DATE: 22/10/2022

WEBLEM 7
INTRODUCTION TO IEDB DATABASE

AIM:
Introduction to IEDB Database

INTRODUCTION:
The IEDB was established in 2004, and over the past 10 years our team has manually curated almost 16 000 published manuscripts and processed 200 direct submissions. As a result, detailed experimental data regarding more than 120 000 epitopes are now freely and easily accessible to the scientific community via most web browsers as a web-based interface. In addition, if one wishes to view 3D structural data using the Epitope Viewer application, Java 6 or 7 is required. The IEDB's primary curation focus is on data from scientific publications available in PubMed focused on infectious diseases, allergy, autoimmunity and transplantation. Excluded from the primary scope are HIV-derived epitopes captured in the LANLdatabase and cancer epitopes for which there is no resource currently available due to lack of support for such a resource by the National Institutes of Health. As an exception, all publications describing the 3D structure of an epitope in complex with its adaptive immune receptor or major histocompatibility complex(MHC) molecule are included regardless of origin of the epitope in order to provide a complete dataset of this particularly valuable type of information. Details describing the curation process put in place and followed by the curation team, including quality controls for accuracy and consistency, have beendiscussed previously

The IEDB houses epitope-specific experimental assays. That is, every assay reflects the binding of an epitope-specific T cell receptor (TCR), antibody or MHC molecule to an experimentally tested antigen or epitope. The structure entered as the epitope is limited to the exact entity that was actually tested in the assay or was clearly deduced to be the epitope by the authors. In many cases this is not the minimal epitope and may not be limited to the contact residues of the epitope, but is rather a region containing the epitope. The fields of the IEDB describe the details of these experiments in great detail. First, the epitope structure is designated as either peptidic or non-peptidic. Peptidic epitopes are described by their linear amino acid sequence or as discontinuous amino acids by position within their source protein.

WORKING:
IEDB Database icons were designed to highlight the main search components used for epitope related data. Icons were chosen based on a survey of scientists asked to identify the most relevant icon from a set to represent each major search parameter. These icons are similar to ones for hotel, airfare, or car rentalon a travel web site, distinguishing the major types of searches possible. These search sections also serve to restrict search terms to specific database fields and help guide the user as to the types of data that the IEDB contains. For example, in the ‘Host’ section, a variety of hosts including humans, rodents, non- human primates, and an additional nine commonly studied species are presented.

Peptidic epitopes having 3D structural data are described by the residues found to contact the antibody, TCR or MHC molecule. Non-peptidic epitopes are manually curated by staff from the ChEBI team (3) who annotate the complete molecular structures using SMILES annotation. If the epitope was derived froma protein or a larger non-peptidic structure, these are also provided along with the organism in which these structures are found.

Fig 1: The IEDB 3.0 home page has the most commonly used search parameters centered on the page, shown in box (A), with the highly used analysis tools made moreprominent, shown in box (B).

Once a query such as the one populated in Figure 1 has been executed, the search results are presented on a new page with the current search filters displayed at the top of the results table (Figure 1, box A). Any filter can be removed by a single click on the 'X' next to each parameter. The amount of data present withineach of the result set types of Epitopes, Antigens, Assays and References are conveyed by counts and displayed as tabs that allow the user to easily navigate between them (Figure 1, box B). As shown in Figure 1 box C, a search panel added to the left side of the page allows the current result set to be further refined by adding search parameters or to run a new query entirely. These search panels contain the functionality present on the home page plus several additional search features, some of which were previously only present in the IEDB 2.0 'Advanced Search', such as the 'Assay Types.' We plan to continuously monitor the usage of each search parameter to identify additional fields that should be addedto or removed from the search panel on the results page.

