIV infection, or age. It also might fail for genetic reasons if the host's immune system includes no strains of B cells that can generate antibodies suited to reacting effectively and binding to the antigens associated with the pathogen.

Even if the host does develop antibodies, protection might not be adequate; immunity might develop too slowly to be effective in time, the antibodies might not disable the pathogen completely, or there might be multiple strains of the pathogen, not all of which are equally susceptible to the immune reaction. However, even a partial, late, or weak immunity, such as a one resulting from cross-immunity to a strain other than the target strain, may mitigate an infection, resulting in a lower mortality rate, lower morbidity, and faster recovery.

Adjuvants commonly are used to boost immune response, particularly for older people (50–75 years and up), whose immune response to a simple vaccine may have weakened.

The efficacy or performance of the vaccine is dependent on a number of factors:

- the disease itself (for some diseases vaccination performs better than for others)
- the strain of vaccine (some vaccines are specific to, or at least most effective against, particular strains of the disease)
- whether the vaccination schedule has been properly observed.
- idiosyncratic response to vaccination; some individuals are "non- responders" to certain vaccines, meaning that they do not generate antibodies even after being vaccinated correctly.
- assorted factors such as ethnicity, age, or genetic predisposition.

If a vaccinated individual does develop the disease vaccinated against (breakthrough infection), the disease is likely to be less virulent than in unvaccinated victims.

The following are important considerations in the effectiveness of a vaccination program:

1. careful modeling to anticipate the effect that an immunization campaign will have on the epidemiology of the disease in the medium to long term
2. ongoing surveillance for the relevant disease following introduction of a new vaccine
3. maintenance of high immunization rates, even when a disease has become rare.

In 1958, there were 763,094 cases of measles in the United States; 552 deaths resulted. After the introduction of new vaccines, the number ofcases dropped to fewer than 150 per year (median of 56). In early 2008, there were

64 suspected cases of measles. Fifty-four of those infections were associated with importation from another country, although only 13% were actually acquired outside the United States; 63 of the 64 individuals either had never been vaccinated against measles or were uncertain whether they had been vaccinated.

Vaccines led to the eradication of smallpox, one of the most contagious and deadly diseases in humans. Other diseases such as rubella, polio, measles, mumps, chickenpox, and typhoid are nowhere near as common as they were a hundred years ago thanks to widespread vaccination programs. As long as the vast majority of people are vaccinated, it is much more difficult for an outbreak of disease to occur, let alone spread. This effect is called herd immunity. Polio, which is transmitted only between humans, is targeted by an extensive eradication campaign that has seen endemic polio restricted to only parts of three countries (Afghanistan, Nigeria, and Pakistan). However, the difficulty of reaching all children as well as cultural misunderstandings have caused the anticipated eradication date to be missed several times. Vaccines also help prevent the development of antibiotic resistance. For example, by greatly reducing the incidence of pneumonia caused by *Streptococcus pneumoniae*, vaccine programs have greatly reduced the prevalence of infections resistant to penicillin or other first-line antibiotics.

Adverse effects

Vaccination given during childhood is generally safe. Adverse effects, if any, are generally mild. The rate of side effects depends on the vaccine in question. Some common side effects include fever, pain around the injection site, and muscle aches. Additionally, some individuals may be allergic to ingredients in the vaccine.MMR vaccine is rarely associated with febrile seizures [Maglione MA, et.al (August 2014)].

Severe side effects are extremely rare. Varicella vaccine is rarely associated with complications in immunodeficient individuals and rotavirus vaccines are moderately associated with intussusception.

Some countries such as the United Kingdom provide compensation for victims of severe adverse effects via its Vaccine Damage Payment. The United States has the National Childhood Vaccine Injury Act. At least 19 countries have such no-fault compensation.

3. *In silico Vaccine Design*

Infectious diseases caused by bacteria, viruses, fungi, and parasites cause millions of deaths worldwide every year. The efficacy of all known current anti-infective agents is affected by the drug resistant form of the pathogens. Therefore, the development of anti-infective drugs that target drug-resistant pathogens is very urgent. *In silico* models help to understand infectious diseases and develop novel therapeutics to treat them.

With the advent of computer-aided informatics and high-throughput technologies, vaccine research has entered a new era. Rational vaccine prediction is more reasonable than years ago because of the two progressive areas, vaccine database and *in silico* vaccine design models.

In silico models and databases play different but complementary roles in vaccine design. The database collects information about experimentally verified vaccine and vaccine components and *in silico* models use computational methods to predict and design new vaccine and the components.

In the development of small molecules, *in silico* model is important in genome-wide analysis, comparative genomics, pathway analysis, virtual screening and target identification. For vaccine development, *in silico* model can accelerate the computer identification algorithm of the relevant protein candidates with improved expression immunogens.

Advanced DNA sequencing and cellular, molecular and immunological methods provide a large number of vaccine-related data, which has led to an increase in the amount of data associated with vaccines and vaccinations. Advanced bioinformatics tools have been developed to make effective use of these data and vaccine-related papers has increased exponentially since then. Databases that store, reorganize, and classify these data, promotethe discovery of vaccine. Vaccine discovery, in turn, support the accumulation of vaccine data more efficiently in the database.

A variety of *in silico* models for vaccine design have been developed to predict T-cells and B-cell immune epitopes. After predicting candidate proteins, high-throughput assays can be used to evaluate efficacy of the vaccine.

4. *Objectives:*

In context of the previous sections, the objectives of this particular research are:

1. To identify and select homologous sequences in relation to the query nucleotide and protein sequence.
2. To compare both datasets and select the matching protein sequences with the nucleotide sequences.
3. To generate phylogenetic tree of the selected sequences.
4. To analyze the properties of the selected protein using in silico methods.
5. To predict the transmembrane regions, secondary structure content and secondary structure of the query protein using in silico methods.
6. To predict the three-dimensional (3D) structure of the query protein using homology modeling techniques and identify its structural motifs.
7. To generate the three-dimensional (3D) structure of the desired epitope sequence.

CHAPTER 2 MATERIALS AND METHODS

DATA RETRIVAL

The National Centre For Biotechnology Information(NCBI):

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information. Major databases include GenBank for DNA sequences and PubMed, a bibliographic database for the biomedical literature. Other databases include the NCBI Epigenomics database. All these databases are available online through the Entrez search engine.

Various genomic, protein and glycoprotein sequences of Hepatits Virus were identified using NCBI.(Fig:2.1)

URL Link:https://www.ncbi.nlm.nih.gov/

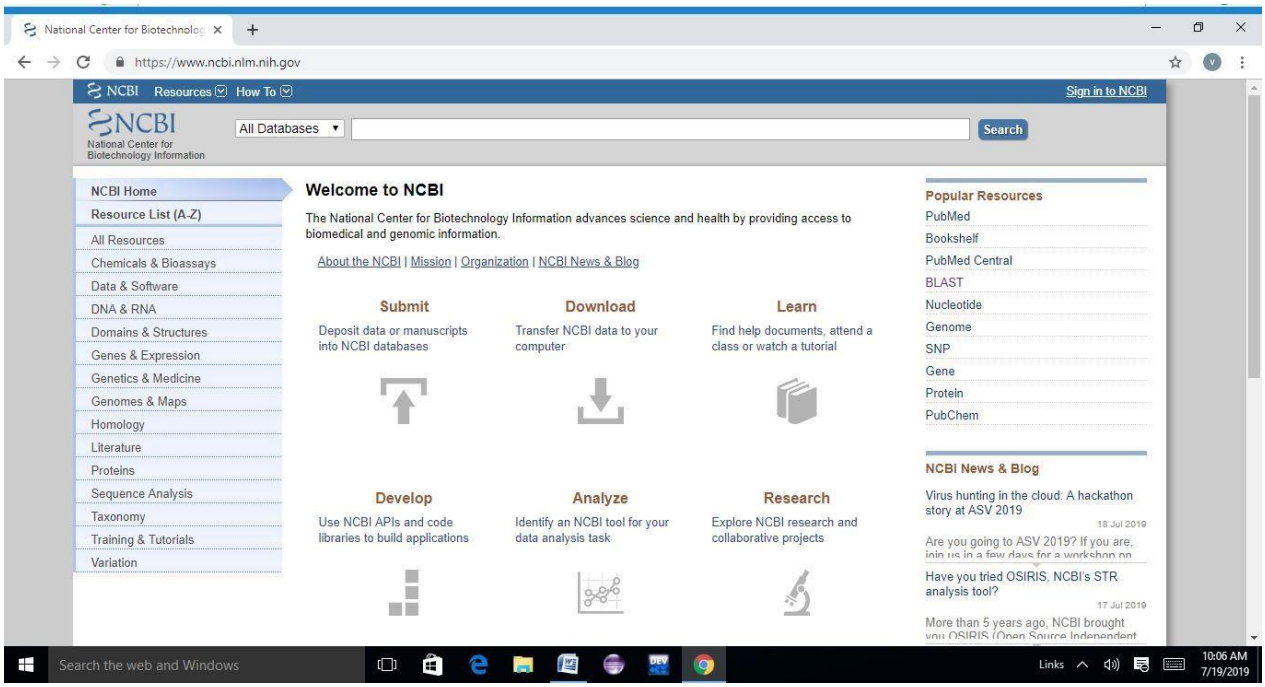


Fig:2.1 NCBI Homepage

EVOLUTIONARY ANALYSIS

CLUSTAL OMEGA

Clustal Omega is a completely rewritten and revised version of widely used Clustal series of programs for multiple sequence alignment. It can deal with very large numbers of DNA/RNA or protein sequences. The accuracy of the program has been considerably enhanced over earlier Clustal programs, through the use of the HHailgn method for aligning profile hidden Markov models. The program currently is used from the command line or can berun on line(Sievers and Higgins,2014)(Figure 2.2.1).

URL link: https://www.ebi.ac.uk/Tools/msa/clustalo/

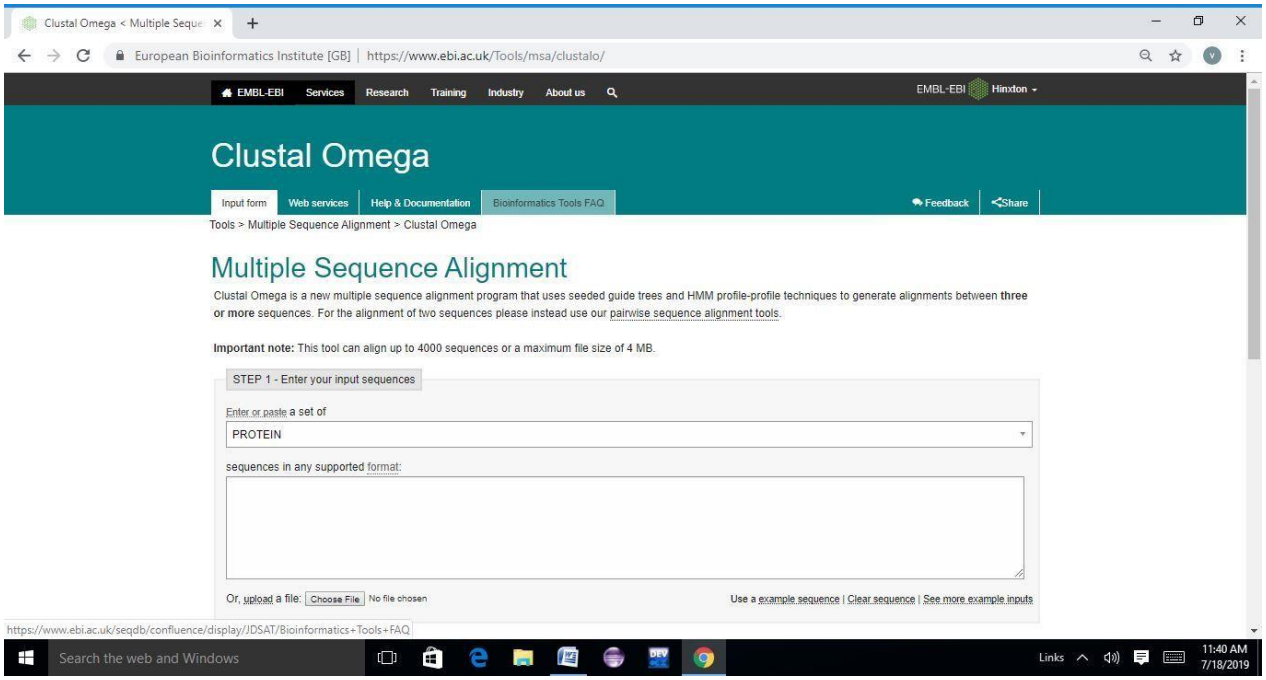


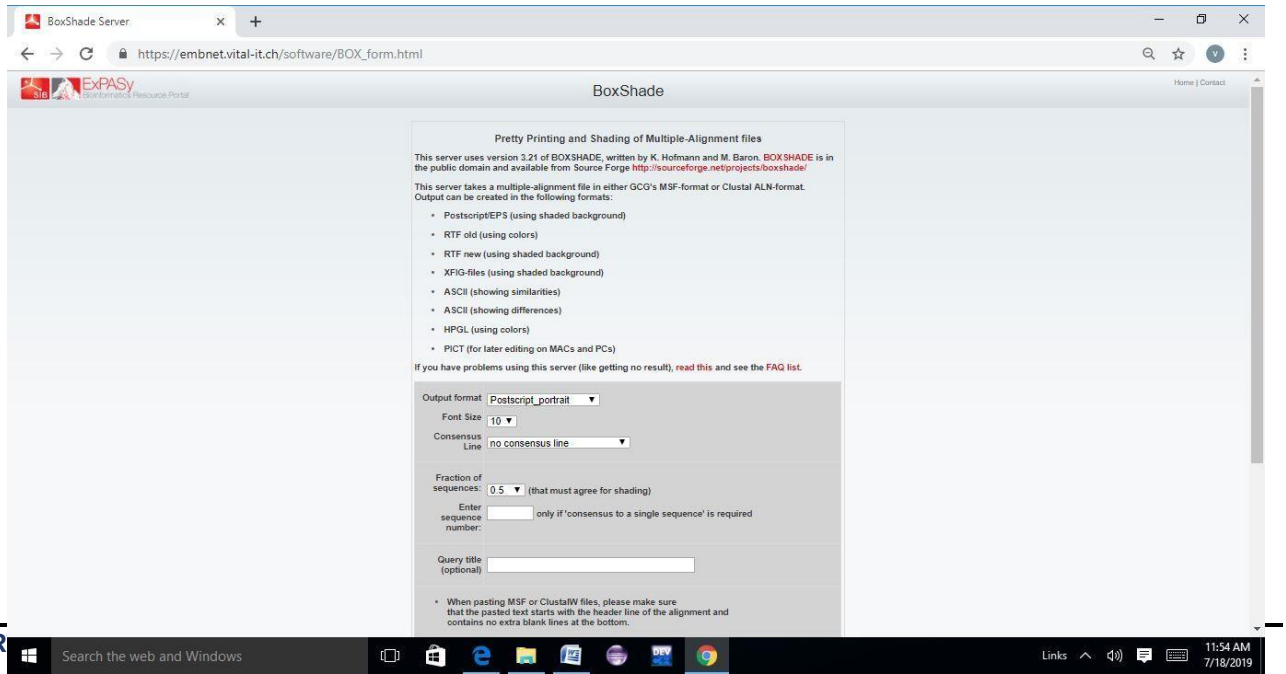
Figure 2.2.1: Clustal Omega homepage

BoxShade:

BoxShade is a program for pretty-printing multiple alignment output. The program itself does not carry out alignment of the selected nucleotide or protein sequences, as such, a multiple sequence alignment (MSA) programs like Clustal Omega or Clustal W2 needs to be used. Following so, the outputs of the programs are used as inputs for BoxShade to attain publishable images of the MSA results. The output format selected for the current study was RTF new(Figure 2.2.2)

URL link: https://embnet.vital-it.ch/software/BOX_form.html

Figure 2.2.2: BoxShade Homepage



Molecular Evolutionary Genetics Analysis(MEGA):

Molecular Evolutionary Genetics Analysis (MEGA) is an integrated tool for conducting sequence alignments, estimating divergence times, inferring phylogenetic trees, online database mining, molecular evolution rate estimation, inferring ancestral sequences and testing evolutionary hypotheses. It is used by biologists for reconstruction of evolutionary histories of species and hypothesizing/theorizing the extent and nature of the selective forces that shape the evolution of genes as well as species. The software is available online and can be downloaded (Figure 2.2.3)

URL link: https://www.megasoftware.net/

In this current study, MEGA X (Tamura et al., 2013) was used to generate phylogenetic trees for both nucleotide and protein sequences.

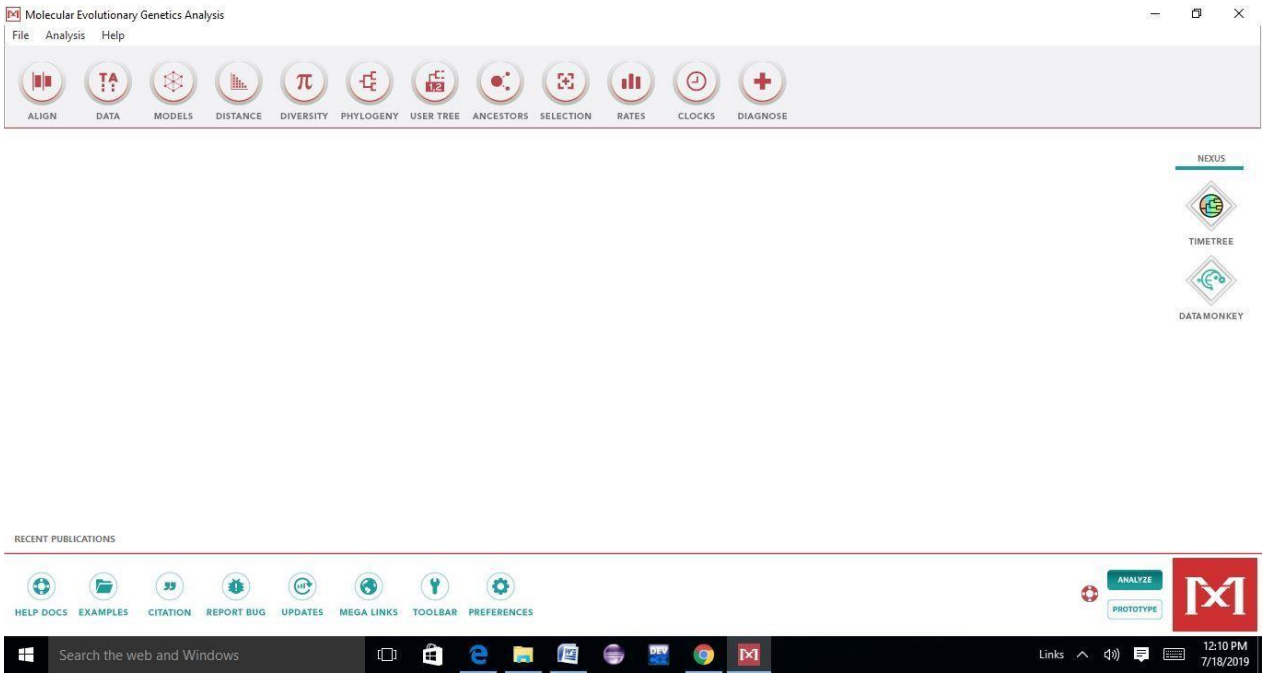


Figure 2.2.3: Mega X opening window

Nucleotide sequence data:

For the generation of nucleotide sequence based phylogenetic trees, in the input data section Nucleotide sequences were selected and it was confirmed as the proteincoding nucleotide sequence data. For the selection of genetic code, the Standard option was selected (Figure 2.2.3.1).

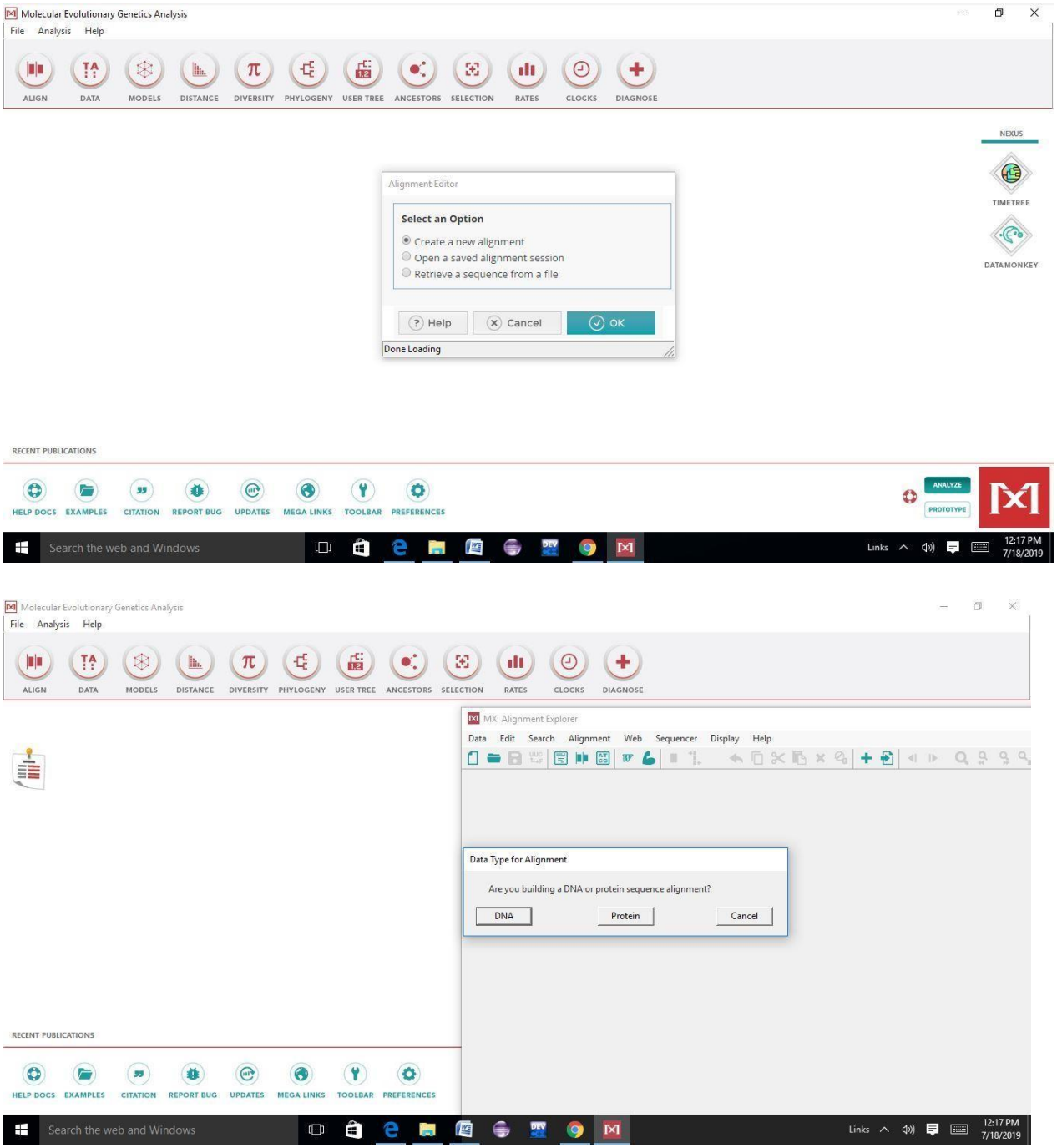


Figure 2.2.3.1: Steps taken for setting parameters for nucleotide sequence data.(a) Input data, (b)Confirmation

Using the Phylogeny option in MEGA 6, a Neighbor-Joining (NJ) phylogenetic tree was constructed. In the Analysis Preferences section of the software, in the statistical method, NJ method was selected. For the test of phylogeny the Bootstrap method was selected to determine the robustness with replicates set at 500. This was done as the NJ method does not have any clade support measure. The Nucleotide Substitutions type was selected. The model used was Maximum Composite Likelihood. Transitions and Transversions were selected as the substitutions to be included. For Rates and Patterns, the Rates were set to Gamma distributed (G) and the gamma parameter was set to 2. The Pattern among lineages was set to Homogenous and for gaps and missing data, Complete deletion was selected (Figure 2.2.3.1).

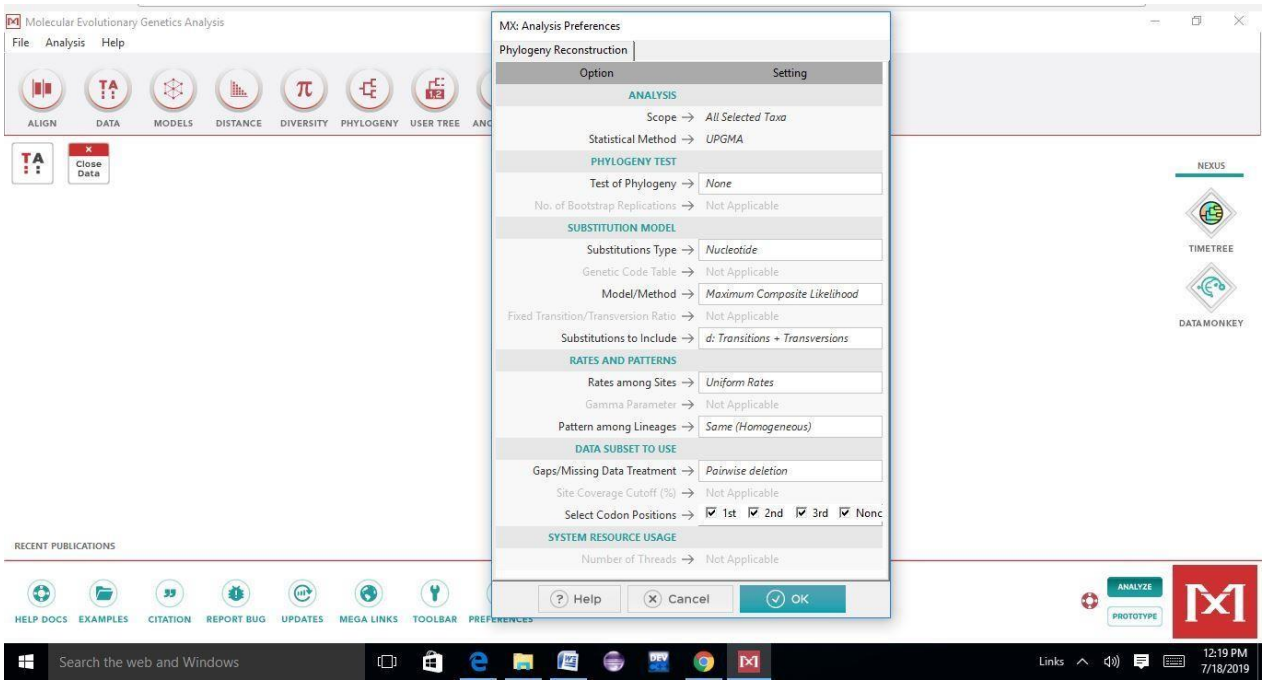


Figure 2.2.3.1(c): Parameters used for the contruction of phylogenetic trees for nucleotide sequence data

Protein sequence data:

For the generation of protein sequence based phylogenetic trees, in the input data section Protein sequences were selected (Figure 2.2.3.2).

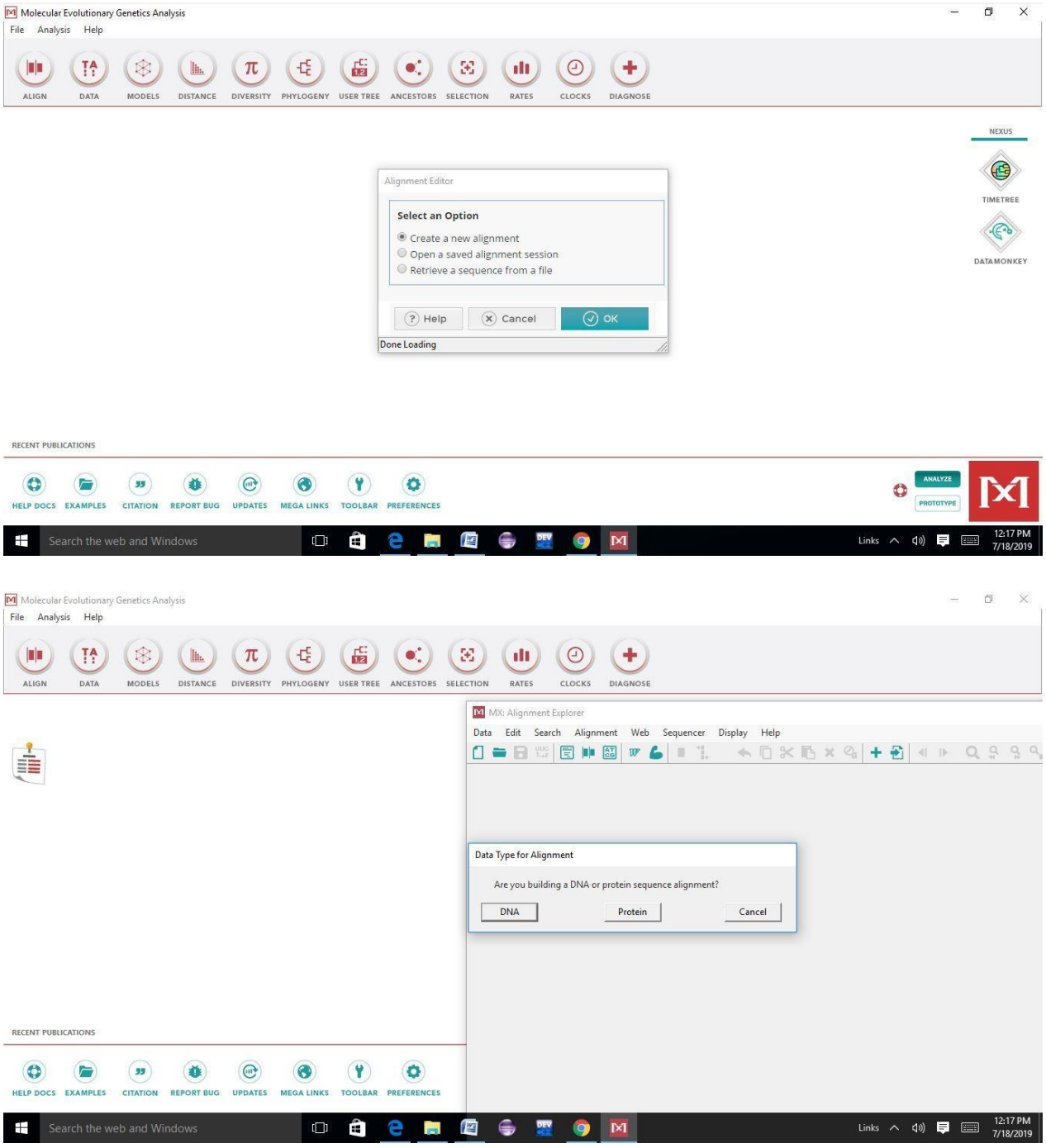


Figure 2.2.3.2: Input data type selection for protein data (a)Input data

(b) Confirmation

Using the Phylogeny option in MEGA 6, a Neighbor-Joining (NJ) phylogenetic tree was constructed. In the Analysis Preference section, in the statistical method NJ method was selected. The Bootstrap method was selected for the test of phylogeny with replicates set at 500. Amino acid Substitutions type was selected. The model used was Poisson model. For Rates and

Patterns, the Rates were set to Gamma distributed (G) and the gamma parameter was set to 2. The Pattern among lineages was set to Homogenous and for gaps and missing data, Complete deletion was selected (Figure 2.2.3.2).

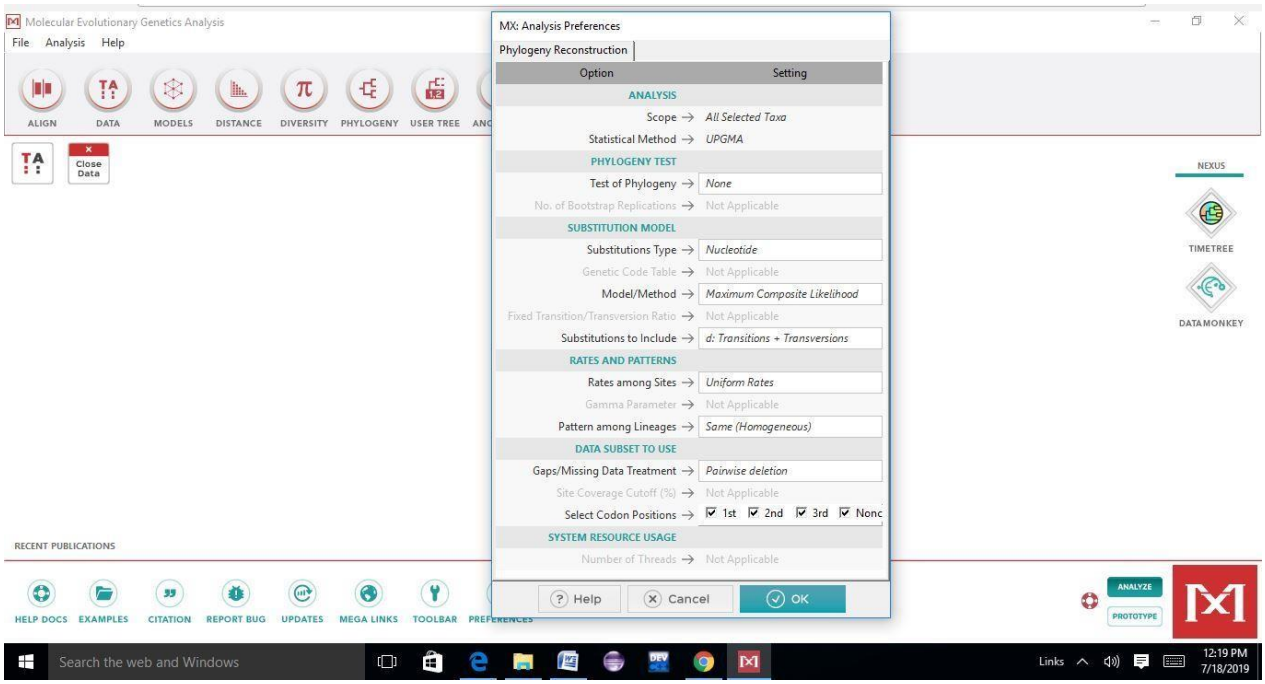


Figure 2.2.3.2 (c): Parameters used for the contruction of phylogenetic trees for protein sequence data

PREDICTION OF ANTIGENS AND SUBUNIT VACCINES

2.3.1.Vaxijen 2.0

The first server for alignment-independent prediction of protective antigens of bacterial, viral and tumour origin. VaxiJen contains models derived by auto- and cross-covariance pre-processing of amino acids properties. The predictive ability of our models was tested by internal leave-one-out cross- validation on training sets and by external validation on test sets. The models showed remarkable stability, as tested by combinations of the positive set and five different negative sets. Thus, VaxiJen is a reliable and consistent tool for the prediction of protective antigens. It can be used singly or in combination with other bioinformatics tools used for reverse vaccinology.

For determining the protective antigen property of various amino acid sequences, we have used the VaxiJenanalysis.(Figure 2.3.1)

URL Link: http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html

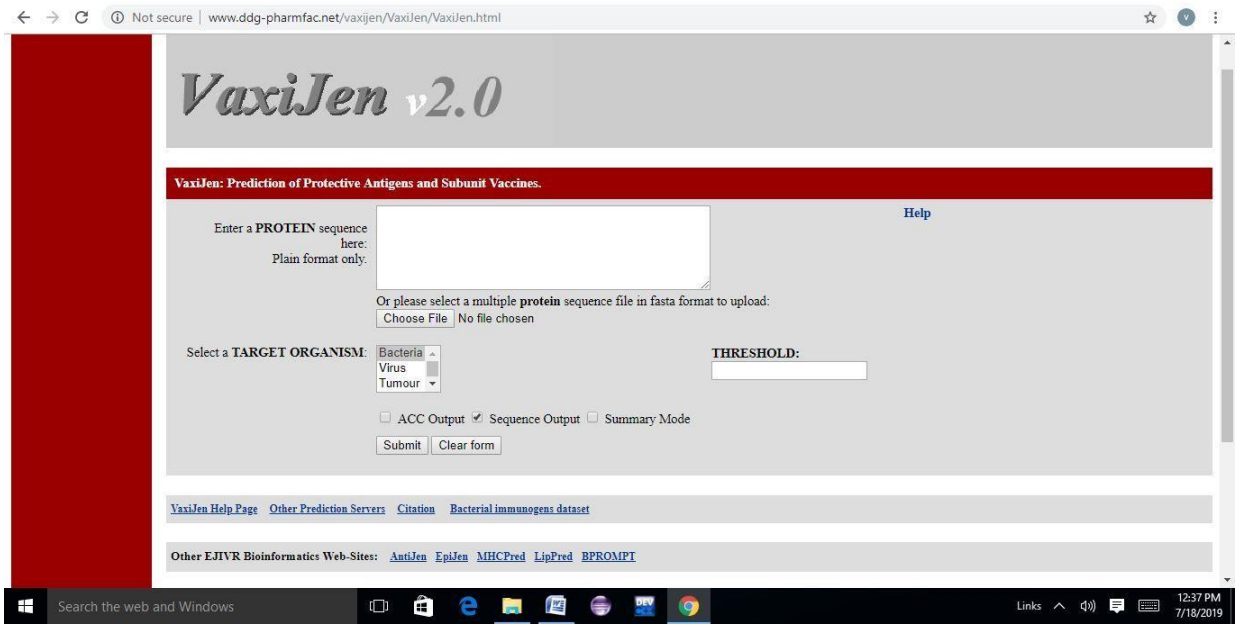


Figure 2.3.1.: VaxiJen 2.0 Homepage 2.4PREDICTION OF T-CELL

EPITOPE

NetCTL 1.2 Server

NetCTL 1.2 server predicts CTL epitopes in protein sequences. The current version 1.2 is an update to the version 1.0. The version 1.2 expands the MHC class I binding predicition to 12 MHC supertypes including the supertypes A26 and B39. The accuracy of the MHC class I peptide binding affinity is significantly improved compared to the earlier version. Also the prediction of proteasonal cleavage has been improved and is now identical to the predictions obtained by the NetChop-3.0 server. The updated version has been trained on a set of 886 known MHC class I ligands.

The method integrates prediction of peptide MHC class I binding, proteasomal C terminal cleavage and TAP transport efficiency. The server allows for predictions of CTL epitopes restricted to 12 MHC class I supertype. MHC class I binding and proteasomal cleavage is performed using artificial neural networks. TAP transport efficiency is predictied using weight matrix.

Using this software, we have predicted the CTL epitopes in both the gene and protein sequences. (Figure:2.4.1)

URL link: http://www.cbs.dtu.dk/services/NetCTL/

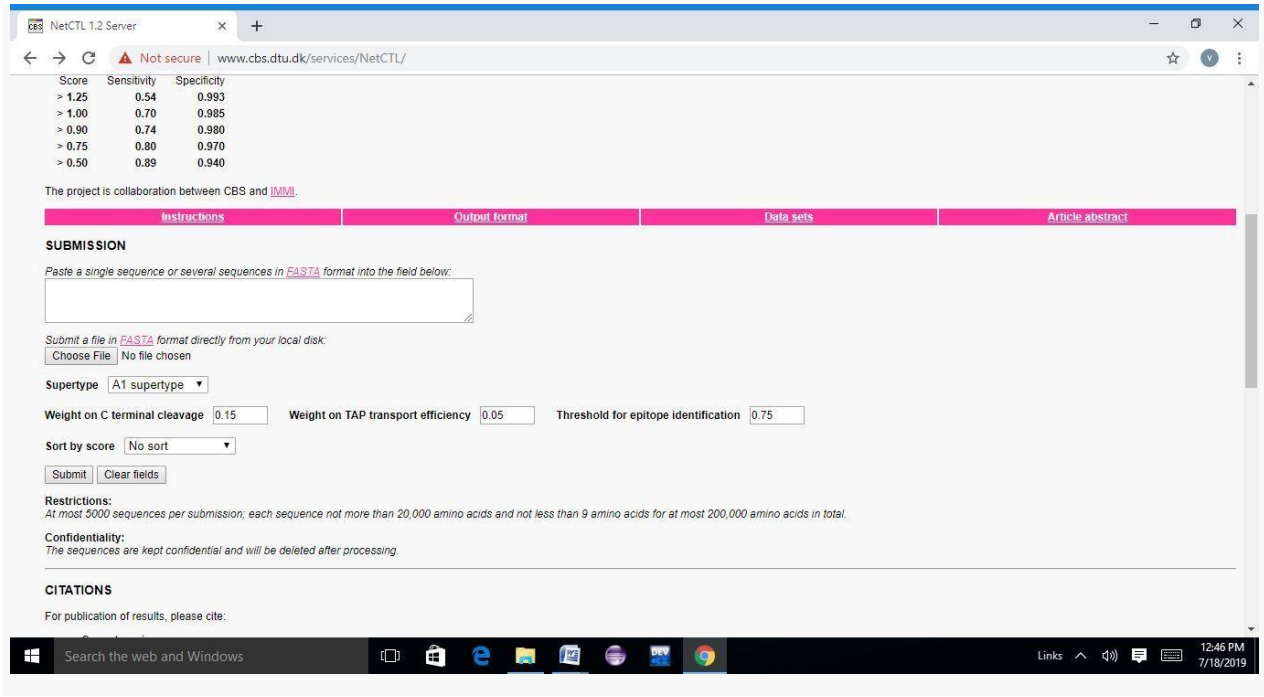


Figure 2.4.1: Net CTL Homepage

IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE(IEDB)

The Immune Epitope Database (IEDB) is a freely available resource funded by NIAID. It catalogs experimental data on antibody and T cell epitopes studied in humans, non-human primates, and other animal species in the context of infectious disease, allergy, autoimmunity and transplantation. The IEDB also hosts tools to assist in the prediction and analysis of epitopes.

Using this software, identification of epitopes and antigens are done. (Figure 2.4.2)

URL link: https://www.iedb.org/

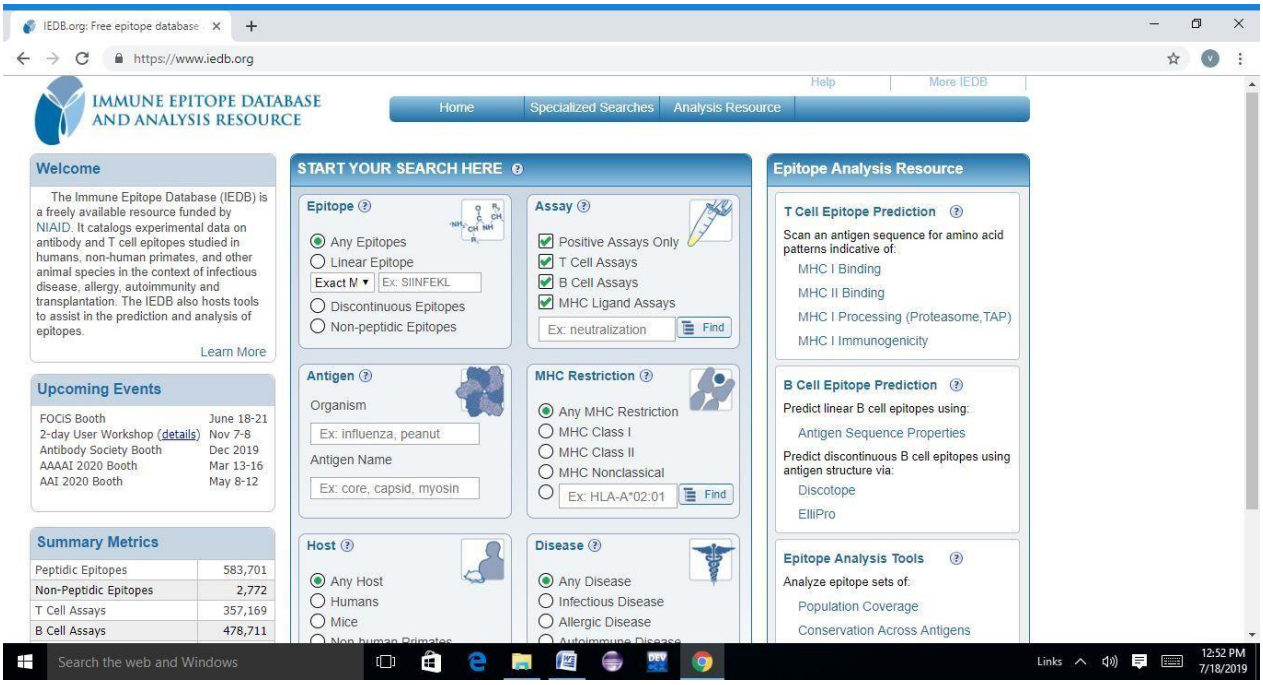


Figure: 2.4.2: IEDB Homepage

PREDICTION OF B-CELL EPITOPE

Antibody epitope prediction:

This is a bioinformatics tool for B-cell epitope prediction.