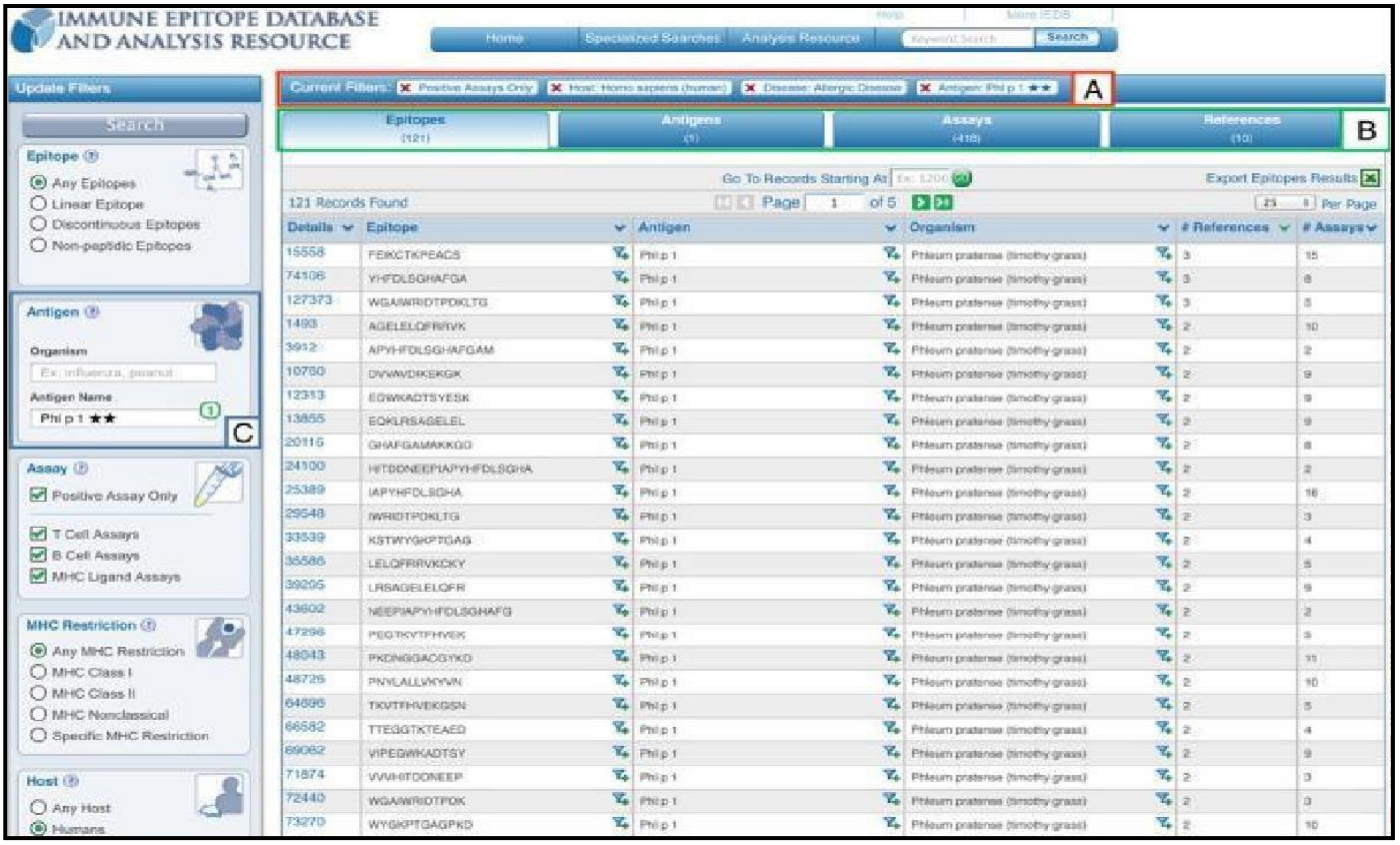


Fig 2: New results presentation format shows current search filters in box (A), counts returned per data type in box (B) and the new left search panel allowing for continuedrefinement or editing of one's query, such as by the epitope source, in box (C).

In addition to the query interface, the presentation of the results has been modified as well. Query results are grouped in four tabs: Epitopes, Antigens, Assays and References that match the current search criteria (Figure2, box B). These different units of information reflect that some users want to utilize the IEDB as, for example, a way to explore the literature (on the reference tab), while others want to see which specific proteins in an organism have been studied for immune reactivity (on the antigen tab). The amount of data hosted in the IEDB has grown dramatically in the last few years, so that typical queries retrieve a very large number of epitopes. To make sure the most relevant epitopes are immediately visible, results are now sorted by how much information is available, such as the number of references with relevant data, as shown in Figure 2, rather than alphabetically, as was previously done. In addition to the left search panel, users can click on an epitope structure or its source to further narrow the result, using a new 'filter' icon present in the results table. Another noteworthy enhancement in the IEDB 3.0 is a new 'Antigen' tab which displays all epitopes that belong to the same antigen in one row

The Molecule Finder has two top-level branches for peptidic and non-peptidic epitopes. Non- peptidic epitopes are assigned to sources in ChEBIand displayed using the ChEBI hierarchy. Peptidic epitopes derived from proteins occurring in nature have their specific source protein identified by GenPept (5) entries. The variety of distinct sequences represented in GenPept (e.g. the five versions of Phl p I is necessary and reflective of the heterogeneity of proteins within individual species; however, the large number of entries and lack of standardized nomenclature previously overwhelmed users, and made it difficult to obtain all epitopes belonging to a single antigen.