The following methods are provided for B-cell epitope prediction:

- Chou&Fasman beta-turn prediction
- Emini surface accessibility prediction
- Karplus &schulz flexibility prediction
- Kolaskar&Tongaonkar antigenicity
- Parkar hydrophilicity prediction
- Bepipred linear epitope prediction

Using this software, prediction of B-cell epitopes is done. (Figure 2.5.1)

URL Link: http://tools.immuneepitope.org/bcell/

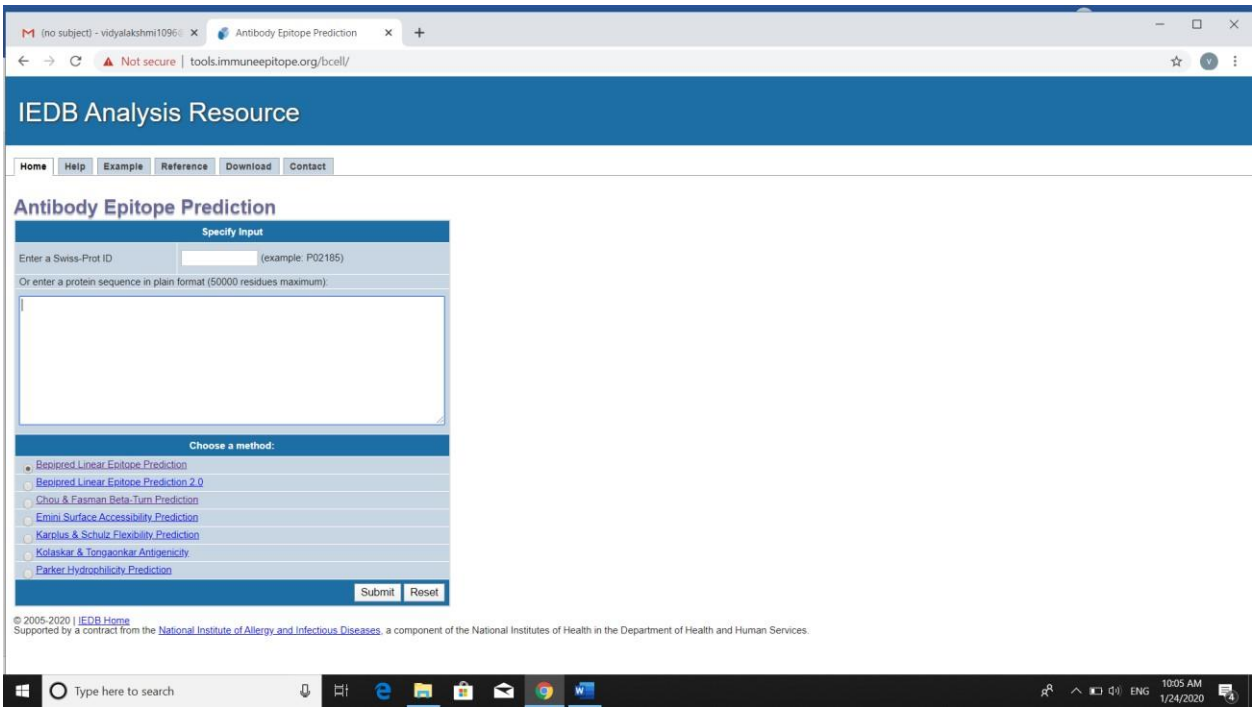


Figure2.5.1: Antibody Epitope Prediction Homepage

BepiPred 1.0:

BepiPredpredicts the location of linear B-cell epitopes using a combination of a hidden Markov model and a propensity scale method.

Using this software, B-cell epitopes are identified. (Figure 2.5.2) URL link:

http://www.cbs.dtu.dk/services/BepiPred-1.0/

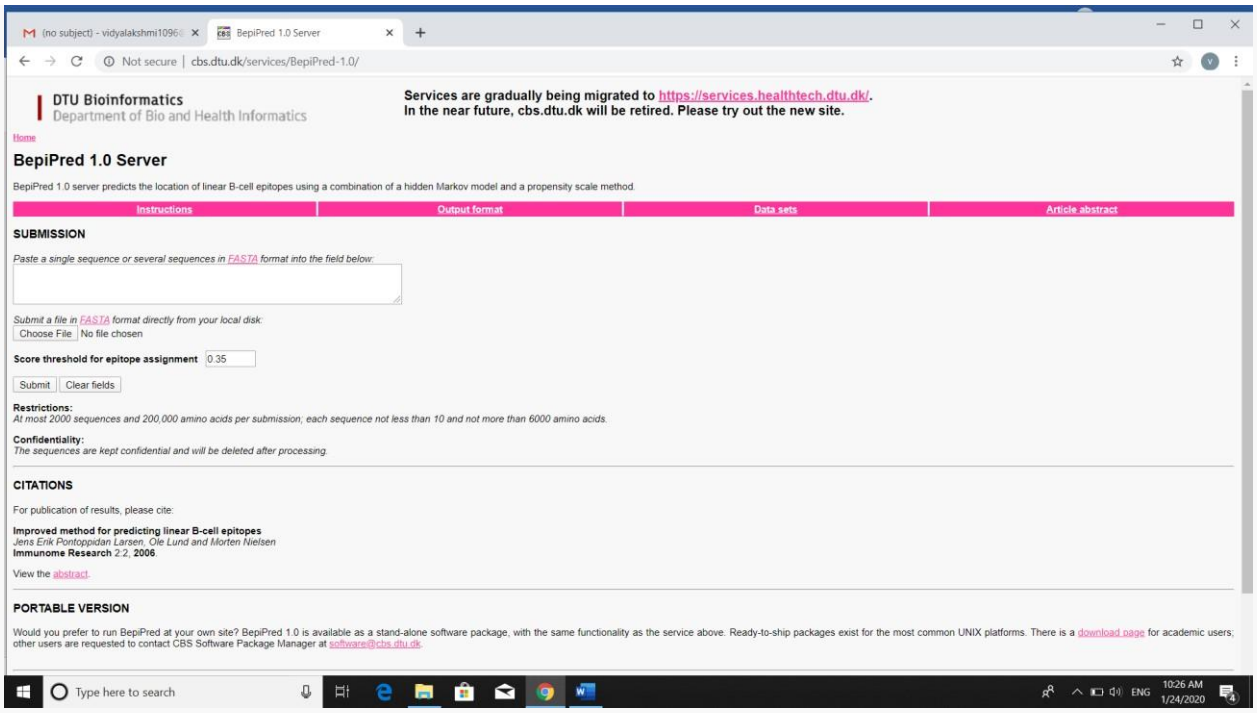


Figure 2.5.2: BepiPred Homepage

ALLERGENCITY PREDICTION

Allertop V 2.0:

This is a Bioinformatics tool for allergenicity prediction based on a novel descriptor fingerprint approach.

With the help of this tool, the allergenicity prediction of various sequences was done. (Figure: 2.6.1)

URL link: http://www.ddg-pharmfac.net/allertop/

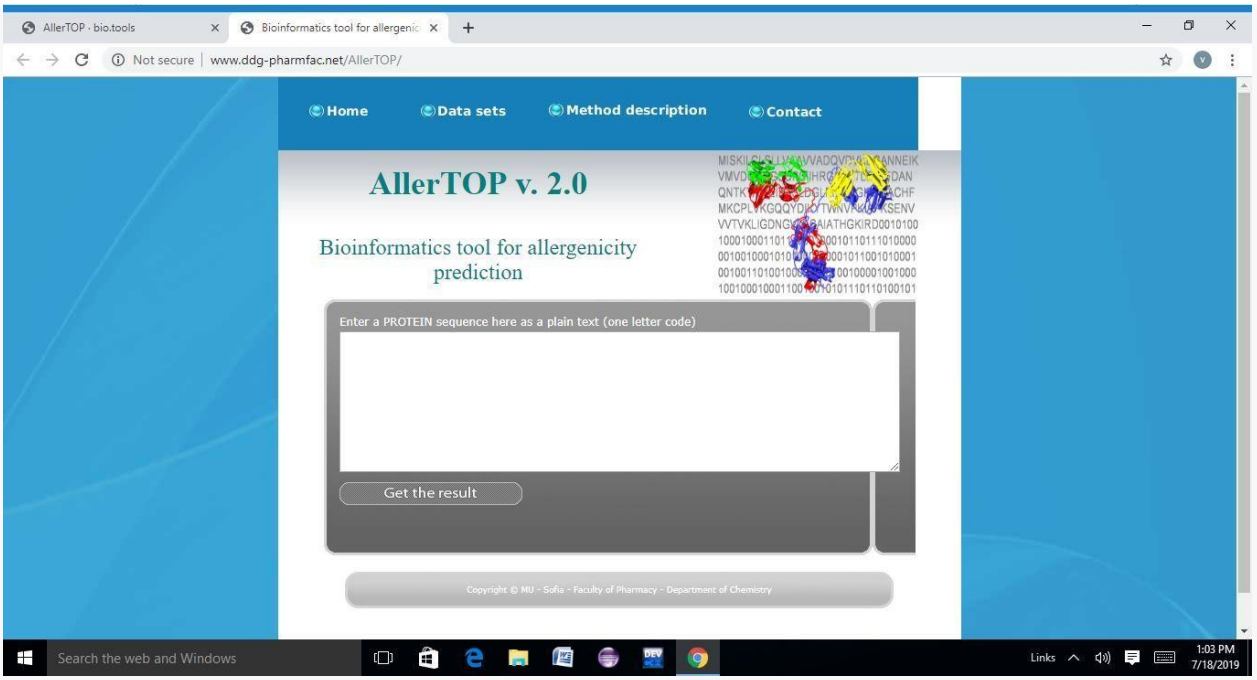


Figure 2.6.1: AllerTOP V 2.0 homepage

AllergenFP v.1.0

This is a Bioinformatics tool for allergenicity prediction based on a novel descriptor fingerprint approach.

With the help of this tool, the allergenicity prediction of various sequences was done. (Figure: 2.6.2)

URL link: http://ddg-pharmfac.net/AllergenFP/

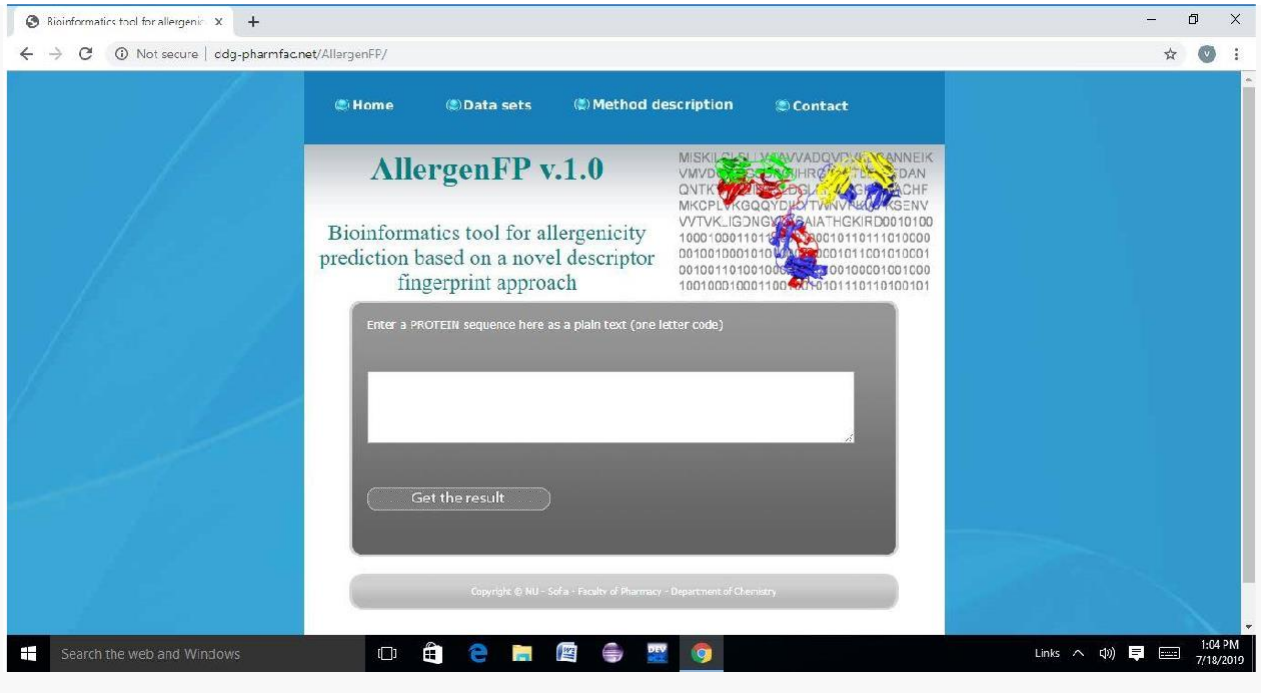


Figure 2.6.2: Allergen FP v 1.0 Homepage

3D MODELLING OF AN EPITOPE

PEP FOLD:

PEP-FOLD is a *de novo* approach aimed at predicting peptide structures from amino acid sequences. This method, based on structural alphabet SA letters to describe the conformations of four consecutive residues, couples the predicted series of SA letters to a greedy algorithm and a coarse-grained force field. PEP-FOLD latest evolution improves performance for linear peptides up to 36 amino acids - best model with an averaged RMSd of 2.1 A from NMR structure, also allows user specified constraints such as disulfide bonds and inter-residue proximities.

Using the PEP-FOLD analysis, the structure of various amino acid sequences were predicted. (Figure 2.7.1)

URL link: http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD/

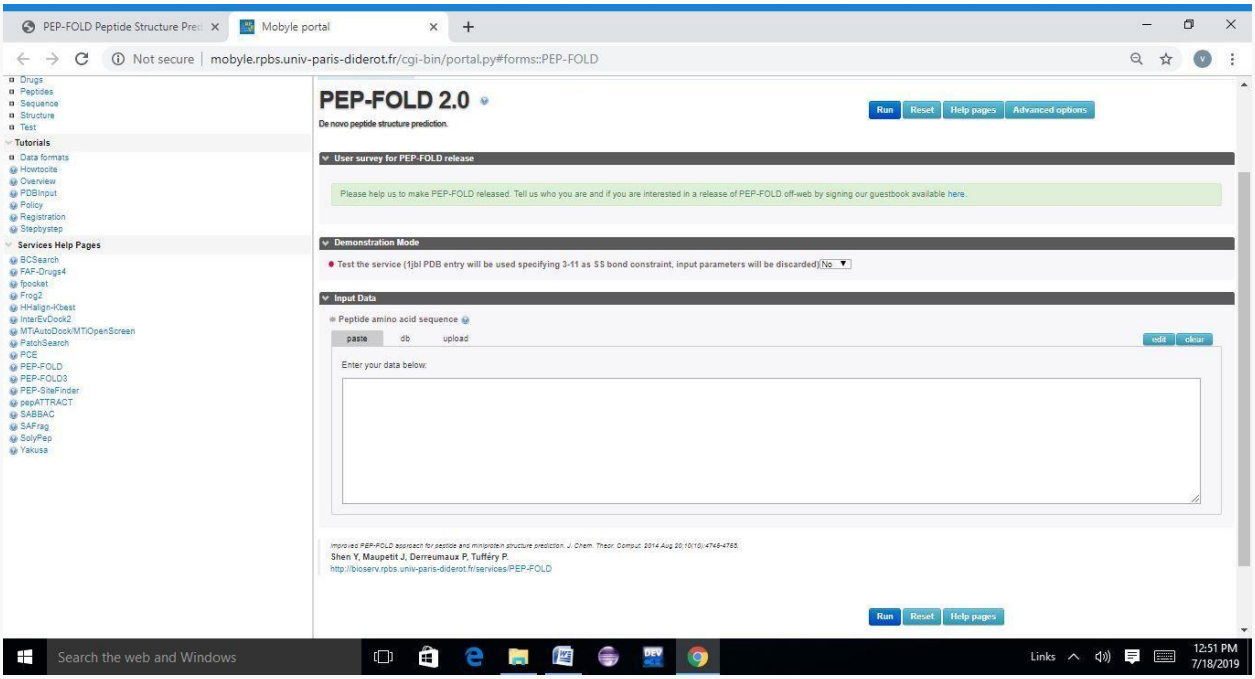


Figure 2.7.1: PEP-FOLD Homepage

PHYRE2:

The Phyre and Phyre2 servers predict the three-dimensional structure of a protein sequence using the principles and techniques of homology modeling. Because the structure of a protein is more conserved in evolution than its amino acid sequence, a protein sequence of interest (the target) can be modeled with reasonable accuracy on a very distantly related sequence of known structure (the template), provided that the relationship between target and template can be discerned through sequence alignment. Currently the most powerful and accurate methods for detecting and aligning remotely related sequences rely on profiles or hidden Markov models (HMMs).

Using this software, the structure of various amino acid sequences were predicted. (Figure 2.7.2)

URL link: http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index

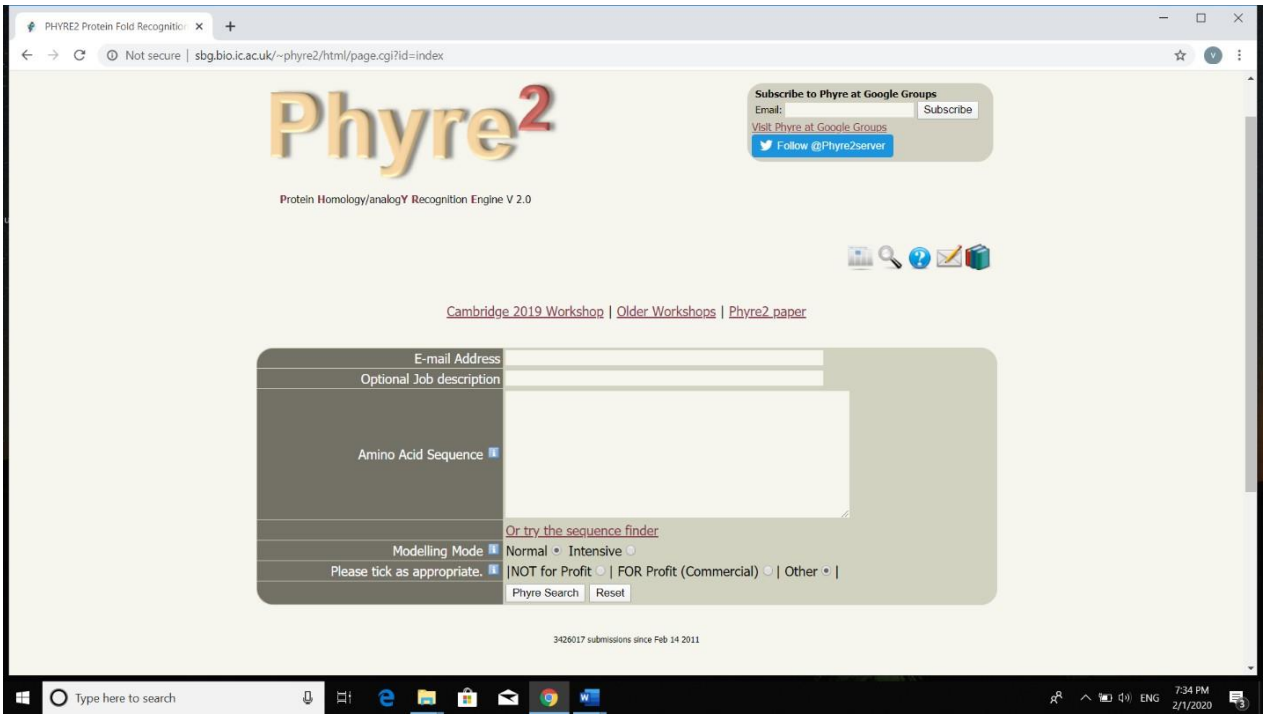


Figure 2.7.2: PHYRE2 Homepage

CHAPTER 3 RESULTS AND DISCUSSION

DATA RETRIVAL

The National Centre For Biotechnology Information (NCBI):

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

Various genomic, protein and glycoprotein sequences of Hepatitis Virus were identified using NCBI.

S.No.	Virus	Accession ID
1.	HEPATITIS A	MK829707
2.	HEPATITIS B	MK075117
3.	HEPATITIS C	MK527509
4.	HEPATITIS D	MH844625
5.	HEPATITIS E	LC436450

Table3.1.1a: List of Hepatitis virus genomes selected for vaccine design

S.No	Virus	Accession ID
1.	HEPATITIS A	MK829707
2.	HEPATITIS B	MK075117
3.	HEPATITIS C	MK527509
4.	HEPATITIS D	MH844625

5.	HEPATITIS E	LC436450

Table3.1.1b: List of Hepatitis virus genome Translation sequences selected for vaccine design

S.No	Virus	Accession ID
1.	HEPATITIS A	AAC39862
2.	HEPATITIS B	AAX44104
3.	HEPATITIS C	AAL25079
5.	HEPATITIS E	ATU81864

Table3.1.1c: List of Hepatitis virus glycoprotein sequences selected for vaccine design

EVOLUTIONARY ANALYSIS

CLUSTAL OMEGA:

This is used for multiple sequence alignment.

Various gene sequences of hepatitis virus were uploaded in the clustal omega and the results of phylogenetic tree is mentioned below.

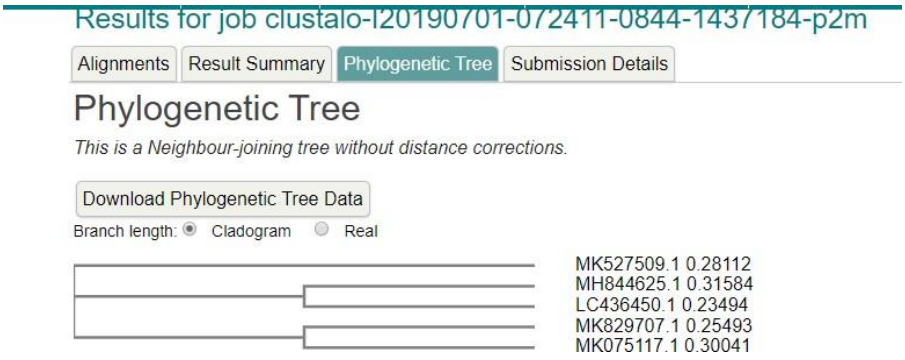


Fig 3.2.1a: Phylogenetic tree of genomic sequences

Various protein sequences of hepatitis virus were uploaded in the clustal omega and the results of phylogenetic tree is mentioned below.

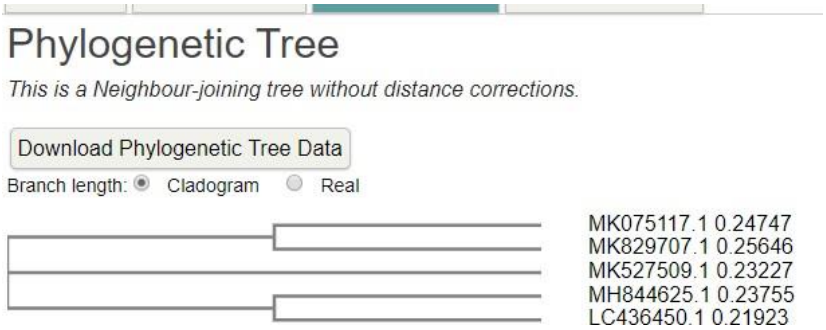


Fig 3.2.1b: Phylogenetic tree of protein sequences

Molecular Evolutionary Genetics Analysis (MEGA):

Using this software, the genome sequence is uploaded and a phylogenetic tree is constructed. The results are displayed below:

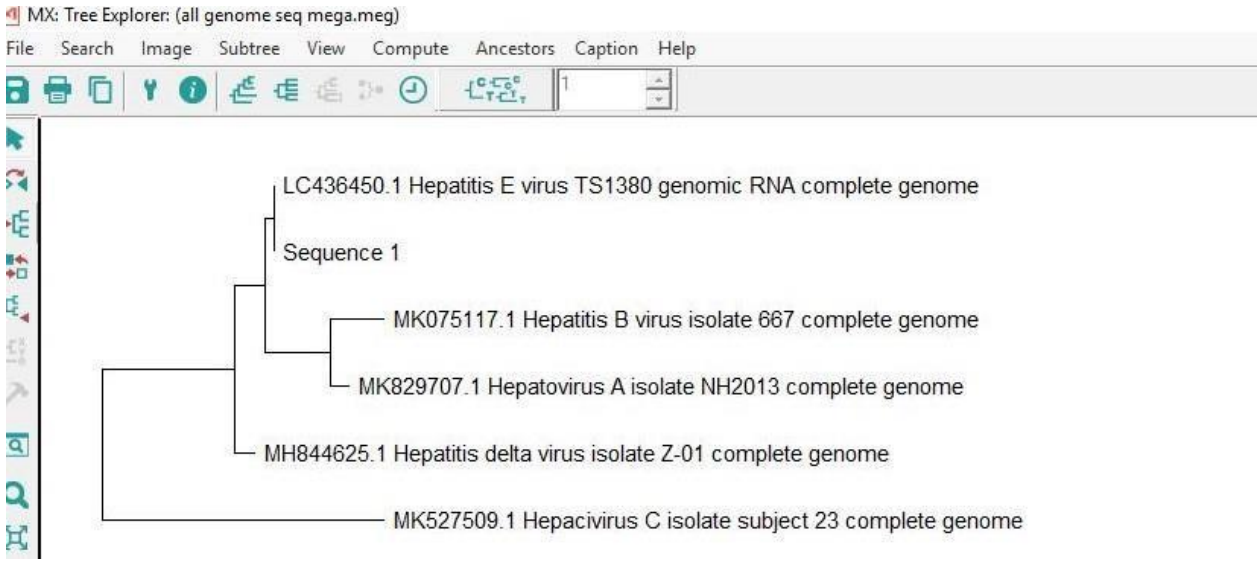


Fig 3.2.2a: Phylogenetic tree of Genomic sequence

Using this software, the glycoprotein sequence is uploaded and a phylogenetic tree is constructed. The results are displayed below. The glycoprotein sequence is retrieved from UNIPROT.

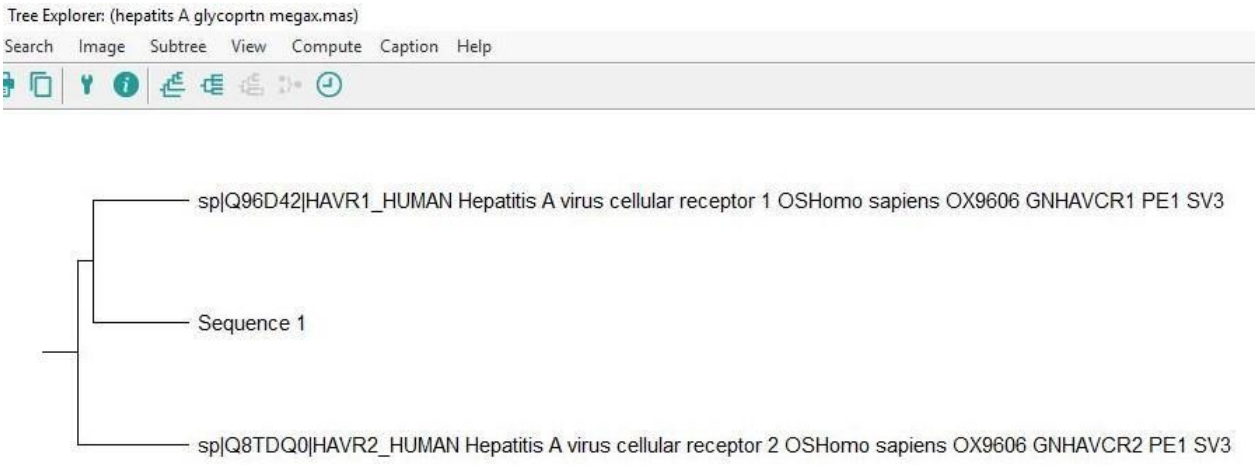


Fig 3.2.2b: Phylogenetic tree of Glycoprotein sequence.

PREDICTION OF ANTIGENS AND SUBUNIT VACCINES

3.3.1. Vaxijen 2.0

For determining the protective antigen property of various amino acid sequences, we have used the VaxiJen analysis and the results are displayed below.

S.No	Virus	Overall prediction for protective Antigen	Nature
1.	HEPATITIS A	0.5464	Probable Antigen
2.	HEPATITIS B	0.5509	Probable Antigen
3.	HEPATITIS C	0.4597	Probable non-antigen
4.	HEPATITIS D	0.4759	Probable non-antigen

5.	HEPATITIS E	0.3817	Probable non-antigen

Table3.3.1a: Determination of Antigenic property of translation sequences using Vaxijen 2.0

For determining the protective antigen property of various Glycoprotein sequences, we have used the VaxiJen analysis and the results are displayed below.

S.No	**Virus**	**Overall prediction for protective Antigen**	**Nature**
1.	HEPATITIS A	0.5605	Probable Antigen
2.	HEPATITIS B	0.4493	Probable non-antigen
3.	HEPATITIS C	0.5728	Probable Antigen
4.	HEPATITIS E	0.5975	Probable Antigen

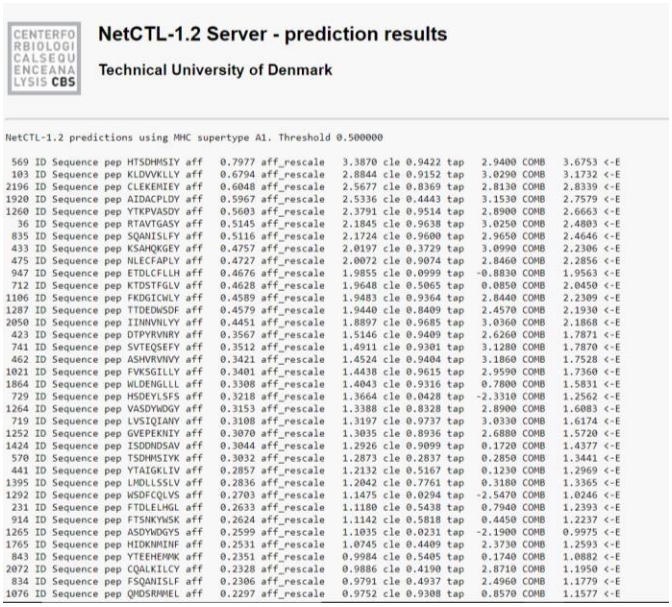
Table3.3.1b: Determination of Antigenic property of Glycoprotein sequences using Vaxijen 2.0

PREDICTION OF T-CELL EPITOPE

3.4.1. NetCTL 1.2 Server:

NetCTL 1.2 server predicts CTL epitopes in protein sequences.

Using this software, we have predicted the CTL epitopes in the gene sequences. The gene sequences of various Hepatitis Virus are uploaded in the NetCTL Server and the results are diaplayed below. (Figure:3.4.1)



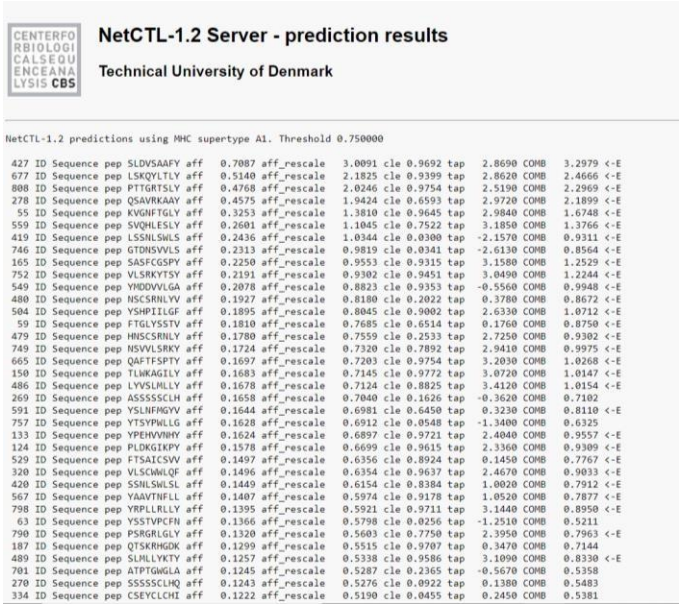
NetCTL-1.2 Server - prediction results

Technical University of Denmark

NetCTL-1.2 predictions using MHC supertype A1. Threshold 0.500000

569	ID	Sequence	pep	HTSDHMSIY	aff	0.7977	aff_rescale	3.3870	cle	0.9422	tap	2.9400	COMB	3.6753 <-E
103	ID	Sequence	pep	KLDVVKLLY	aff	0.6794	aff_rescale	2.8844	cle	0.9152	tap	3.0290	COMB	3.1732 <-E
2196	ID	Sequence	pep	CLEKEMIEY	aff	0.6048	aff_rescale	2.5677	cle	0.8369	tap	2.8130	COMB	2.8339 <-E
1920	ID	Sequence	pep	AIDACPLDY	aff	0.5967	aff_rescale	2.5336	cle	0.4443	tap	3.1530	COMB	2.7579 <-E
1260	ID	Sequence	pep	YTKPVASDY	aff	0.5603	aff_rescale	2.3791	cle	0.9514	tap	2.8900	COMB	2.6663 <-E
36	ID	Sequence	pep	RTAVTGASY	aff	0.5145	aff_rescale	2.1845	cle	0.9638	tap	3.0250	COMB	2.4803 <-E
835	ID	Sequence	pep	SQANISLFY	aff	0.5116	aff_rescale	2.1724	cle	0.9600	tap	2.9650	COMB	2.4646 <-E
433	ID	Sequence	pep	KSAHQKGEY	aff	0.4757	aff_rescale	2.0197	cle	0.3729	tap	3.0990	COMB	2.2306 <-E
475	ID	Sequence	pep	NLECFAPLY	aff	0.4727	aff_rescale	2.0072	cle	0.9074	tap	2.8460	COMB	2.2856 <-E
947	ID	Sequence	pep	ETDLCFLLH	aff	0.4676	aff_rescale	1.9855	cle	0.0999	tap	-0.8830	COMB	1.9563 <-E
712	ID	Sequence	pep	KTDSTFGLV	aff	0.4628	aff_rescale	1.9648	cle	0.5065	tap	0.0850	COMB	2.0450 <-E
1106	ID	Sequence	pep	FKDGICWLY	aff	0.4589	aff_rescale	1.9483	cle	0.9364	tap	2.8440	COMB	2.2309 <-E
1287	ID	Sequence	pep	TTDEDWSDF	aff	0.4579	aff_rescale	1.9440	cle	0.8409	tap	2.4570	COMB	2.1930 <-E
2050	ID	Sequence	pep	IINNVNLYY	aff	0.4451	aff_rescale	1.8897	cle	0.9685	tap	3.0360	COMB	2.1868 <-E
423	ID	Sequence	pep	DTPYRVNRY	aff	0.3567	aff_rescale	1.5146	cle	0.9409	tap	2.6260	COMB	1.7871 <-E
741	ID	Sequence	pep	SVTEQSEFY	aff	0.3512	aff_rescale	1.4911	cle	0.9301	tap	3.1280	COMB	1.7870 <-E
462	ID	Sequence	pep	ASHVRVNVY	aff	0.3421	aff_rescale	1.4524	cle	0.9404	tap	3.1860	COMB	1.7528 <-E
1021	ID	Sequence	pep	FVKSGILLY	aff	0.3401	aff_rescale	1.4438	cle	0.9615	tap	2.9590	COMB	1.7360 <-E
1864	ID	Sequence	pep	WLDENGLLL	aff	0.3308	aff_rescale	1.4043	cle	0.9316	tap	0.7800	COMB	1.5831 <-E
729	ID	Sequence	pep	HSDEYLSFS	aff	0.3218	aff_rescale	1.3664	cle	0.0428	tap	-2.3310	COMB	1.2562 <-E
1264	ID	Sequence	pep	VASDYWDGY	aff	0.3153	aff_rescale	1.3388	cle	0.8328	tap	2.8900	COMB	1.6083 <-E
719	ID	Sequence	pep	LVSIQIANY	aff	0.3108	aff_rescale	1.3197	cle	0.9737	tap	3.0330	COMB	1.6174 <-E
1252	ID	Sequence	pep	GVEPEKNIY	aff	0.3070	aff_rescale	1.3035	cle	0.8936	tap	2.6880	COMB	1.5720 <-E
1424	ID	Sequence	pep	ISDDNDSAV	aff	0.3044	aff_rescale	1.2926	cle	0.9099	tap	0.1720	COMB	1.4377 <-E
570	ID	Sequence	pep	TSDHMSIYK	aff	0.3032	aff_rescale	1.2873	cle	0.2837	tap	0.2850	COMB	1.3441 <-E
441	ID	Sequence	pep	YTAIGKLIV	aff	0.2857	aff_rescale	1.2132	cle	0.5167	tap	0.1230	COMB	1.2969 <-E
1395	ID	Sequence	pep	LMDLLSSLV	aff	0.2836	aff_rescale	1.2042	cle	0.7761	tap	0.3180	COMB	1.3365 <-E
1292	ID	Sequence	pep	WSDFCQLVS	aff	0.2703	aff_rescale	1.1475	cle	0.0294	tap	-2.5470	COMB	1.0246 <-E
231	ID	Sequence	pep	FTDLELHGL	aff	0.2633	aff_rescale	1.1180	cle	0.5438	tap	0.7940	COMB	1.2393 <-E
914	ID	Sequence	pep	FTSNKYWSK	aff	0.2624	aff_rescale	1.1142	cle	0.5818	tap	0.4450	COMB	1.2237 <-E
1265	ID	Sequence	pep	ASDYWDGYS	aff	0.2599	aff_rescale	1.1035	cle	0.0231	tap	-2.1900	COMB	0.9975 <-E
1765	ID	Sequence	pep	HIDKNMINF	aff	0.2531	aff_rescale	1.0745	cle	0.4409	tap	2.3730	COMB	1.2593 <-E
843	ID	Sequence	pep	YTEEHEMMK	aff	0.2351	aff_rescale	0.9984	cle	0.5405	tap	0.1740	COMB	1.0882 <-E
2072	ID	Sequence	pep	CQALKILCY	aff	0.2328	aff_rescale	0.9886	cle	0.4190	tap	2.8710	COMB	1.1950 <-E
834	ID	Sequence	pep	FSQANISLF	aff	0.2306	aff_rescale	0.9791	cle	0.4937	tap	2.4960	COMB	1.1779 <-E
1076	ID	Sequence	pep	QMDSRMMEL	aff	0.2297	aff_rescale	0.9752	cle	0.9308	tap	0.8570	COMB	1.1577 <-E

Fig 3.4.1a: NetCTL-1.2 Server results of Hepatitis A



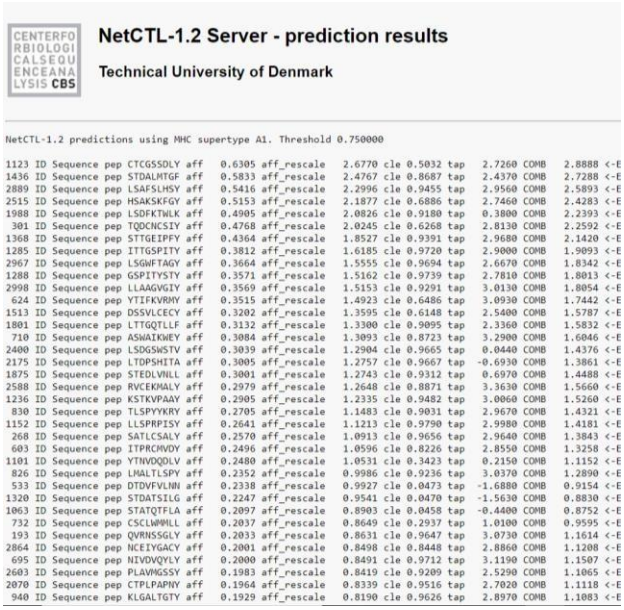
NetCTL-1.2 Server - prediction results

Technical University of Denmark

NetCTL-1.2 predictions using MHC supertype A1. Threshold 0.750000

427	ID	Sequence	pep	SLDVSAAFY	aff	0.7087	aff_rescale	3.0091	cle	0.9692	tap	2.8690	COMB	3.2979 <-E
677	ID	Sequence	pep	LSKQYLTLY	aff	0.5140	aff_rescale	2.1825	cle	0.9399	tap	2.8620	COMB	2.4666 <-E
808	ID	Sequence	pep	PTTGRTSLY	aff	0.4768	aff_rescale	2.0246	cle	0.9754	tap	2.5190	COMB	2.2969 <-E
278	ID	Sequence	pep	QSAVRKAAY	aff	0.4575	aff_rescale	1.9424	cle	0.6593	tap	2.9720	COMB	2.1899 <-E
55	ID	Sequence	pep	KVGNFTGLY	aff	0.3253	aff_rescale	1.3810	cle	0.9645	tap	2.9840	COMB	1.6748 <-E
559	ID	Sequence	pep	SVQHLESLY	aff	0.2601	aff_rescale	1.1045	cle	0.7522	tap	3.1850	COMB	1.3766 <-E
419	ID	Sequence	pep	LSSNLSWLS	aff	0.2436	aff_rescale	1.0344	cle	0.0300	tap	-2.1570	COMB	0.9311 <-E
746	ID	Sequence	pep	GTDNSVVLS	aff	0.2313	aff_rescale	0.9819	cle	0.0341	tap	-2.6130	COMB	0.8564 <-E
165	ID	Sequence	pep	SASFCGSPY	aff	0.2250	aff_rescale	0.9553	cle	0.9315	tap	3.1580	COMB	1.2529 <-E
752	ID	Sequence	pep	VLSRKYTSY	aff	0.2191	aff_rescale	0.9302	cle	0.9451	tap	3.0490	COMB	1.2244 <-E
549	ID	Sequence	pep	YMDDVVLGA	aff	0.2078	aff_rescale	0.8823	cle	0.9353	tap	-0.5560	COMB	0.9948 <-E
480	ID	Sequence	pep	NSCSRNLYV	aff	0.1927	aff_rescale	0.8180	cle	0.2022	tap	0.3780	COMB	0.8672 <-E
504	ID	Sequence	pep	YSHPIILGF	aff	0.1895	aff_rescale	0.8045	cle	0.9002	tap	2.6330	COMB	1.0712 <-E
59	ID	Sequence	pep	FTGLYSSTV	aff	0.1810	aff_rescale	0.7685	cle	0.6514	tap	0.1760	COMB	0.8750 <-E
479	ID	Sequence	pep	HNSCSRNLY	aff	0.1780	aff_rescale	0.7559	cle	0.2533	tap	2.7250	COMB	0.9302 <-E
749	ID	Sequence	pep	NSVVLSRKY	aff	0.1724	aff_rescale	0.7320	cle	0.7892	tap	2.9410	COMB	0.9975 <-E
665	ID	Sequence	pep	QAFTFSPTY	aff	0.1697	aff_rescale	0.7203	cle	0.9754	tap	3.2030	COMB	1.0268 <-E
150	ID	Sequence	pep	TLWKAGILY	aff	0.1683	aff_rescale	0.7145	cle	0.9772	tap	3.0720	COMB	1.0147 <-E
486	ID	Sequence	pep	LYVSLMLLY	aff	0.1678	aff_rescale	0.7124	cle	0.8825	tap	3.4120	COMB	1.0154 <-E
269	ID	Sequence	pep	ASSSSSCLH	aff	0.1658	aff_rescale	0.7040	cle	0.1626	tap	-0.3620	COMB	0.7102
591	ID	Sequence	pep	YSLNFNGYV	aff	0.1644	aff_rescale	0.6981	cle	0.6450	tap	0.3230	COMB	0.8110 <-E
757	ID	Sequence	pep	YTSYPWLLG	aff	0.1628	aff_rescale	0.6912	cle	0.0548	tap	-1.3400	COMB	0.6325
133	ID	Sequence	pep	YPEHVVNHY	aff	0.1624	aff_rescale	0.6897	cle	0.9721	tap	2.4040	COMB	0.9557 <-E
124	ID	Sequence	pep	PLDKGIKPY	aff	0.1578	aff_rescale	0.6699	cle	0.9615	tap	2.3360	COMB	0.9309 <-E
529	ID	Sequence	pep	FTSAICSVV	aff	0.1497	aff_rescale	0.6356	cle	0.8924	tap	0.1450	COMB	0.7767 <-E
320	ID	Sequence	pep	VLSCWWLQF	aff	0.1496	aff_rescale	0.6354	cle	0.9637	tap	2.4670	COMB	0.9033 <-E
420	ID	Sequence	pep	SSNLSWLSL	aff	0.1449	aff_rescale	0.6154	cle	0.8384	tap	1.0020	COMB	0.7912 <-E
567	ID	Sequence	pep	YAAVTNFLL	aff	0.1407	aff_rescale	0.5974	cle	0.9178	tap	1.0520	COMB	0.7877 <-E
798	ID	Sequence	pep	YRPLLRLLY	aff	0.1395	aff_rescale	0.5921	cle	0.9711	tap	3.1440	COMB	0.8950 <-E
63	ID	Sequence	pep	YSSTVPCFN	aff	0.1366	aff_rescale	0.5798	cle	0.0256	tap	-1.2510	COMB	0.5211
790	ID	Sequence	pep	PSRGRLGLY	aff	0.1320	aff_rescale	0.5603	cle	0.7750	tap	2.3950	COMB	0.7963 <-E
187	ID	Sequence	pep	QTSKRHGDK	aff	0.1299	aff_rescale	0.5515	cle	0.9707	tap	0.3470	COMB	0.7144
489	ID	Sequence	pep	SLMLLYKTY	aff	0.1257	aff_rescale	0.5338	cle	0.9586	tap	3.1090	COMB	0.8330 <-E
701	ID	Sequence	pep	ATPTGWGLA	aff	0.1245	aff_rescale	0.5287	cle	0.2365	tap	-0.5670	COMB	0.5358
270	ID	Sequence	pep	SSSSSCLHQ	aff	0.1243	aff_rescale	0.5276	cle	0.0922	tap	0.1380	COMB	0.5483
334	ID	Sequence	pep	CSEYCLCHI	aff	0.1222	aff_rescale	0.5190	cle	0.0455	tap	0.2450	COMB	0.5381