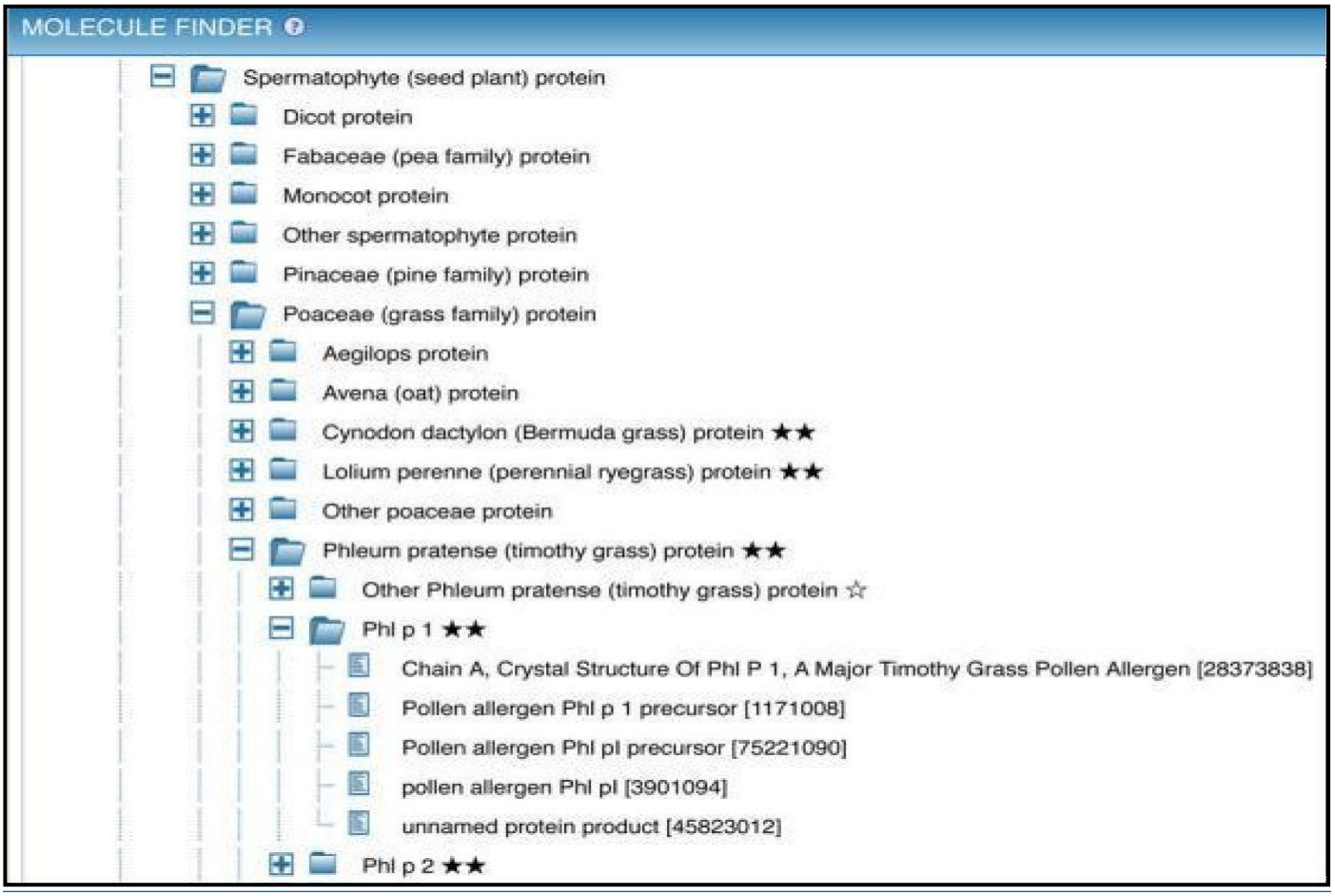


Fig 3: The Molecule Finder provides a hierarchical organization of proteins that allowsnarrowing the search to epitopes derived from a specific antigen, such as the common allergen Phl p 1. The reference proteome protein 'Phl p 1' is the parent of five individual GenPept entries for this protein from Timothy grass.

DISCOTOPE SERVER 1.1

DiscoTope is the first method to focus explicitly on discontinuous epitopes. We show that the new structure-based method has a better performance for predicting residues of discontinuous epitopes than methods based solely on sequence information, and that it can successfully predict epitope residues that have been identified by different techniques. DiscoTope detects 15.5% of residues located in discontinuous epitopes with a specificity of 95%. At this level of specificity, the conventional Parker hydrophilicity scale for predicting linear B-cell epitopes identifies only 11.0% of residues located in discontinuous epitopes. Predictions by the DiscoTope method can guide experimental epitope mapping in both rational vaccine design and development of diagnostic tools, and may lead to more efficient epitope identification.

Discontinuous epitopes, B-cell epitope, antibody, vaccine design, protein structure, antigen, accessibility, hydrophilicity.

Table view lists following columns:

Chain ID: The chain id of the protein chain used in prediction (specified by the user)

Residue ID: PDB Residue id

Residue Name: Name of the residue

Contact Number: The residue contact number is the number of Cα atoms in the antigen within a distance of 10 Å of the residue's Cα atom. A low contact number correlates with localization of the residue close tothe surface or in protruding regions of the antigen's structures.

Propensity Score: This score tells you about the probability/tendency of being part of an epitope for that particular residue. The propensity is reflected in amino acid epitope log-odds ratios, which were calculated on a set of 75 antigens. The propensity score is calculated by sequentially averaging epitope log-odds ratios within a window of 9 residues. Then the scores are summed up based on the proximity in the 3D structure of the antigen. For any given residue, the sequentially averaged log-odds scores from all residues within 10Å are summed to give the propensity score.

Discotope Score: This score is calculated by combining the contact numbers with propensity score. DiscoTope score above the threshold value indicates positive predictions and that below the threshold value indicates negative predictions.