Fig 3.4.1b:NetCTL-1.2 Server results of Hepatitis B



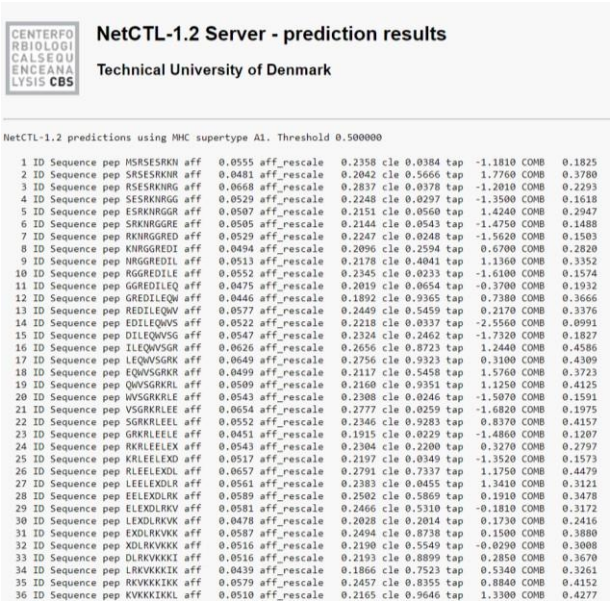
NetCTL-1.2 Server - prediction results

Technical University of Denmark

NetCTL-1.2 predictions using MHC supertype A1. Threshold 0.750000

1123	ID	Sequence	pep	CTCGSSDLY	aff	0.6305	aff_rescale	2.6770	cle	0.5032	tap	2.7260	COMB	2.8888	<-E
1436	ID	Sequence	pep	STDALMTGF	aff	0.5833	aff_rescale	2.4767	cle	0.8687	tap	2.4370	COMB	2.7288	<-E
2889	ID	Sequence	pep	LSAFSLHSY	aff	0.5416	aff_rescale	2.2996	cle	0.9455	tap	2.9560	COMB	2.5893	<-E
2515	ID	Sequence	pep	HSAKSKFGY	aff	0.5153	aff_rescale	2.1877	cle	0.6886	tap	2.7460	COMB	2.4283	<-E
1988	ID	Sequence	pep	LSDFKTWLK	aff	0.4905	aff_rescale	2.0826	cle	0.9180	tap	0.3800	COMB	2.2393	<-E
301	ID	Sequence	pep	TQDCNCSIY	aff	0.4768	aff_rescale	2.0245	cle	0.6268	tap	2.8130	COMB	2.2592	<-E
1368	ID	Sequence	pep	STTGEIPFY	aff	0.4364	aff_rescale	1.8527	cle	0.9391	tap	2.9680	COMB	2.1420	<-E
1285	ID	Sequence	pep	ITTGSPITY	aff	0.3812	aff_rescale	1.6185	cle	0.9720	tap	2.9000	COMB	1.9093	<-E
2967	ID	Sequence	pep	LSGWFTAGY	aff	0.3664	aff_rescale	1.5555	cle	0.9694	tap	2.6670	COMB	1.8342	<-E
1288	ID	Sequence	pep	GSPITYSTY	aff	0.3571	aff_rescale	1.5162	cle	0.9739	tap	2.7810	COMB	1.8013	<-E
2998	ID	Sequence	pep	LLAAGVGIY	aff	0.3569	aff_rescale	1.5153	cle	0.9291	tap	3.0130	COMB	1.8054	<-E
624	ID	Sequence	pep	YTIFKVRMY	aff	0.3515	aff_rescale	1.4923	cle	0.6486	tap	3.0930	COMB	1.7442	<-E
1513	ID	Sequence	pep	DSSVLCECY	aff	0.3202	aff_rescale	1.3595	cle	0.6148	tap	2.5400	COMB	1.5787	<-E
1801	ID	Sequence	pep	LTTGQTLLF	aff	0.3132	aff_rescale	1.3300	cle	0.9095	tap	2.3360	COMB	1.5832	<-E
710	ID	Sequence	pep	ASWAIKWEY	aff	0.3084	aff_rescale	1.3093	cle	0.8723	tap	3.2900	COMB	1.6046	<-E
2400	ID	Sequence	pep	LSDGSWSTV	aff	0.3039	aff_rescale	1.2904	cle	0.9665	tap	0.0440	COMB	1.4376	<-E
2175	ID	Sequence	pep	LTDPSHITA	aff	0.3005	aff_rescale	1.2757	cle	0.9667	tap	-0.6930	COMB	1.3861	<-E
1875	ID	Sequence	pep	STEDLVNLL	aff	0.3001	aff_rescale	1.2743	cle	0.9312	tap	0.6970	COMB	1.4488	<-E
2588	ID	Sequence	pep	RVCEKMALY	aff	0.2979	aff_rescale	1.2648	cle	0.8871	tap	3.3630	COMB	1.5660	<-E
1236	ID	Sequence	pep	KSTKVPAAY	aff	0.2905	aff_rescale	1.2335	cle	0.9482	tap	3.0060	COMB	1.5260	<-E
830	ID	Sequence	pep	TLSPYYKRY	aff	0.2705	aff_rescale	1.1483	cle	0.9031	tap	2.9670	COMB	1.4321	<-E
1152	ID	Sequence	pep	LLSPRPISY	aff	0.2641	aff_rescale	1.1213	cle	0.9790	tap	2.9980	COMB	1.4181	<-E
268	ID	Sequence	pep	SATLCSALY	aff	0.2570	aff_rescale	1.0913	cle	0.9656	tap	2.9640	COMB	1.3843	<-E
603	ID	Sequence	pep	ITPRCMVDY	aff	0.2496	aff_rescale	1.0596	cle	0.8226	tap	2.8550	COMB	1.3258	<-E
1101	ID	Sequence	pep	YTNVDQDLV	aff	0.2480	aff_rescale	1.0531	cle	0.3423	tap	0.2150	COMB	1.1152	<-E
826	ID	Sequence	pep	LMALTLSPY	aff	0.2352	aff_rescale	0.9986	cle	0.9236	tap	3.0370	COMB	1.2890	<-E
533	ID	Sequence	pep	DTDVFVLNN	aff	0.2338	aff_rescale	0.9927	cle	0.0473	tap	-1.6880	COMB	0.9154	<-E
1320	ID	Sequence	pep	STDATSILG	aff	0.2247	aff_rescale	0.9541	cle	0.0470	tap	-1.5630	COMB	0.8830	<-E
1063	ID	Sequence	pep	STATQTFLA	aff	0.2097	aff_rescale	0.8903	cle	0.0458	tap	-0.4400	COMB	0.8752	<-E
732	ID	Sequence	pep	CSCLWMHLL	aff	0.2037	aff_rescale	0.8649	cle	0.2937	tap	1.0100	COMB	0.9595	<-E
193	ID	Sequence	pep	QVRNSSGLY	aff	0.2033	aff_rescale	0.8631	cle	0.9647	tap	3.0730	COMB	1.1614	<-E
2864	ID	Sequence	pep	NCEIYGACY	aff	0.2001	aff_rescale	0.8498	cle	0.8448	tap	2.8860	COMB	1.1208	<-E
695	ID	Sequence	pep	NIVDVQYLY	aff	0.2000	aff_rescale	0.8491	cle	0.9712	tap	3.1190	COMB	1.1507	<-E
2603	ID	Sequence	pep	PLAVMGSSY	aff	0.1983	aff_rescale	0.8419	cle	0.9209	tap	2.5290	COMB	1.1065	<-E
2070	ID	Sequence	pep	CTPLPAPNY	aff	0.1964	aff_rescale	0.8339	cle	0.9516	tap	2.7020	COMB	1.1118	<-E
940	ID	Sequence	pep	KLGALTGTY	aff	0.1929	aff_rescale	0.8190	cle	0.9626	tap	2.8970	COMB	1.1083	<-E

Fig 3.4.1c: NetCTL-1.2 Server results of Hepatitis C



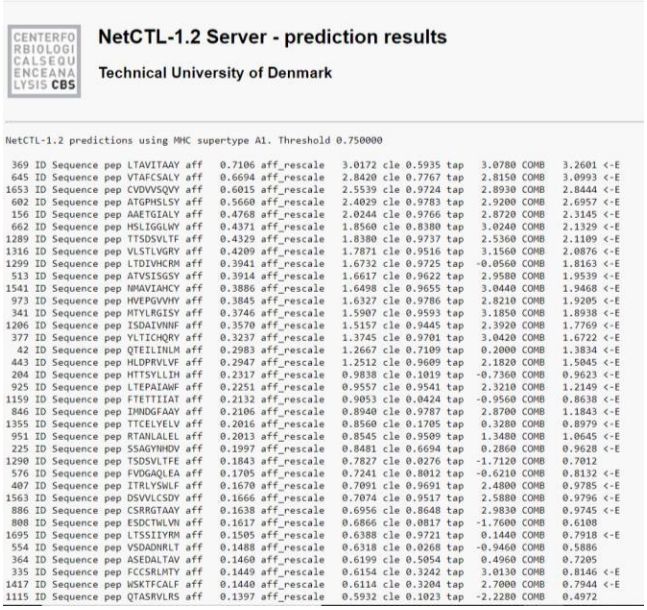
NetCTL-1.2 Server - prediction results

Technical University of Denmark

NetCTL-1.2 predictions using MHC supertype A1. Threshold 0.500000

1	ID	Sequence	pep	MSRSESRKN	aff	0.0555	aff_rescale	0.2358	cle	0.0384	tap	-1.1810	COMB	0.1825
2	ID	Sequence	pep	SRSESRKNR	aff	0.0481	aff_rescale	0.2042	cle	0.5666	tap	1.7760	COMB	0.3780
3	ID	Sequence	pep	RSESRKNRG	aff	0.0668	aff_rescale	0.2837	cle	0.0378	tap	-1.2010	COMB	0.2293
4	ID	Sequence	pep	SESRKNRGG	aff	0.0529	aff_rescale	0.2248	cle	0.0297	tap	-1.3500	COMB	0.1618
5	ID	Sequence	pep	ESRKNRGGR	aff	0.0507	aff_rescale	0.2151	cle	0.0560	tap	1.4240	COMB	0.2947
6	ID	Sequence	pep	SRKNRGGRE	aff	0.0505	aff_rescale	0.2144	cle	0.0543	tap	-1.4750	COMB	0.1488
7	ID	Sequence	pep	RKNRGGRED	aff	0.0529	aff_rescale	0.2247	cle	0.0248	tap	-1.5620	COMB	0.1503
8	ID	Sequence	pep	KNRGGREDI	aff	0.0494	aff_rescale	0.2096	cle	0.2594	tap	0.6700	COMB	0.2820
9	ID	Sequence	pep	NRGGREDIL	aff	0.0513	aff_rescale	0.2178	cle	0.4041	tap	1.1360	COMB	0.3352
10	ID	Sequence	pep	RGGREDILE	aff	0.0552	aff_rescale	0.2345	cle	0.0233	tap	-1.6100	COMB	0.1574
11	ID	Sequence	pep	GGREDILEQ	aff	0.0475	aff_rescale	0.2019	cle	0.0654	tap	-0.3700	COMB	0.1932
12	ID	Sequence	pep	GREDILEQW	aff	0.0446	aff_rescale	0.1892	cle	0.9365	tap	0.7380	COMB	0.3666
13	ID	Sequence	pep	REDILEQWV	aff	0.0577	aff_rescale	0.2449	cle	0.5459	tap	0.2170	COMB	0.3376
14	ID	Sequence	pep	EDILEQWVS	aff	0.0522	aff_rescale	0.2218	cle	0.0337	tap	-2.5560	COMB	0.0991
15	ID	Sequence	pep	DILEQWVSG	aff	0.0547	aff_rescale	0.2324	cle	0.2462	tap	-1.7320	COMB	0.1827
16	ID	Sequence	pep	ILEQWVSGR	aff	0.0626	aff_rescale	0.2656	cle	0.8723	tap	1.2440	COMB	0.4586
17	ID	Sequence	pep	LEQWVSGRK	aff	0.0649	aff_rescale	0.2756	cle	0.9323	tap	0.3100	COMB	0.4309
18	ID	Sequence	pep	EQWVSGRKR	aff	0.0499	aff_rescale	0.2117	cle	0.5458	tap	1.5760	COMB	0.3723
19	ID	Sequence	pep	QWVSGRKRL	aff	0.0509	aff_rescale	0.2160	cle	0.9351	tap	1.1250	COMB	0.4125
20	ID	Sequence	pep	WVSGRKRLE	aff	0.0543	aff_rescale	0.2308	cle	0.0246	tap	-1.5070	COMB	0.1591
21	ID	Sequence	pep	VSGRKRLEE	aff	0.0654	aff_rescale	0.2777	cle	0.0259	tap	-1.6820	COMB	0.1975
22	ID	Sequence	pep	SGRKRLEEL	aff	0.0552	aff_rescale	0.2346	cle	0.9283	tap	0.8370	COMB	0.4157
23	ID	Sequence	pep	GRKRLEELE	aff	0.0451	aff_rescale	0.1915	cle	0.0229	tap	-1.4860	COMB	0.1207
24	ID	Sequence	pep	RKRLEELEX	aff	0.0543	aff_rescale	0.2304	cle	0.2200	tap	0.3270	COMB	0.2797
25	ID	Sequence	pep	KRLEELEXD	aff	0.0517	aff_rescale	0.2197	cle	0.0349	tap	-1.3520	COMB	0.1573
26	ID	Sequence	pep	RLEELEXDL	aff	0.0657	aff_rescale	0.2791	cle	0.7337	tap	1.1750	COMB	0.4479
27	ID	Sequence	pep	LEELEXDLR	aff	0.0561	aff_rescale	0.2383	cle	0.0455	tap	1.3410	COMB	0.3121
28	ID	Sequence	pep	EELEXDLRK	aff	0.0589	aff_rescale	0.2502	cle	0.5869	tap	0.1910	COMB	0.3478
29	ID	Sequence	pep	ELEXDLRKV	aff	0.0581	aff_rescale	0.2466	cle	0.5310	tap	-0.1810	COMB	0.3172
30	ID	Sequence	pep	LEXDLRKVK	aff	0.0478	aff_rescale	0.2028	cle	0.2014	tap	0.1730	COMB	0.2416
31	ID	Sequence	pep	EXDLRKVKK	aff	0.0587	aff_rescale	0.2494	cle	0.8738	tap	0.1500	COMB	0.3880
32	ID	Sequence	pep	XDLRKVKKK	aff	0.0516	aff_rescale	0.2190	cle	0.5549	tap	-0.0290	COMB	0.3008
33	ID	Sequence	pep	DLRKVKKKI	aff	0.0516	aff_rescale	0.2193	cle	0.8899	tap	0.2850	COMB	0.3670
34	ID	Sequence	pep	LRKVKKKIK	aff	0.0439	aff_rescale	0.1866	cle	0.7523	tap	0.5340	COMB	0.3261
35	ID	Sequence	pep	RKVKKKIKK	aff	0.0579	aff_rescale	0.2457	cle	0.8355	tap	0.8840	COMB	0.4152
36	ID	Sequence	pep	KVKKKIKKL	aff	0.0510	aff_rescale	0.2165	cle	0.9646	tap	1.3300	COMB	0.4277

Fig 3.4.1d :NetCTL-1.2 Server results of Hepatitis D



NetCTL-1.2 Server - prediction results

Technical University of Denmark

NetCTL-1.2 predictions using MHC supertype A1. Threshold 0.750000

369	ID	Sequence	pep	LTAVITAAY	aff	0.7106	aff_rescale	3.0172	cle	0.5935	tap	3.0780	COMB	3.2601 <-E
645	ID	Sequence	pep	VTAFCSALY	aff	0.6694	aff_rescale	2.8420	cle	0.7767	tap	2.8150	COMB	3.0993 <-E
1653	ID	Sequence	pep	CVDVVSQVV	aff	0.6015	aff_rescale	2.5539	cle	0.9724	tap	2.8930	COMB	2.8444 <-E
602	ID	Sequence	pep	ATGPMSLSY	aff	0.5660	aff_rescale	2.4029	cle	0.9783	tap	2.9200	COMB	2.6957 <-E
156	ID	Sequence	pep	AAETGIALY	aff	0.4768	aff_rescale	2.0244	cle	0.9766	tap	2.8720	COMB	2.3145 <-E
662	ID	Sequence	pep	HSLIGGLWY	aff	0.4371	aff_rescale	1.8560	cle	0.8380	tap	3.0240	COMB	2.1329 <-E
1289	ID	Sequence	pep	TTSDSVLTF	aff	0.4329	aff_rescale	1.8380	cle	0.9737	tap	2.5360	COMB	2.1109 <-E
1316	ID	Sequence	pep	VLSTLVGRY	aff	0.4209	aff_rescale	1.7871	cle	0.9516	tap	3.1560	COMB	2.0876 <-E
1299	ID	Sequence	pep	LTDIVHCRM	aff	0.3941	aff_rescale	1.6732	cle	0.9725	tap	-0.0560	COMB	1.8163 <-E
513	ID	Sequence	pep	ATVSISGSY	aff	0.3914	aff_rescale	1.6617	cle	0.9622	tap	2.9580	COMB	1.9539 <-E
1541	ID	Sequence	pep	NMAVIAHCY	aff	0.3886	aff_rescale	1.6498	cle	0.9655	tap	3.0440	COMB	1.9468 <-E
973	ID	Sequence	pep	HVEPGVVHY	aff	0.3845	aff_rescale	1.6327	cle	0.9786	tap	2.8210	COMB	1.9205 <-E
341	ID	Sequence	pep	MTYLRGISY	aff	0.3746	aff_rescale	1.5907	cle	0.9593	tap	3.1850	COMB	1.8938 <-E
1206	ID	Sequence	pep	ISDAIVNNF	aff	0.3570	aff_rescale	1.5157	cle	0.9445	tap	2.3920	COMB	1.7769 <-E
377	ID	Sequence	pep	YLTICHQRY	aff	0.3237	aff_rescale	1.3745	cle	0.9701	tap	3.0420	COMB	1.6722 <-E
42	ID	Sequence	pep	QTEILINLM	aff	0.2983	aff_rescale	1.2667	cle	0.7109	tap	0.2000	COMB	1.3834 <-E
443	ID	Sequence	pep	HLDPRVLVF	aff	0.2947	aff_rescale	1.2512	cle	0.9609	tap	2.1820	COMB	1.5045 <-E
204	ID	Sequence	pep	HTTSYLLIH	aff	0.2317	aff_rescale	0.9838	cle	0.1019	tap	-0.7360	COMB	0.9623 <-E
925	ID	Sequence	pep	LTEPAIAWF	aff	0.2251	aff_rescale	0.9557	cle	0.9541	tap	2.3210	COMB	1.2149 <-E
1159	ID	Sequence	pep	FTETTIIAT	aff	0.2132	aff_rescale	0.9053	cle	0.0424	tap	-0.9560	COMB	0.8638 <-E
846	ID	Sequence	pep	IMNDGFAAY	aff	0.2106	aff_rescale	0.8940	cle	0.9787	tap	2.8700	COMB	1.1843 <-E
1355	ID	Sequence	pep	TTCELYELV	aff	0.2016	aff_rescale	0.8560	cle	0.1705	tap	0.3280	COMB	0.8979 <-E
951	ID	Sequence	pep	RTANLALEL	aff	0.2013	aff_rescale	0.8545	cle	0.9509	tap	1.3480	COMB	1.0645 <-E
225	ID	Sequence	pep	SSAGYNHDV	aff	0.1997	aff_rescale	0.8481	cle	0.6694	tap	0.2860	COMB	0.9628 <-E
1290	ID	Sequence	pep	TSDSVLTFE	aff	0.1843	aff_rescale	0.7827	cle	0.0276	tap	-1.7120	COMB	0.7012
576	ID	Sequence	pep	FVDGAQLEA	aff	0.1705	aff_rescale	0.7241	cle	0.8012	tap	-0.6210	COMB	0.8132 <-E
407	ID	Sequence	pep	ITRLYSWLF	aff	0.1670	aff_rescale	0.7091	cle	0.9691	tap	2.4800	COMB	0.9785 <-E
1563	ID	Sequence	pep	DSVVLCSDY	aff	0.1666	aff_rescale	0.7074	cle	0.9517	tap	2.5880	COMB	0.9796 <-E
886	ID	Sequence	pep	CSRRGTAAY	aff	0.1638	aff_rescale	0.6956	cle	0.8648	tap	2.9830	COMB	0.9745 <-E
808	ID	Sequence	pep	ESDCTWLVN	aff	0.1617	aff_rescale	0.6866	cle	0.0817	tap	-1.7600	COMB	0.6108
1695	ID	Sequence	pep	LTSSIIYRM	aff	0.1505	aff_rescale	0.6388	cle	0.9721	tap	0.1440	COMB	0.7918 <-E
554	ID	Sequence	pep	VSDAQNRLT	aff	0.1488	aff_rescale	0.6318	cle	0.0268	tap	-0.9460	COMB	0.5886
364	ID	Sequence	pep	ASEDALTAV	aff	0.1460	aff_rescale	0.6199	cle	0.5054	tap	0.4960	COMB	0.7205
335	ID	Sequence	pep	FCCSRLMTY	aff	0.1449	aff_rescale	0.6154	cle	0.3242	tap	3.0130	COMB	0.8146 <-E
1417	ID	Sequence	pep	WSKTFCALF	aff	0.1440	aff_rescale	0.6114	cle	0.3204	tap	2.7000	COMB	0.7944 <-E
1115	ID	Sequence	pep	QTASRVLRS	aff	0.1397	aff_rescale	0.5932	cle	0.1023	tap	-2.2280	COMB	0.4972

Fig 3.4.1e :NetCTL-1.2 Server results of Hepatitis E

2.4.2. IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE (IEDB):

Here, the amino acid sequences of Hepatitis viruses are taken and the identification of Epitopes are done. The results are displayed below.(Figure: 3.4.2)

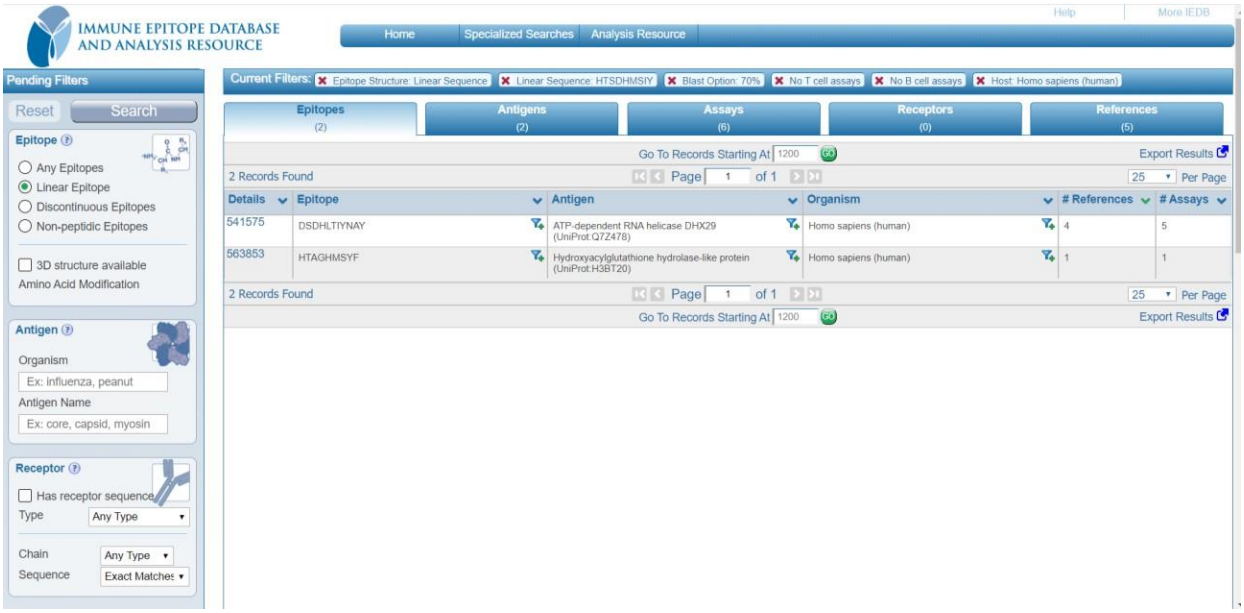


Fig 3.4.2a: IEDB results of Hepatitis A

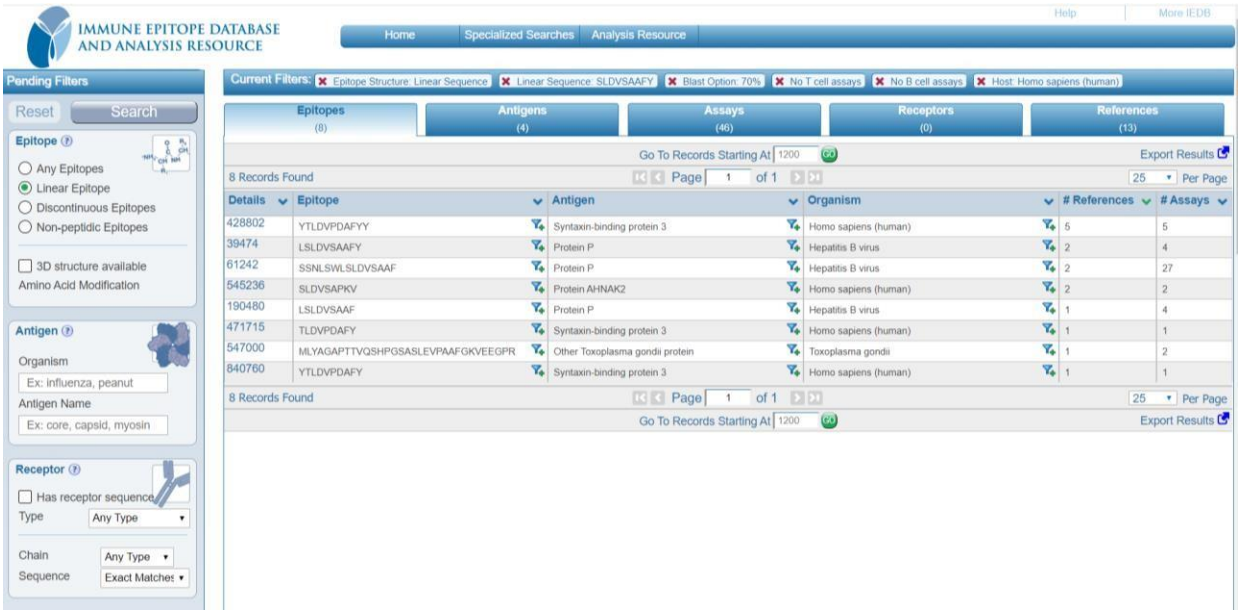


Fig 3.4.2b: IEDB results of Hepatitis B

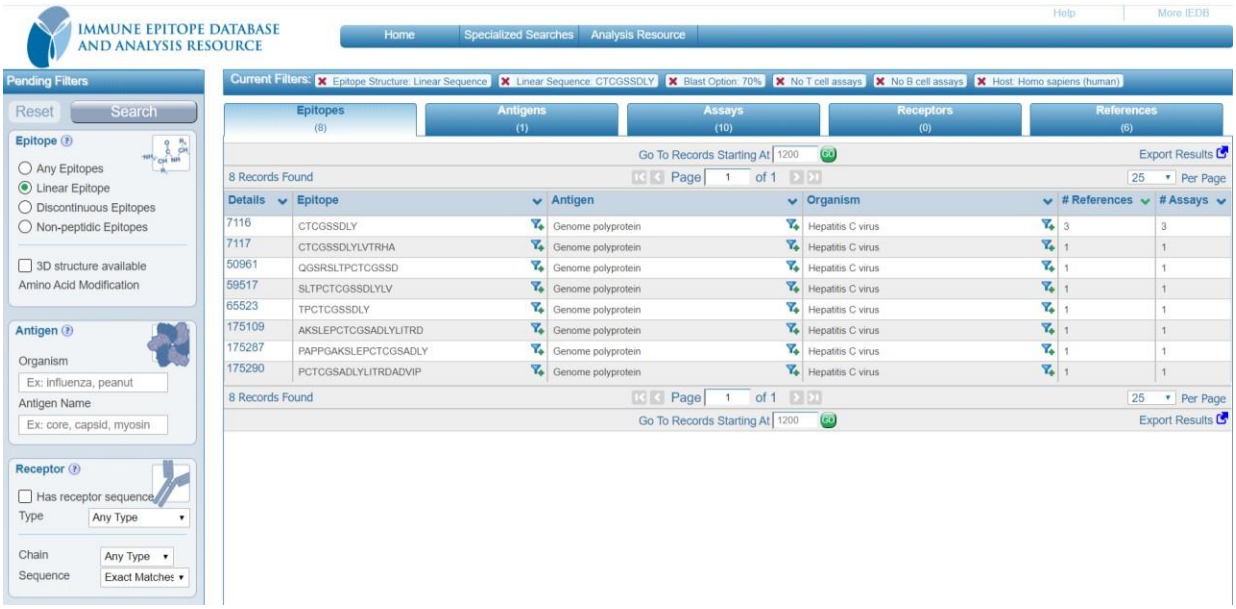


Fig 3.4.2c: IEDB results of Hepatitis C

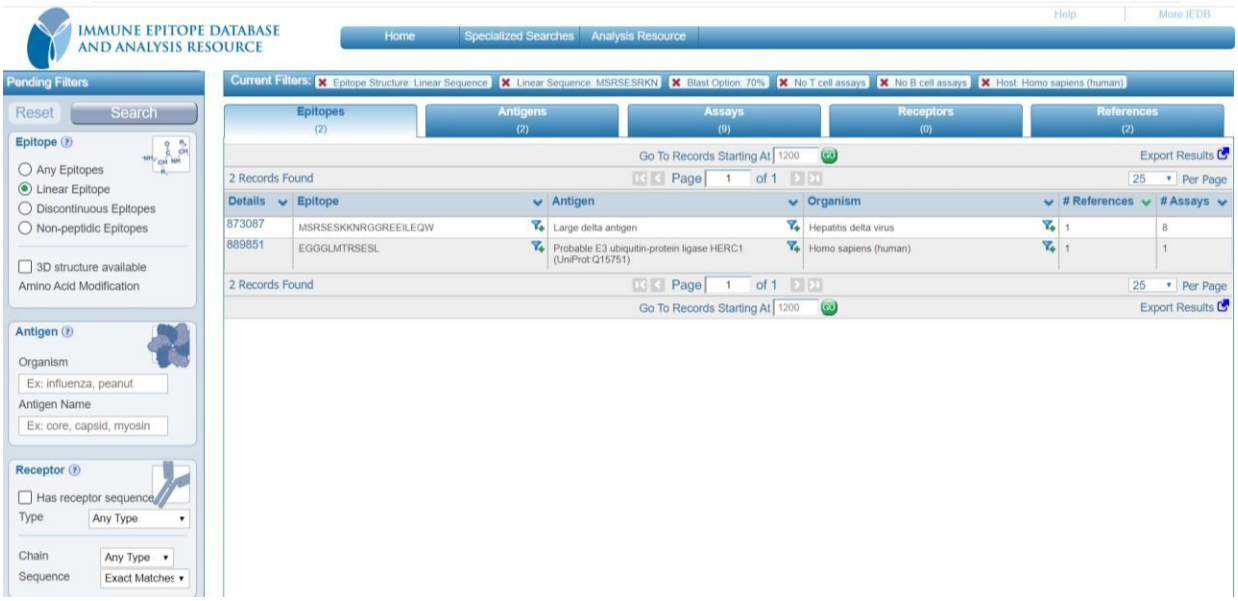


Fig 3.4.2d: IEDB results of Hepatitis D

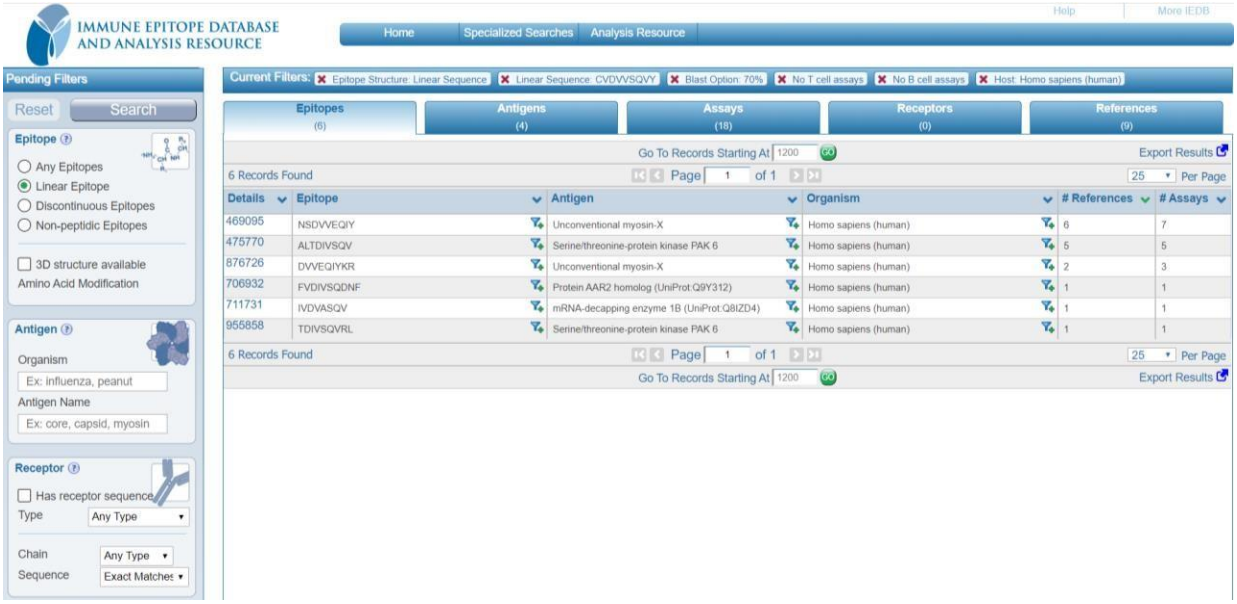


Fig 3.4.2e: IEDB results of Hepatitis E

PREDICTION OF B-CELL EPITOPE

3.5.1. Antibody epitope prediction:

Here, the amino acid sequences of Hepatitis viruses are taken and the identification of B cell Epitopes are done. The results are displayed below. (Figure: 3.5.1)

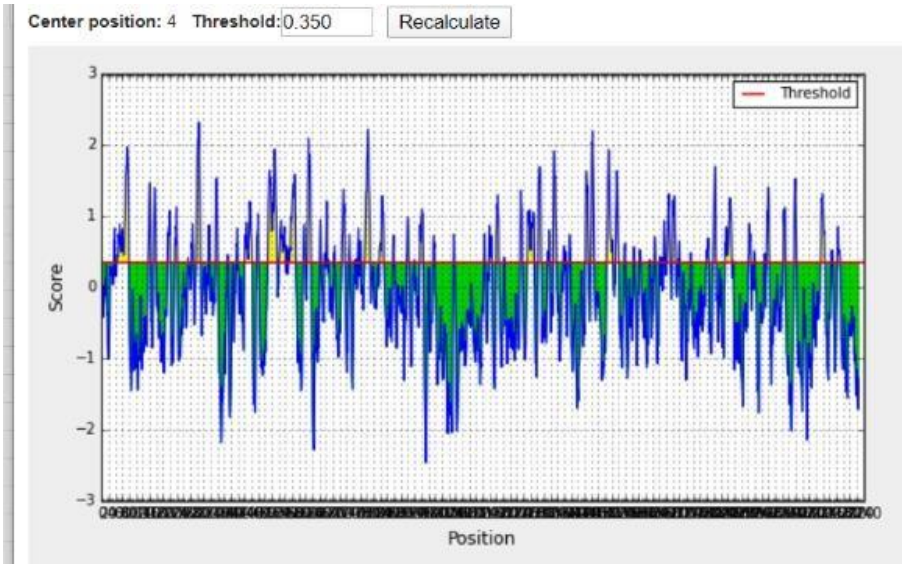


Figure: 3.5.1a Results of Hepatitis A virus

Predicted peptides:

No.	Start	End	Peptide	Length
2	45	79	FTSVDQSSVHTAEVGSHQVEPLRTSVDKPGSKRTQ	35
16	489	516	TTQVGDDSGGFSTTVSTEQNVPDPQVGI	28
14	421	442	ISDTPYRVNRYTKSAHQKGEYT	22
19	550	571	VLAKKVPETFPELKPGESRHTS	22
17	521	541	DLKGKANRGKMDVSGVQAPVG	21
47	1252	1272	GVEPEKNIYTKPVASDYWDGY	21
74	1834	1853	TGAPGIDAINMDSSPGFPYV	20
57	1486	1503	HKEEEPIPTEGVYHGVTK	18
34	774	790	DLESSVDDPRSEEDKRF	17
35	817	830	EELSNEVLPPPRKM	14

Figure: 3.5.1b Results of Hepatitis A virus

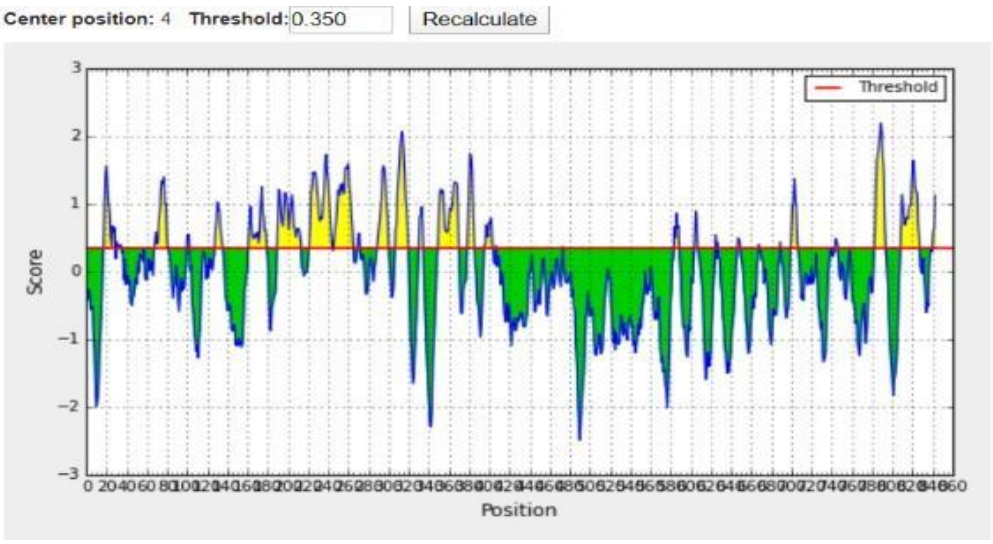


Figure: 3.5.1c: Results of Hepatitis B virus

Predicted peptides:

No.	Start	End	Peptide	Length
7	190	213	KRHGDKSFCPQSPGILPRSSVGPC	24
8	222	244	RLGPQPAQRQLAGRQQGGSGSIR	23
14	349	369	WGPCTEHGEHRIRTPRTPARV	21
9	246	265	RVHPSPWGTVGVEPSGSGHI	20
30	809	827	TTGRTSLYADSPSVPSHLP	19
6	161	178	ESTRSASFCGSPYSWEQD	18
3	68	80	PCFNPKWQTPSFP	13
12	307	319	FPPNSSRSQSQGP	13
1	17	28	EAGPLEEELPRL	12
11	289	300	ISTSQGHSSSGH	12
29	783	793	ALNPADDPSRG	11
16	395	403	FSRGNTRVS	9
5	128	135	GIKPYYPE	8
15	379	385	NPHNTTE	7

Figure: 3.5.1d: Results of Hepatitis B virus

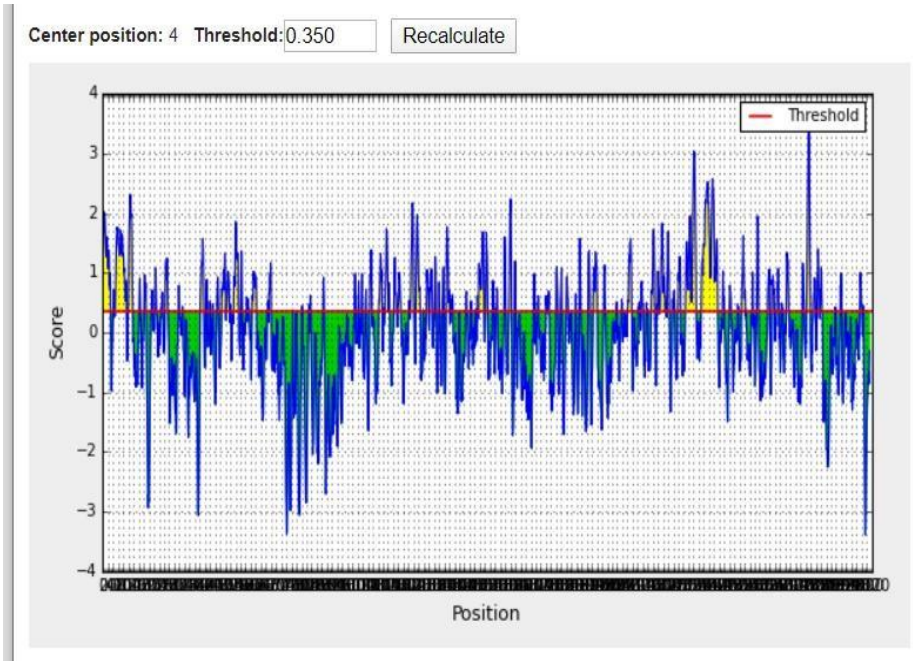


Figure: 3.5.1e: Results of Hepatitis C virus

Predicted peptides:

No.	Start	End	Peptide	Length
98	2351	2415	SFGSSSTSGITGDNTTTSSEPAPSGCPPDSDVESYSSMPPLEGEPGDPDLSDGSWSTVSSGADTE	65
3	49	92	TRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCG	44
97	2288	2330	WARPDYNPPLVETWKKPDYEPPVVHGCPLPPPRSPPVPPPRKK	43
1	1	30	MSTNPKPQRKTKRNTNRRPQDVKFPGGGQI	30
92	2188	2212	RRLARGSPPSMASSSASQLSAPSLK	25
61	1475	1497	TTLPQDAVSRTQRRGRTGRGKPG	23
21	513	534	PVVVGTTDRSGAPTYSWGANDT	22
4	100	117	PRGSRPSWGPTDPRRRSR	18
25	586	603	FRKHPEATYSRCGSGPWI	18
18	467	483	QGWGPISYANGSGLDER	17
46	1230	1246	APTGSGKSTKVPAAYAA	17
78	1926	1942	GNHVSPTHYVPESDAAA	17
82	2061	2077	PINAYTTGPCTPLPAPN	17
89	2154	2169	YPVGSQLPCEPEPDVA	16
102	2509	2524	CSLTPPHSAKSKFGYG	16
113	2764	2779	TRYSAPPGDPPQPEYD	16
45	1208	1222	PVFTDNSSPPAVPQS	15

Figure: 3.5.1f: Results of Hepatitis C virus

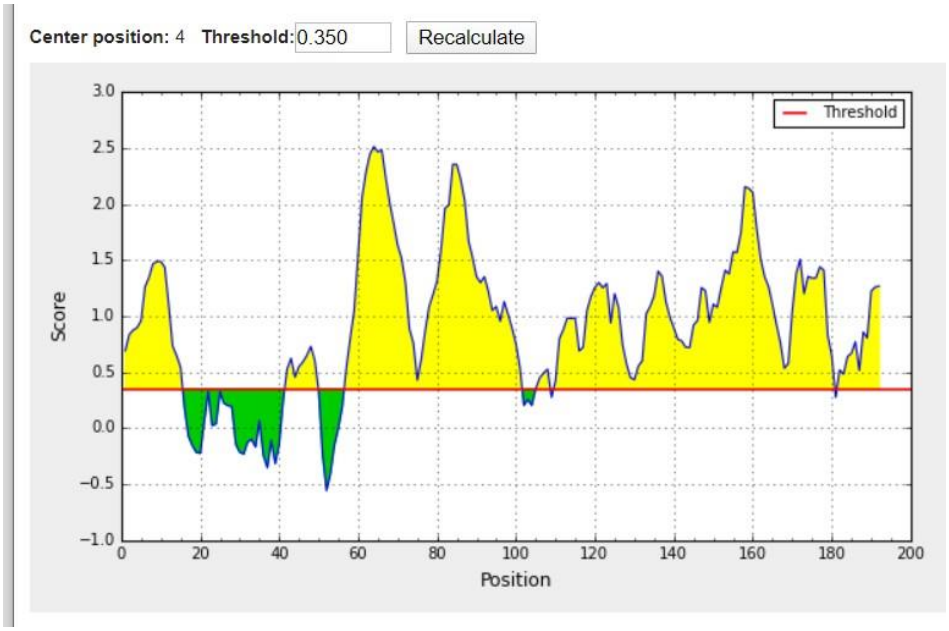


Figure: 3.5.1g: Results of Hepatitis D virus

Predicted peptides:

No.	Start	End	Peptide	Length
5	110	180	KKQLTSGGKNLSKEEEEELGRLTKEDEERKRRVAGPRVGDVNPFEGGSRGAPGGGFVPNMQGVPESPFTRT	71
3	57	101	GKKDKDGEGAPPAKRARTDRMEVDSGPGKRPSRGGFTEQERRDHR	45
1	1	15	MSRSESRKNRGGRED	15
2	42	49	KLEEDPWL	8
4	105	108	ALEN	4

Figure: 3.5.1h: Results of Hepatitis D virus

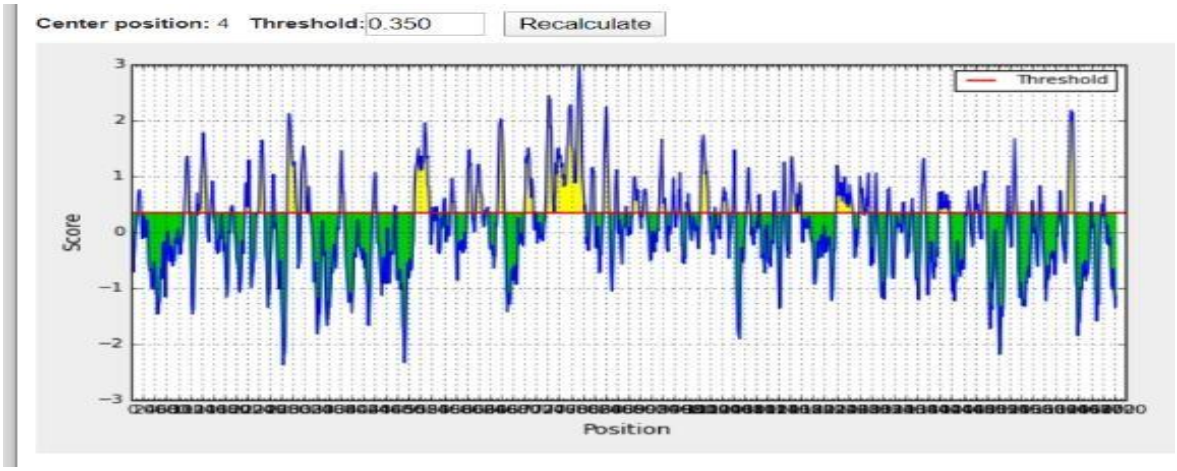


Figure: 3.5.1i: Results of Hepatitis E virus

Predicted peptides:

No.	Start	End	Peptide	Length
28	717	786	LEPCNQDPPPSVNIDLPVNTPSSVTPVQTQAPALEQVAPPSDLPDGGAGPMLGAPIPPPTPPQSITRLSE	70
50	1220	1249	QVGQHRPSVIPRGTIDNNVDTLDAFPPSCQ	30
17	490	516	GRVGEQGYDNEAFEGSDVDPAEEATVS	27
59	1397	1416	NKFTTGETIAHGKVGQGISA	20
11	270	286	APEPSPMPYVPYPRSTE	17
26	681	696	PFSPGHSWESANPFCG	16
5	116	130	QRWYTAPTRGPAANC	15
40	985	999	AGVPGSGKSRSIQQG	15
25	633	645	PPGVATPSAPGEV	13
36	913	925	WERNHRPGDELYL	13
69	1621	1633	SEKNWGPGPERAE	13
23	596	607	PAKQTMATGPHS	12
48	1141	1152	AKAANPGAITVH	12
9	220	230	TYEGDSSAGYN	11
12	293	303	FGPGGSPSLFP	11
33	868	878	DYRVEHNPKRL	11

Figure: 3.5.1j: Results of Hepatitis E virus

ALLERGENCITY PREDICTION

Allertop V 2.0:

With the help of this tool, the allergenicity prediction of various sequences was done. The epitope sequence of various Hepatitis Viruses was uploaded in the Allertop V 2.0 and the results are displayed below.