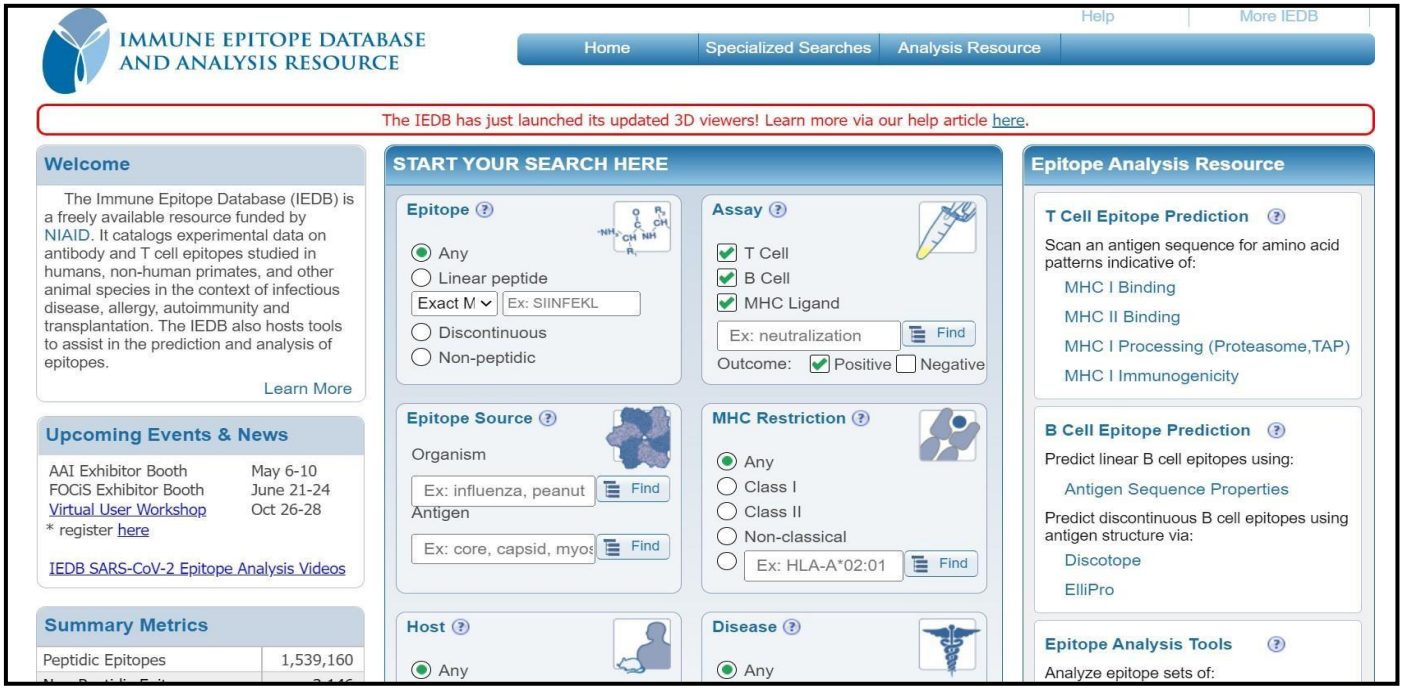


Fig 4: Homepage of IEDB Database

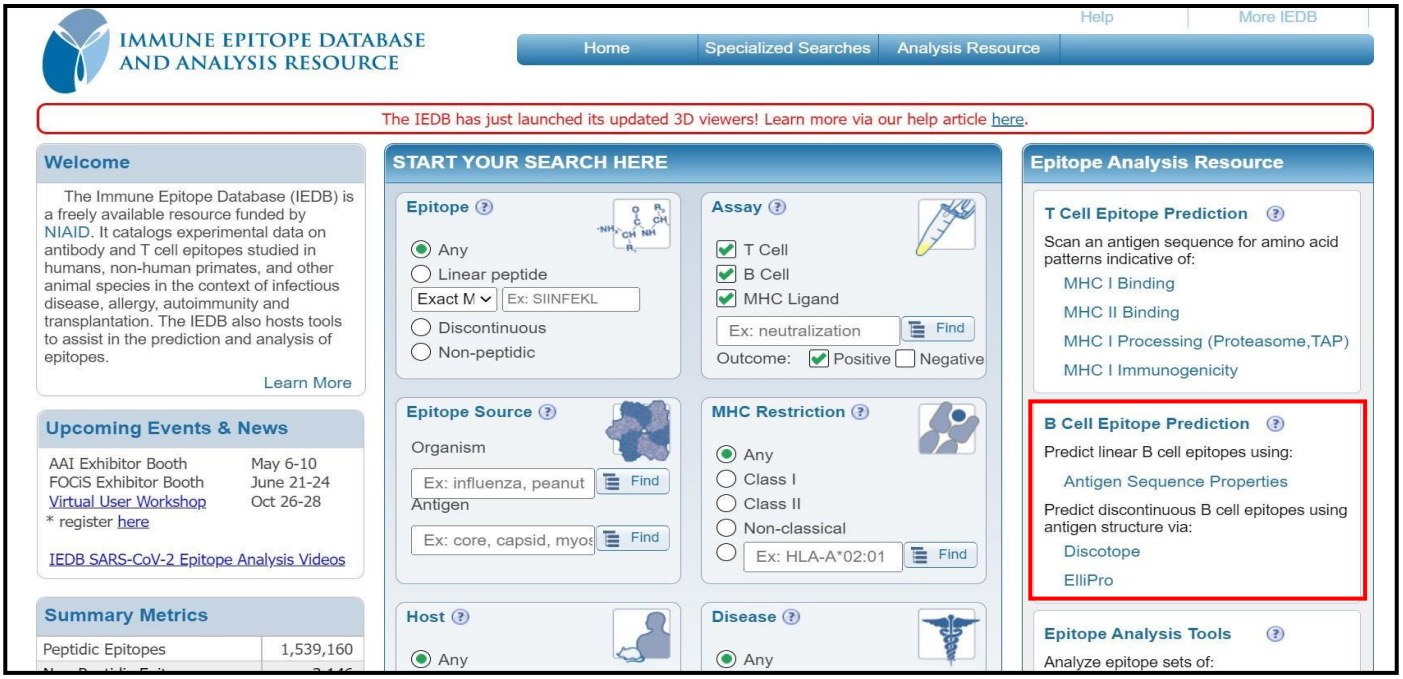


Fig 5: B Cell Epitope Prediction section: Discotope server option

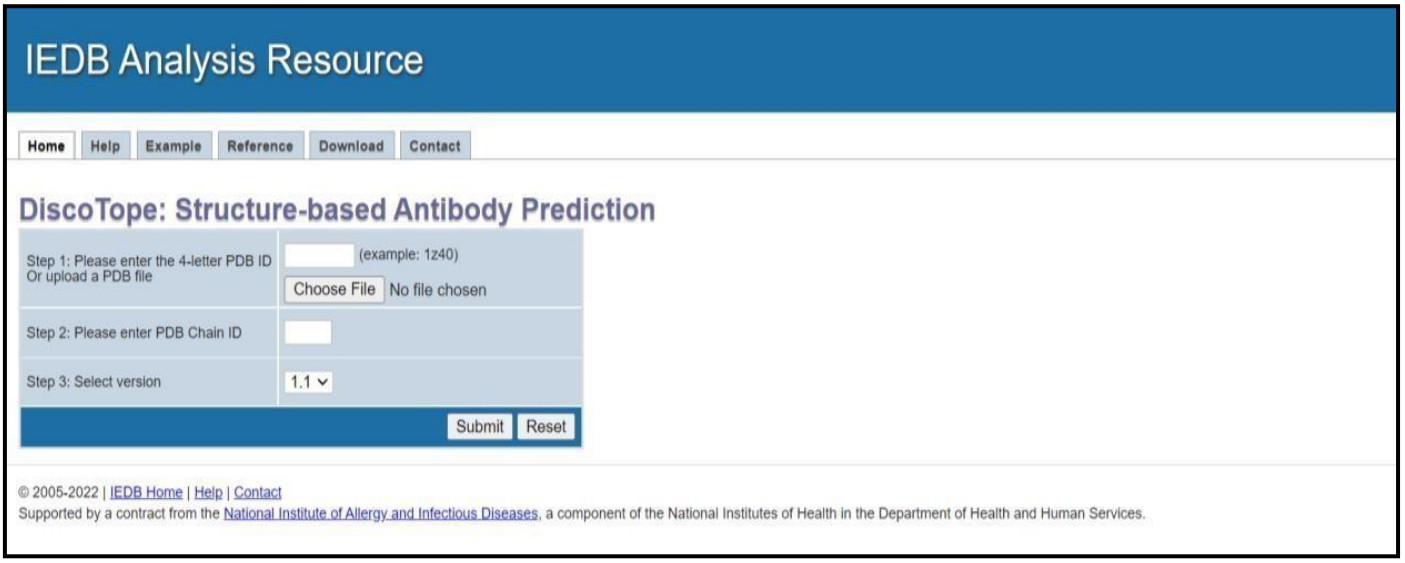


Fig 6: Homepage of DiscoTope

TEPITOOL:

Computational prediction of T-cell epitope candidates is currently being used in several applications including vaccine discovery studies, development of diagnostics and removal of unwanted immune responses against protein therapeutics. There have been continuous improvements on the performance of MHC binding prediction tools but their general adoption by immunologists has been slow due to the lack of user-friendly interfaces and guidelines. Current tools only provide minimal advice on what alleles to include, what lengths to consider, how to deal with homologous peptides and what cutoffs should be considered relevant. This protocol provides step-by-step instructions with necessary recommendations for prediction of the best T-cell epitope candidates in line with the newly developed online tool called TepiTool. The TepiTool, part of IEDB, provides some of the top MHC binding prediction algorithms for number of species including humans, chimpanzees, bovines, gorillas, macaques, mice and pigs. The TepiTool is freely accessible at http://tools.iedb.org/tepitool/.

The binding of a peptide to an MHC molecule is necessary for its ability to activate T cell responses. Peptides bind MHC molecules in the "peptide binding groove", forming a peptide- MHC complex which in turn is recognized by the T cell receptors. Peptides recognized by T cells are called epitopes. Epitopes bound to class I and class II MHC molecules are recognized by $CD8^+$ and $CD4^+$ T cells, respectively.

Generally, MHC binding prediction tools scan amino acid sequences to estimate the binding affinity of each component peptide to a specific MHC. MHC class I molecules have a binding groove that is closedat its ends, limiting the size of its ligands to roughly 8-11 residues in length. Class II molecules, on the other hand, have an open binding groove, allowing them to bind longer peptides, typically 12-20 residues in length. The strength of binding (affinity) of a peptide to an MHC molecule is an important factor that determines potential immunogenicity.