S.no	Virus	Epitopes	Nature
1.	HEPATITIS A	HTSDHMSIY	Probable Allergen
2.	HEPATITIS B	SLDVSAAFY	Probable Non-allergen
3.	HEPATITIS C	CTCGSSDLY	Probable Allergen
4.	HEPATITIS D	MSRSESRKN	Probable Non-allergen
5.	HEPATITIS E	CVDVVSQVY	Probable Allergen

Table 3.6.1: Determination of Allergenicity for Epitope sequences of Hepatitis virus using Allertop 2.0

AllergenFP v.1.0:

With the help of this tool, the allergenicity prediction of various sequences was done. The epitope sequences of various Hepatitis Virus was uploaded and the results are displayed below.

S.no	Virus	Epitopes	Result
1.	HEPATITIS A	HTSDHMSIY	Probable Allergen

2.	HEPATITIS B	SLDVSAAFY	Probable Non-allergen
3.	HEPATITIS C	CTCGSSDLY	Probable Non-allergen
4.	HEPATITIS D	MSRSESRKN	Probable Allergen
5.	HEPATITIS E	CVDVVSQVY	Probable Allergen

Table 3.6.2: Determination of Allergenicity for Epitope sequences of Hepatitis virus using Allergen

3D MODELLING OF AN EPITOPE

PEP FOLD:

Using the PEP-FOLD analysis, the structure of various amino acid sequences were predicted. Here, the amino acid sequences are uploaded and the results are displayed below. (figure: 3.7.1)

The sequence HTSDHMSIY of Hepatitis A virus displayed the following results:

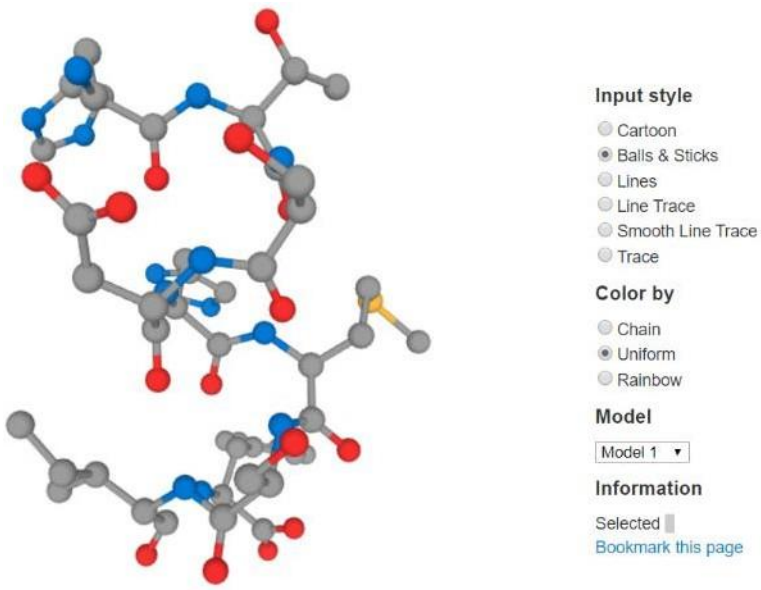


Fig 3.7.1a: result of HTSDHMSIY of Hepatitis A

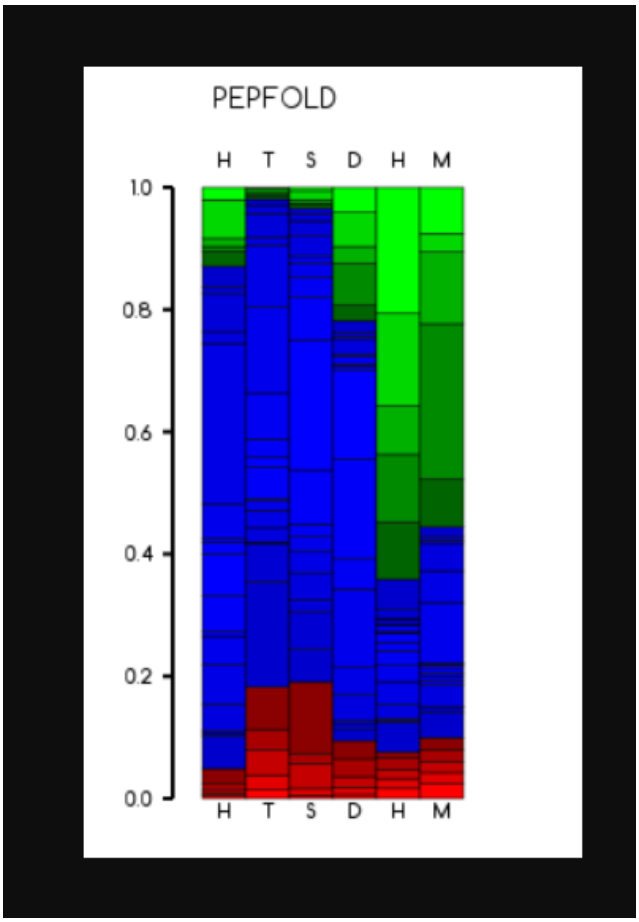

PEPFOLD
H T S D H M
1.0
0.8
0.6
0.4
0.2
0.0
H T S D H M

Fig 3.7.1b: result of HTSDHMSIY of Hepatitis A

The sequence SLDVSAAFY of Hepatitis B virus displayed the following results:

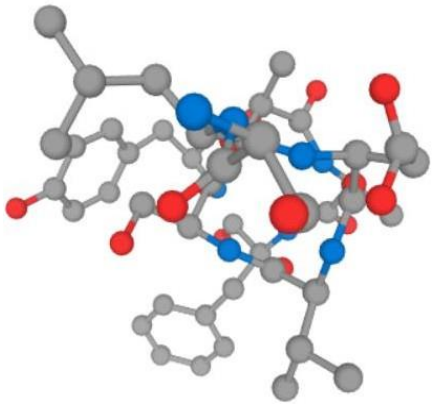


Fig 3.7.1c: result of SLDVSAAFY of Hepatitis B

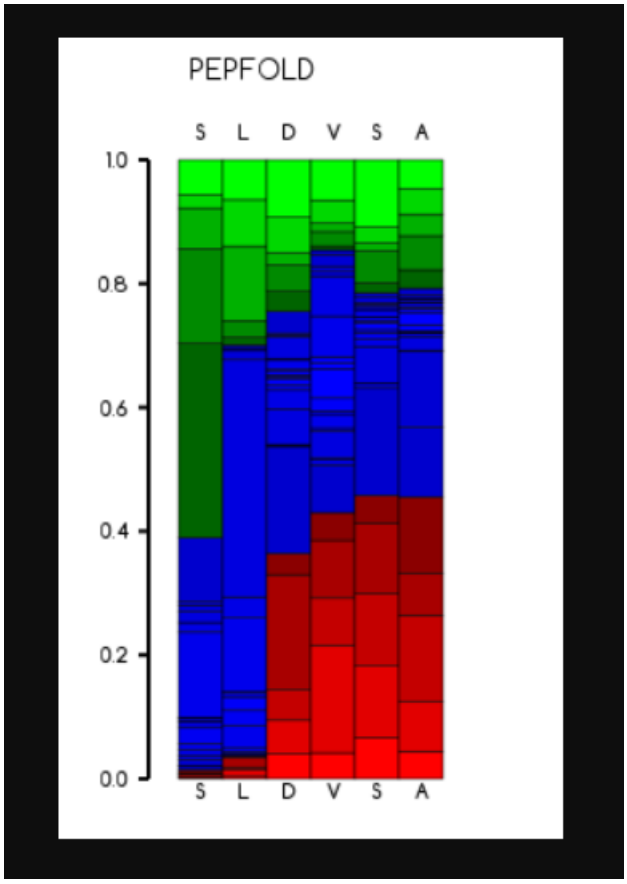


Fig 3.7.1d: result of SLDVSAAFY of Hepatitis B

The sequence CTCGSSDLY of Hepatitis C virus displayed the following results:

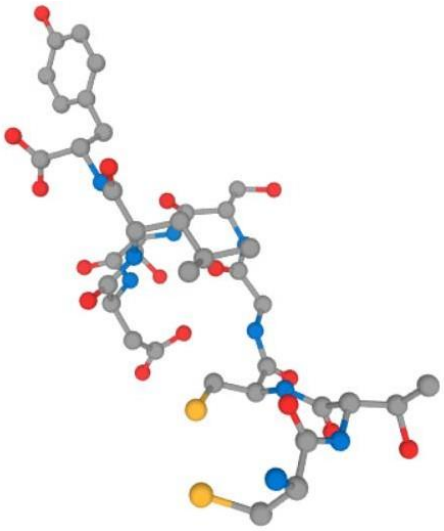


Fig 3.7.1e: result of CTCGSSDLY of Hepatitis C

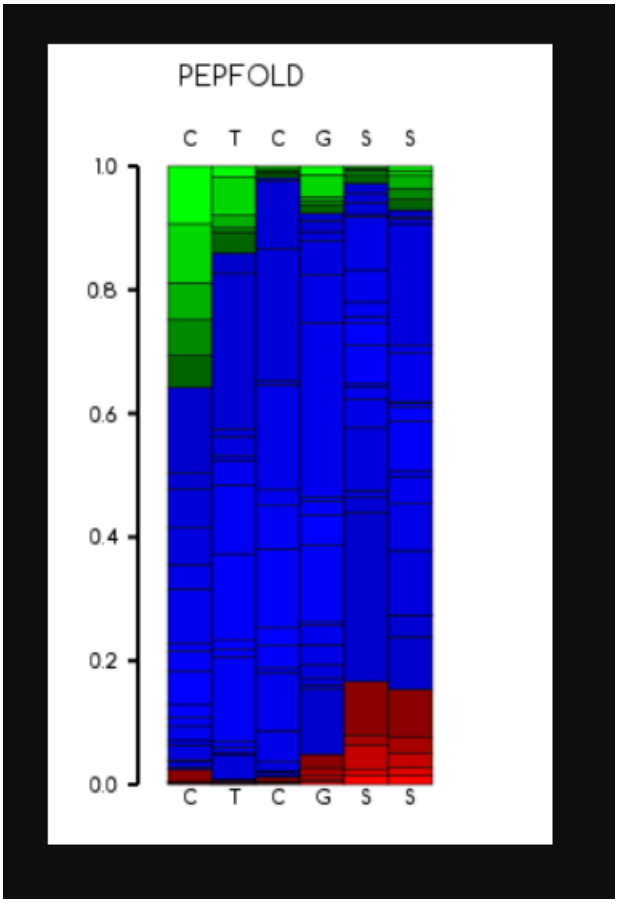


Fig 3.7.1f: result of CTCGSSDLY of Hepatitis C

The sequence MSRSESRKN of Hepatitis D virus displayed the following results:

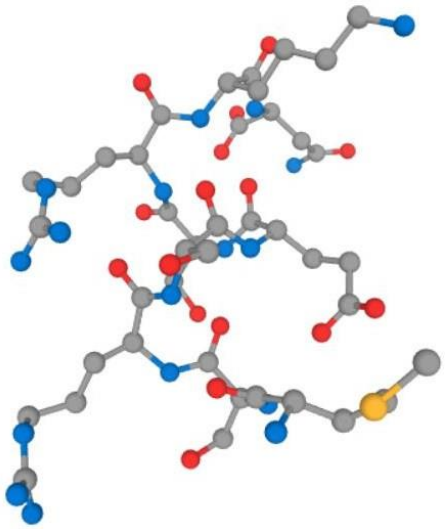


FIG 3.7.1g: Result of MSRSESRKN of Hepatitis D

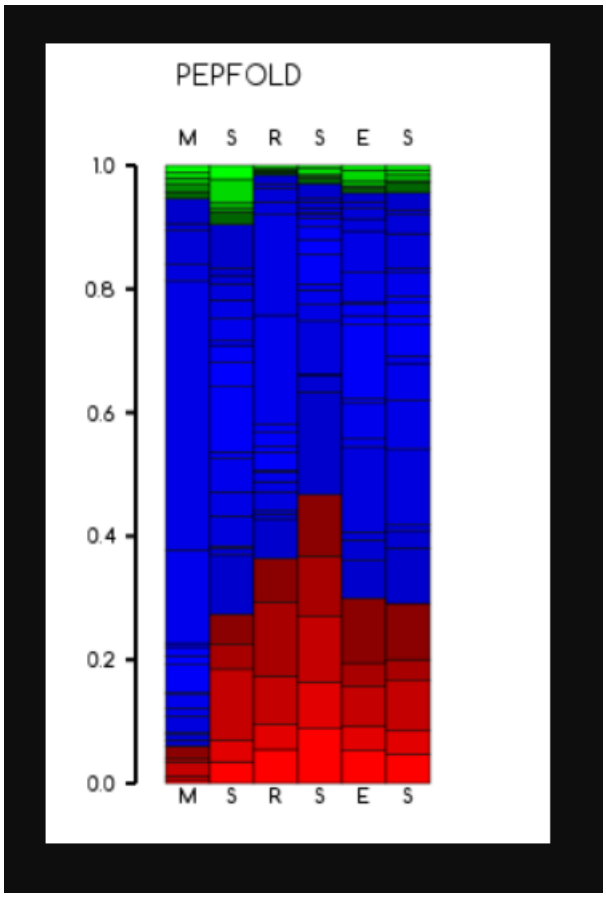


FIG 3.7.1h: Result of MSRSESRKN of Hepatitis D

The sequence CVDVVSQVY of Hepatitis E displayed the following results:

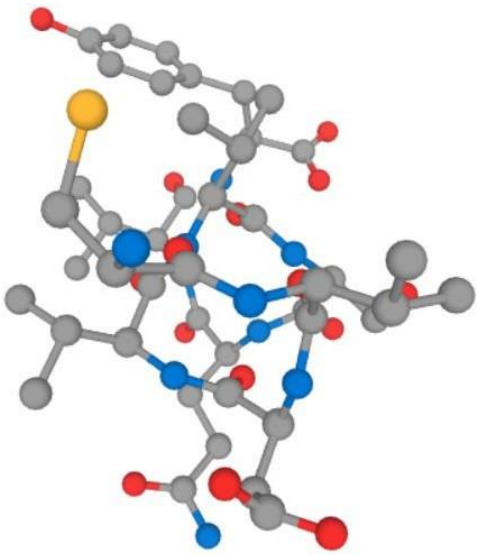


FIG 3.7.1i: result of CVDVVSQVY of Hepatitis E

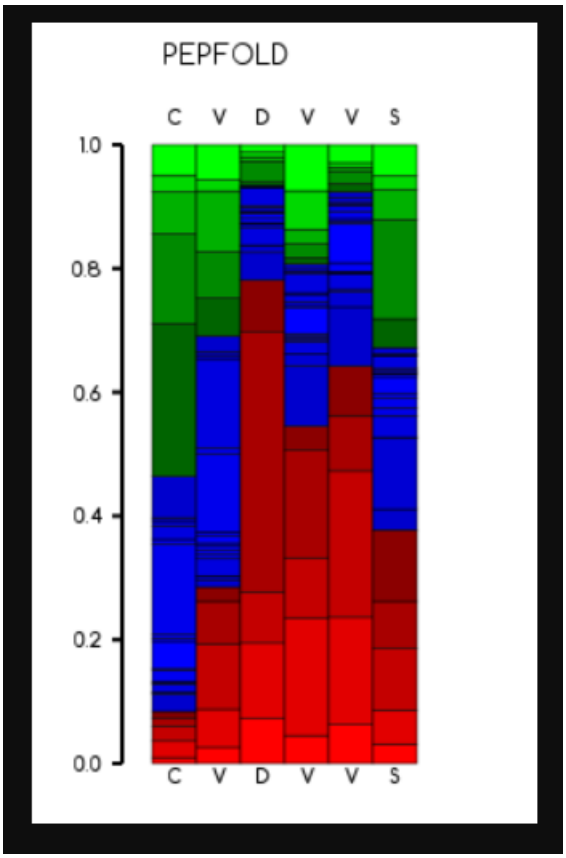


FIG 3.7.1j: result of CVDVVSQVY of Hepatitis E

PHYRE2:

Using the PHYRE2 analysis, the structure of various amino acid sequences were predicted. Here, the amino acid sequences are uploaded and the results are displayed below. (figure: 3.7.1)

The amino acid sequence of Hepatitis A virus displayed the following results:

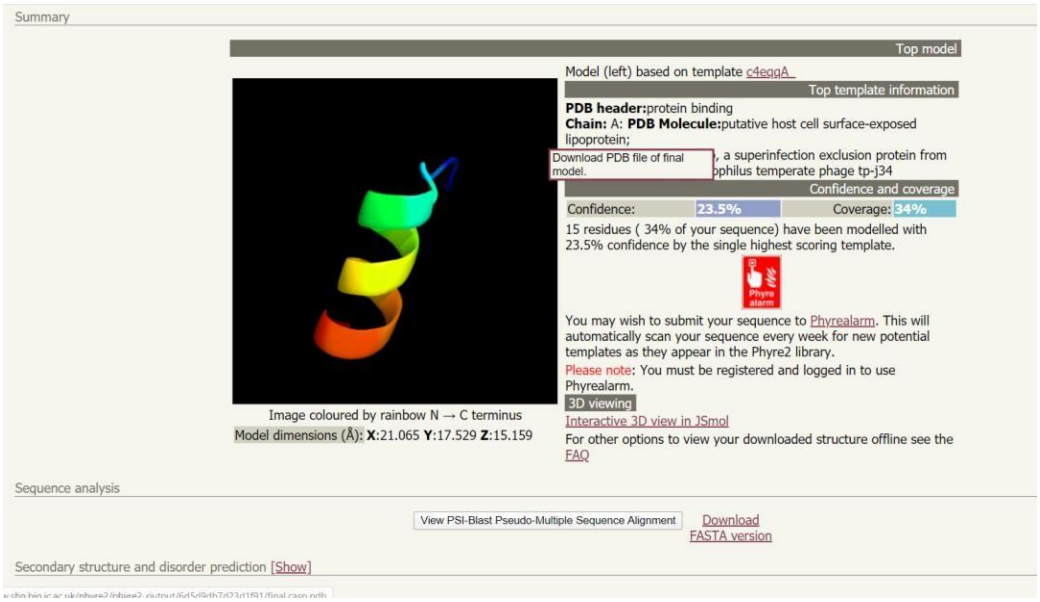


Figure: 3.7.2a: Results of Hepatitis A virus

The amino acid sequence of Hepatitis B virus displayed the following results:

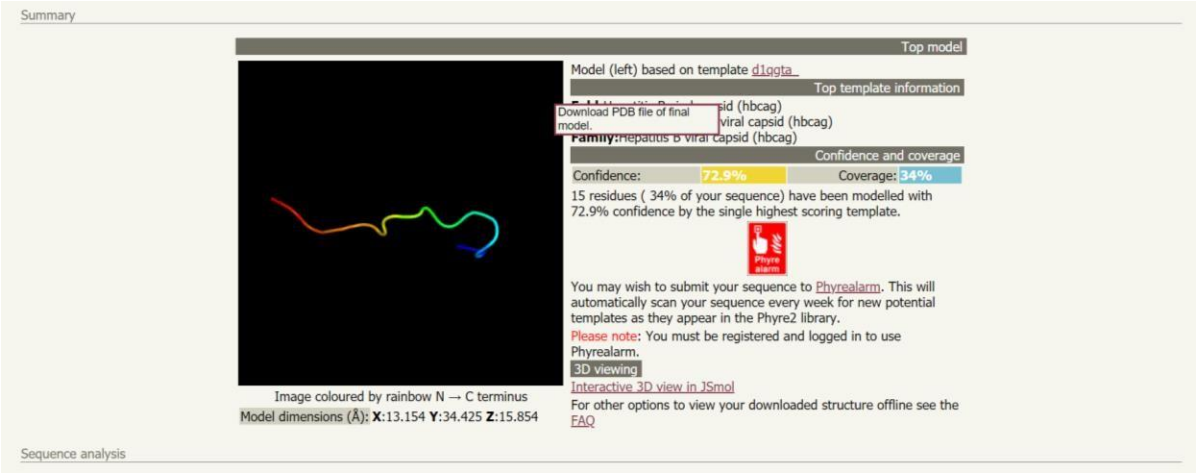


Figure: 3.7.2b: Results of Hepatitis B virus

The amino acid sequence of Hepatitis C virus displayed the following results:

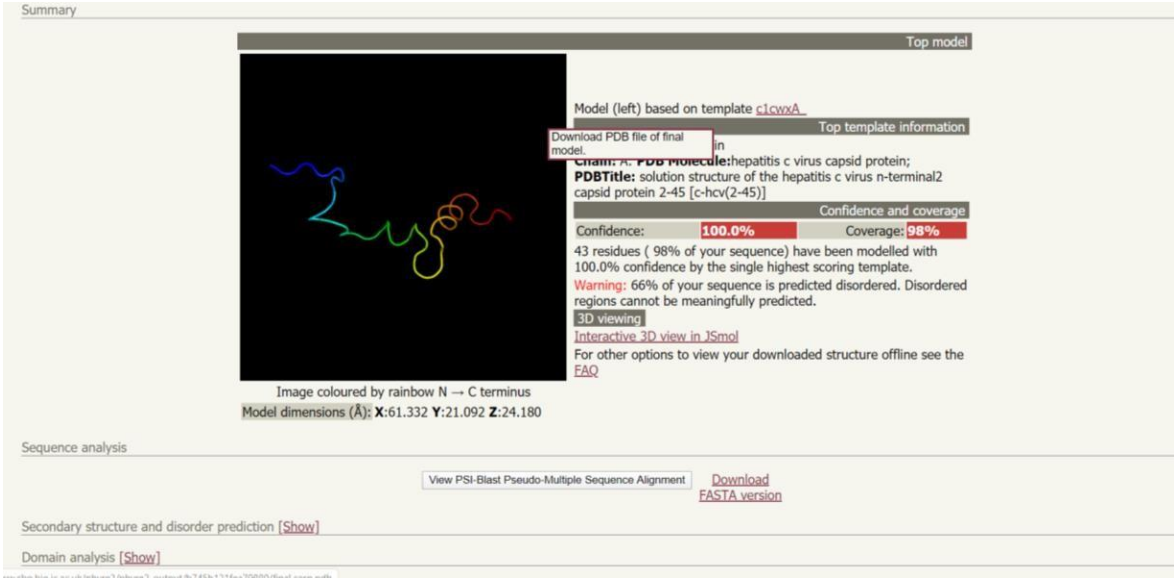


Figure: 3.7.2c: Results of Hepatitis C virus

The amino acid sequence of Hepatitis D virus displayed the following results:

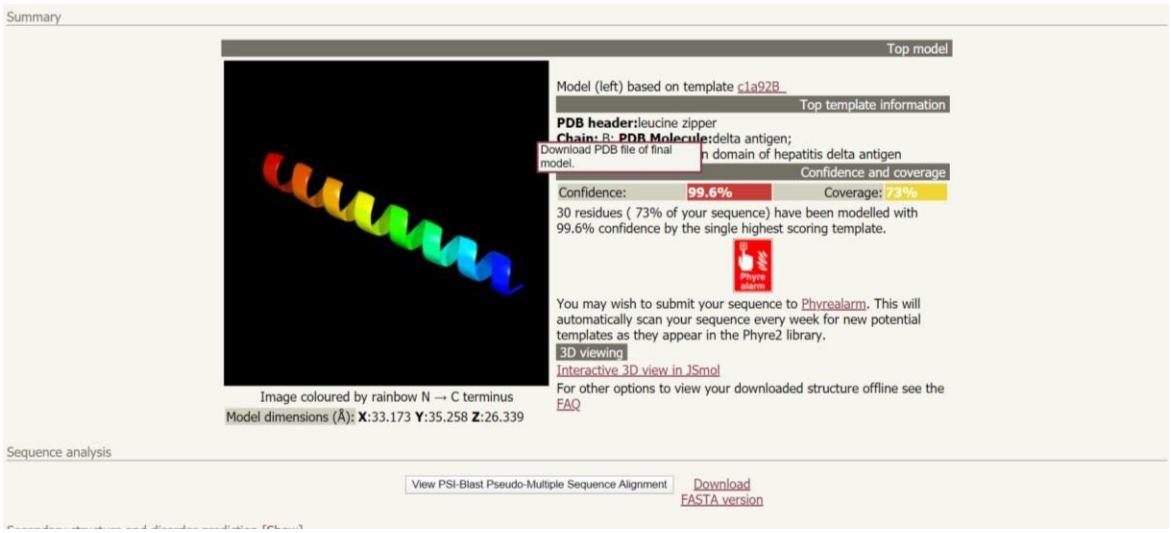


Figure: 3.7.2d: Results of Hepatitis D virus

The amino acid sequence of Hepatitis E virus displayed the following results:

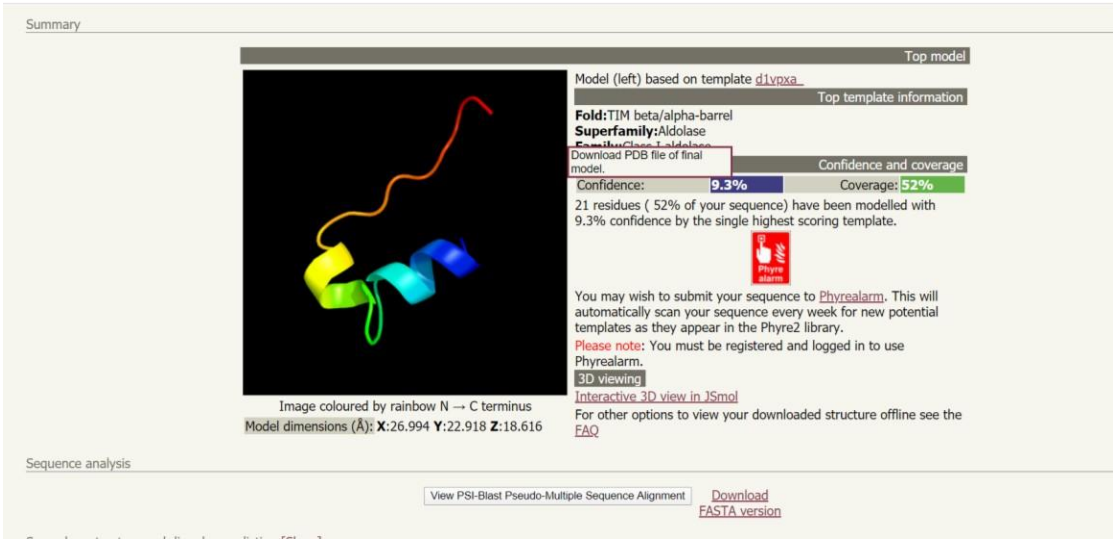


Figure: 3.7.2e: Results of Hepatitis E virus

The peptide sequence of Hepatitis C Virus displayed the following results:

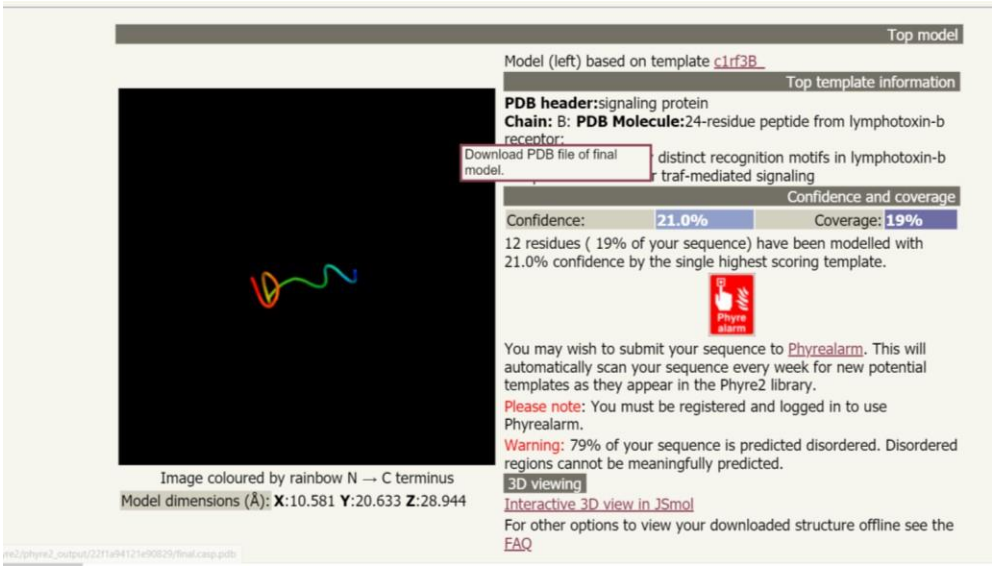


Figure: 3.7.2 f: Results of Hepatitis C virus

The peptide sequence of Hepatitis C Virus displayed the following results:

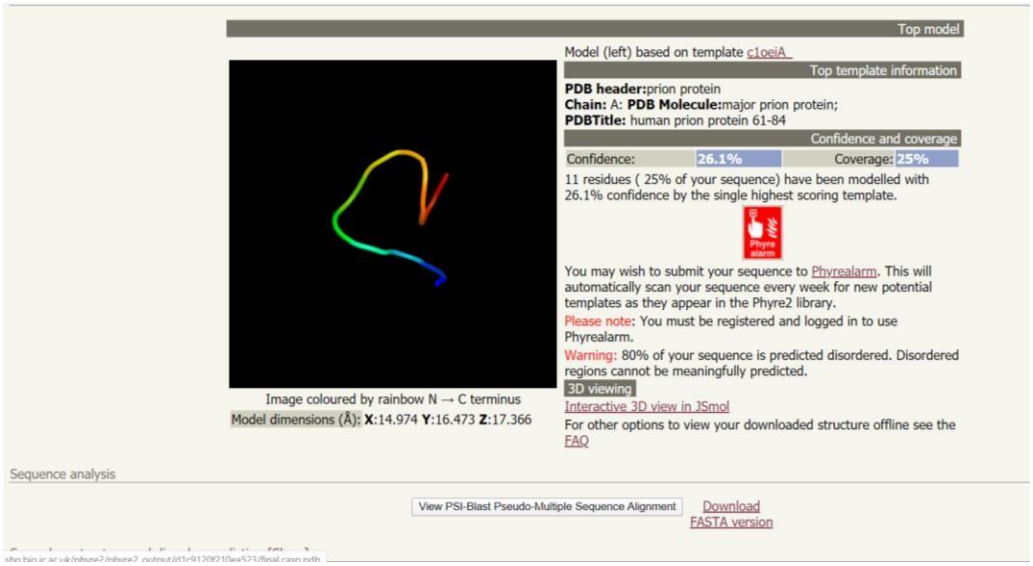


Figure: 3.7.2 g: Results of Hepatitis C virus

The peptide sequence of Hepatitis D Virus displayed the following results:

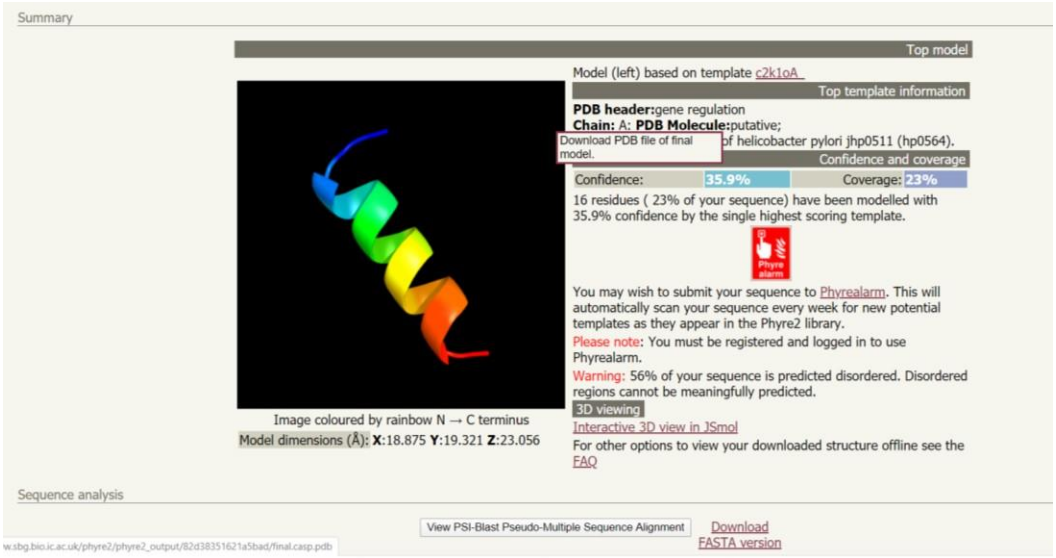


Figure: 3.7.2 h: Results of Hepatitis D virus

The peptide sequence of Hepatitis E Virus displayed the following results:

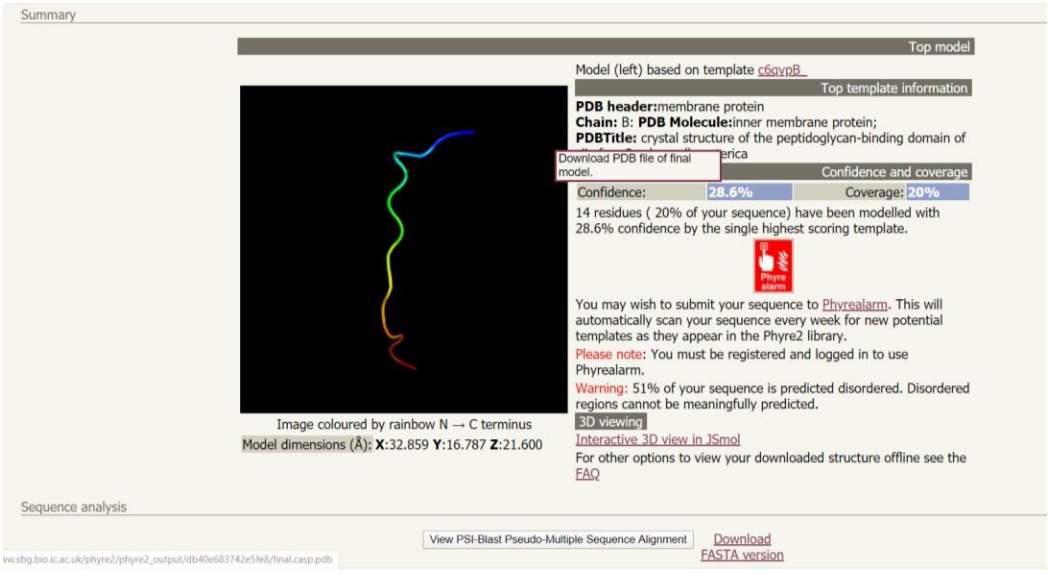


Figure: 3.7.2 i: Results of Hepatitis E virus

Chapter 4 CONCLUSION

The overall intention of this study is to find out a novel epitope as an antigenic determinant against Hepatitis Virus. Upon the infection of the virus, the virus produces nucleocapsid proteins, which tend to play a significant role in number of processes including viral attachment, membrane fusion, and entry into the host cell. The host cell generates viral neutralizing antibodies and cytotoxic T cells in response to the viral infection as a part of immunogenic activity. The principal target for the neutralizing antibodies is the nucleocapsid protein for the abruption of Hepatitis virus infection. It has been shown that the defense against Hepatitis infection in human system is mostly antibody dependent, and virus- neutralizing antibodies are sufficient to impart

protection. The peptides present on Hepatitis envelope proteins are able to propose the specific antigenic determinants which are capable enough to induce Hepatitis neutralizing antibodies.

The B-cell and T-cell epitopes on Hepatitis nucleocapsid protein can prove to be an important tool for the protection against virus infection. The small peptidic epitopes can be used to derive the peptide vaccines, which if used strategically and sensibly can be used to induce immune reaction in human systems. Recently, the concept of peptide vaccines has gained considerable attention as a possible means for treating infectious diseases by promoting the patient's individual immune system. Even the study has shown that the whole protein is not necessary for raising the immune response in the immune reaction to the protein antigen development but short division of protein or the epitopes were sufficient for eliciting the preferred immune response.

Development of in silico tools and techniques has intensively mobilized the process of vaccine development in the current era and is frequently implemented to minimize the time required for the identification of the candidate peptide as a vaccine and analyzing the structure and fuctional relationship of virus protein at molecular level.

Hepatitis is a severe viral infection of the liver. Movement towards the techniques which can enhance the research methodology against Hepatitis and computational approaches are the most demanding in the field od science which can help in the development of vaccines against Hepatitis. In the current study, we have theorized that a successful immunization strategy

against Hepatitis infection apart from conventional vaccines, could also involve the concept of peptide vaccines.

The three dimensional structure of various proteins are known with the help of various softwares. We modeled the 3D structures of various Epitopes of the Hepatitis virus, using the established homology modelling techniques in the PEP FOLD and PHYRE2.0 software. The B-cell epitopes were successfully mapped using Antibody Epitope Prediction software, and the predicted epitopes were carefully documented. Furthermore, as our knowledge of the immune responses to a protein antigen has progressed epitope identification. The T-cell epitopes were mapped using NETCTL and Immune Epitope Database and Analysis Resource (IEDB) softwares, and the predicted epitopes were carefully documented. The T-cell epitopes were identified as HTSDHMSIY, SLDVSAAFY, CTCGSSDLY, MSRSESRKN,

CVDVVSQVY. Further, Vaxijen analysis software was used to to predict the segments from a protein sequence that are likely to be antigenic by eliciting an antibody response. Further, the allergenicity of the predicted epitope was identified using Allertop 2.0, AllergenFP v1.0. The predicted epitopes are modelled using the PEPFOLD and PHYRE2.0 softwares.

Design and development of short peptides as vaccine candidate for Hepatitis has gained momentum in the recent years. In conclusion, in the present study we have predicted the epitope like region, responsible for imparting an immune response. The information of immune response to a protein antigen progresses cell epitope recognition and has proved to be a challenging immune-informatics problem in vaccine design. Henceforth, the current study could prove to be useful in designing candidates capable of producing antipeptide antibodies which are component of recognising Hepatitis specific nucleocapsid protein.

Further Recommendations

- Identifying the epitopes for each Hepatitis virus.
- Docking the identified epitopes with T-cell antigen.
- Docking the identified epitopes with B-cell antigen.
- Analyzing the three dimensional structure of the epitope and producing a novel drug, that can help in the eradication of the disease.

CHAPTER 5

REFERENCES

References:

1. https://www.healthline.com/health/hepatitis
2. Ryan M.; Bhutta, Zulfiqar A.; Brown, Alexandria; Carter, Austin; Casey, Daniel C.; Charlson, Fiona J.; Chen, Alan Z.; Coggeshall, Megan; Cornaby, Leslie; Dandona, Lalit; Dicker, Daniel J.; Dilegge, Tina; Erskine, Holly E.; Ferrari, Alize J.; Fitzmaurice, Christina; Fleming, Tom; Forouzanfar, Mohammad H.; Fullman, Nancy; Gething, Peter W.; Goldberg, Ellen M.; Graetz, Nicholas; Haagsma, Juanita A.; Hay, Simon I.; Johnson, Catherine O.; Kassebaum, Nicholas J.; Kawashima, Toana; et al. (October 2016). "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries,

1990–2015: a systematic analysis for the Global Burden of Disease Study 2015". The Lancet. 388 (10053): 1545–1602

3. "Statistics & Surveillance Division of Viral Hepatitis CDC". CDC. Archived from the original on 11 November 2016. Retrieved 10 November 2016.
4. Rutherford, A; Dienstag, JL (2016). "Chapter 40: Viral Hepatitis". In Greenberger, NJ; Blumberg, RS; Burakoff, R (eds.). CURRENT Diagnosis & Treatment: Gastroenterology, Hepatology, & Endoscopy, 3e. New York, NY: McGraw-Hill.
5. Fontana, Robert; Hayashi, Paul (2014-05-01). "Clinical Features, Diagnosis, and Natural History of Drug-Induced Liver Injury". Seminars in Liver Disease.
6. Lee, WM; Dienstag, JL (2015). "Chapter 361: Toxic and Drug-Induced Hepatitis". In Kasper, D; Fauci, A; Hauser, S; Longo, D; Jameson, J; Loscalzo, J (eds.). Harrison's Principles of Internal Medicine 19e. McGraw-Hill.
7. Lee, William M. (31 July 2003). "Drug-Induced Hepatotoxicity". New England Journal of Medicine.
8. Krawitt, Edward-L (2008). "Clinical features and management of autoimmune hepatitis". World Journal of Gastroenterology.
9. Allan, Richard; Thoirs, Kerry; Phillips, Maureen (2010-07- 28). "Accuracy of ultrasound to identify chronic liver disease". World Journal of Gastroenterology.
10. Hepatitis A And Hepatitis B Vaccine (Intramuscular Route)". Mayo Clinic. Retrieved 25 January 2018.
11. Zhang; et al. (March 5, 2015). "Long-Term Efficacy of aHepatitis E Vaccine".

12. Rosen, Hugo R. (2011-06-23). "Clinical practice. Chronic hepatitis C infection". The New England Journal of Medicine.

13. Ryu W (2017). Molecular Virology of Human Pathogenic Viruses. Academic Press. pp. 247–260.

14. Zuckerman AJ (1996). "Chapter 70: Hepatitis Viruses". In Baron S; et al. (eds.). Baron's Medical Microbiology (4th ed.). Univ of Texas Medical Branch.

15. Kay A, Zoulim F (August 2007). "Hepatitis B virus genetic variability and evolution". Virus Research.

16. Bouchard MJ, Schneider RJ (December 2004). "The enigmatic X gene of hepatitis B virus". Journal of Virology.

17. Ryan KJ, Ray CG, eds. (2004). Sherris Medical Microbiology (4th ed.). McGraw Hill.

18. "Hepatitis D | NIDDK". National Institute of Diabetes and Digestive and Kidney Diseases.

19. Melief CJ, van Hall T, Arens R, Ossendorp F, van der Burg SH (September 2015). "Therapeutic cancer vaccines". The Journal of Clinical Investigation.

20. Miller, E.; Beverley, P. C. L.; Salisbury, D. M. (2002-07-01). "Vaccine programmes and policies". British Medical Bulletin.

21. Needham, Joseph. (2000). Science and Civilization in China: Volume 6, Biology and Biological Technology, Part 6, Medicine. Cambridge: Cambridge University Press

22. Chang MH, Chen CJ, Lai MS, Hsu HM, Wu TC, Kong MS, Liang DC, Shau WY, Chen DS (June 1997). "Universal hepatitis B vaccination in Taiwan and the incidence of hepatocellular carcinoma in children. Taiwan Childhood Hepatoma Study Group". The New England Journal of Medicine.

Appendix – I

MATERIALS AND METHODS

Name of software	URL	Description
DATA RETRIVAL		
The National Centre For Biotechnology	https://www.ncbi.nlm.nih.gov/	The NCBI provides access to biomedical and genomic information. Major databases include Genbank and PubMed. NCBI was directed by David Lipman.

Informati on(NCBI)		
EVOLUTIONARY ANALYSIS		
Clustal Omega	https://www.ncbi.nlm.nih.go v/	Clustal is a series of widely used computer programs used in Bioinformatics for multiple sequence alignment. [*Chenna R, et.al (July 2003)]*
BoxShad e	https://embnet.vital-it.ch/software/BOX_form.ht ml	BoxShade is a program for creating visually pleasing images of multiple alignments of protein or DNA sequences. [Kay Hofmann and Michael D. Baron]
Molecula r Evolution ary Genetics Analysis (MEGA)	https://www.megasoftware. net/	Molecular Evolutionary Genetics Analysis (MEGA) is computer software for conducting statistical analysis of molecular evolution and for constructing phylogenetic trees. [Kumar, S. M. Nei (1993)]
PREDICTION OF ANTIGENS AND SUBUNIT VACCINES		
Vaxijen 2.0	: http://www.ddg-pharmfac.net/vaxijen/VaxiJ en/VaxiJen.html	VaxiJen is the first server for alignment-independent prediction of protective antigens. [Doytchinova IA, Flower DR (2007)]
PREDICTION OF T-CELL EPITOPE		

NetCTL 1.2	http://www.cbs.dtu.dk/servi ces/NetCTL/	NetCTLpredicts Cytotoxic T-lymphocyte epitopes in protein sequences. [Larsen MV, et.al]
The Immune Epitope Database (IEDB)	https://www.iedb.org/	This free resource offers easy searching of experimental data characterizing antibody and T cell epitopes studied in humans, non- human primates, and other animal species.
PREDICTION OF B-CELL EPITOPE		
Antibody epitope predictio n	http://tools.immuneepitope. org/bcell/	The following methods are provided for B-cell epitope prediction: ○ Chou&Fasman beta-turn prediction ○ Emini surface accessibility prediction ○ Karplus &schulz flexibility prediction ○ Kolaskar&Tongaonkar antigenicity ○ Parkar hydrophilicity prediction ○ Bepipred linear epitope prediction

BepiPred 1.0	http://www.cbs.dtu.dk/services/BepiPred-1.0/	This derver predicts the location of linear B cell epitope using a combination of hidden Markov model and a propensity scale method (Larsen et al.,2006)
ALLERGENCITY PREDICTION		
Allertop V 2.0	http://www.ddg-pharmfac.net/allertop/	AllerTopis a server for in silico prediction of allergens based on the main physicochemical properties of proteins. [Dimitrov I, et.al]
AllergenFP v 1.0	http://ddg-pharmfac.net/AllergenFP/	It is a bioinformatics tool for allergenicity prediction.
3D MODELLING OF AN EPITOPE		
PEPFOLD	http://bioserv.rpbs.univ- paris-diderot.fr/services/PEP-FOLD/	PEP-FOLD, an online service, aimed at de novo modelling of 3D conformations for peptides between 9 and 25 amino acids in aqueous solution. [Maupetit J, et.al]
PHYRE 2	http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=in dex	**Phyre** and **Phyre2** (**P**rotein **H**om ology/Analog**Y** **R**ecognition **E**ngi ne; pronounced as 'fire') are free web-based services for protein structure prediction. Phyre2 has been designed to ensure a user-

		friendly interface for users inexpert in protein structure predictionmethods. Its development is funded by the Biotechnology and Biological Sciences Research Council. [Lawrence Kelley; Benjamin Jefferys (2011)].