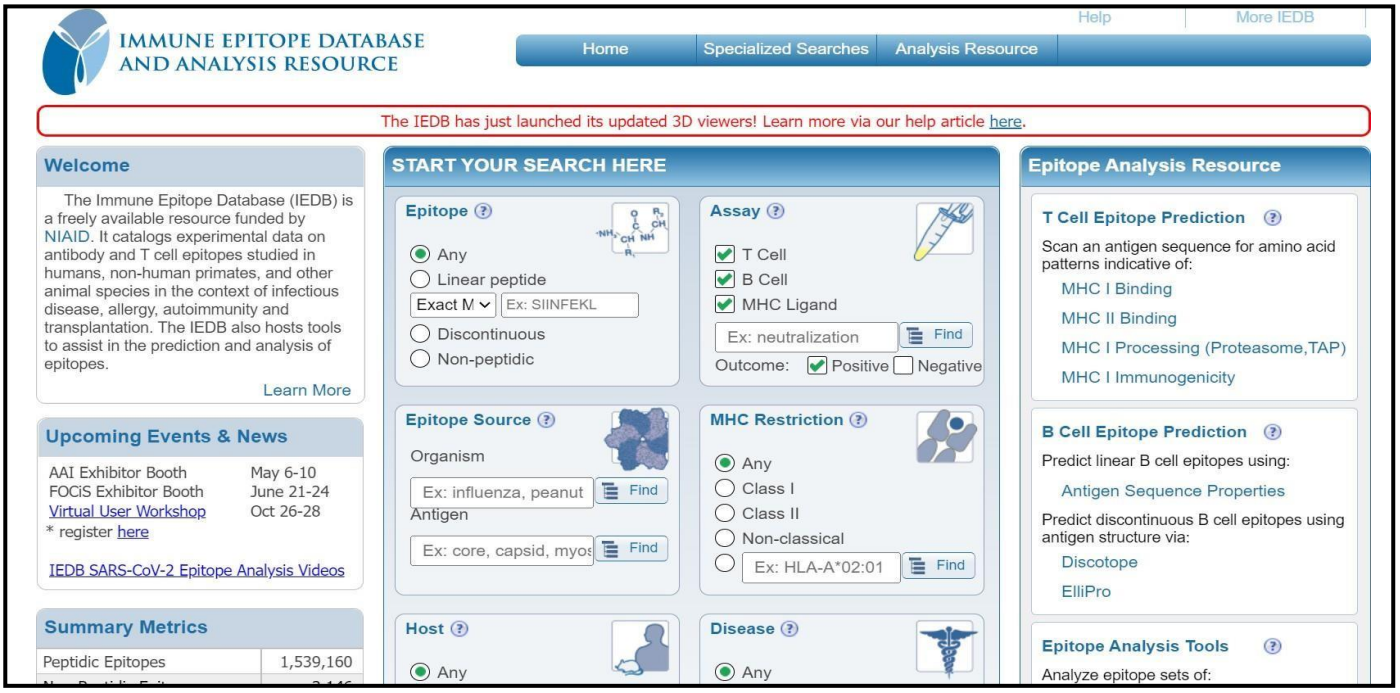


Fig 7: Homepage of IEDB database

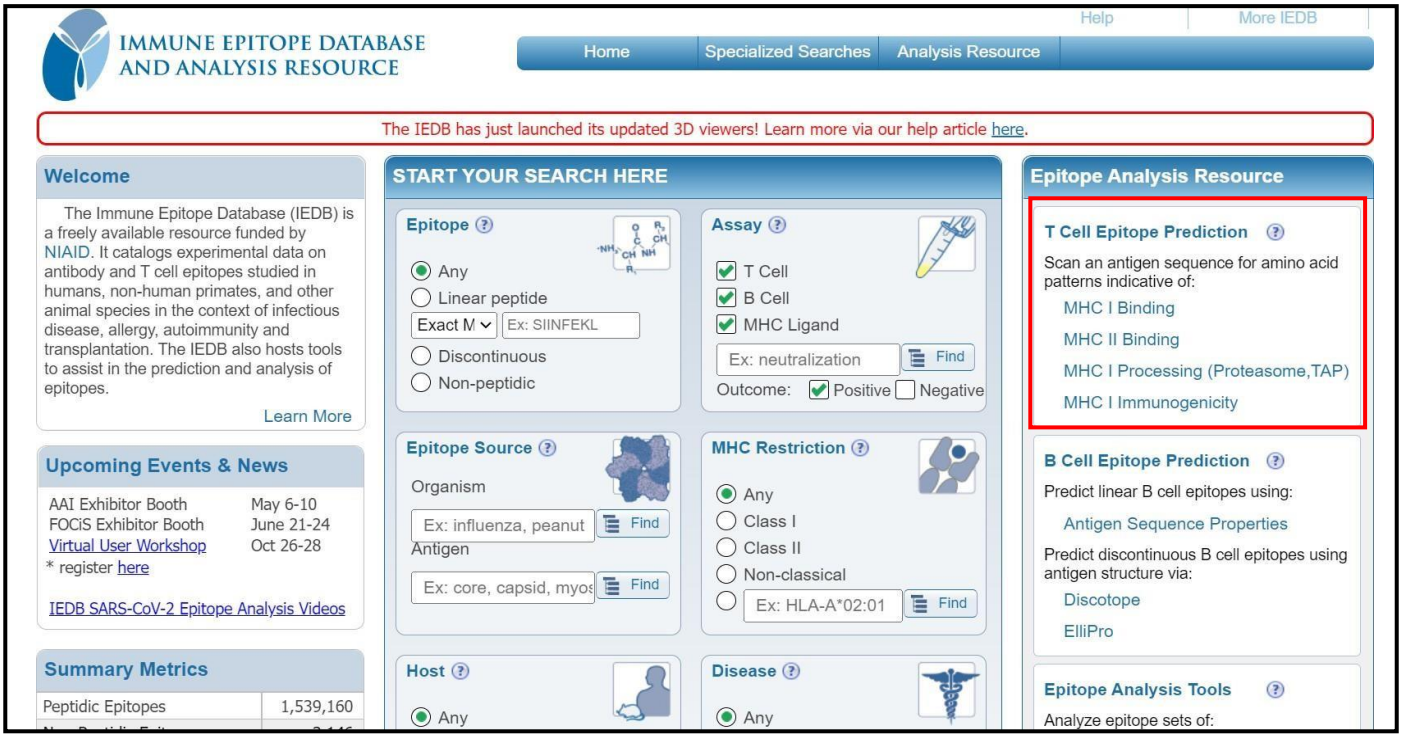


Fig 8: TCE Prediction under IEDB Database

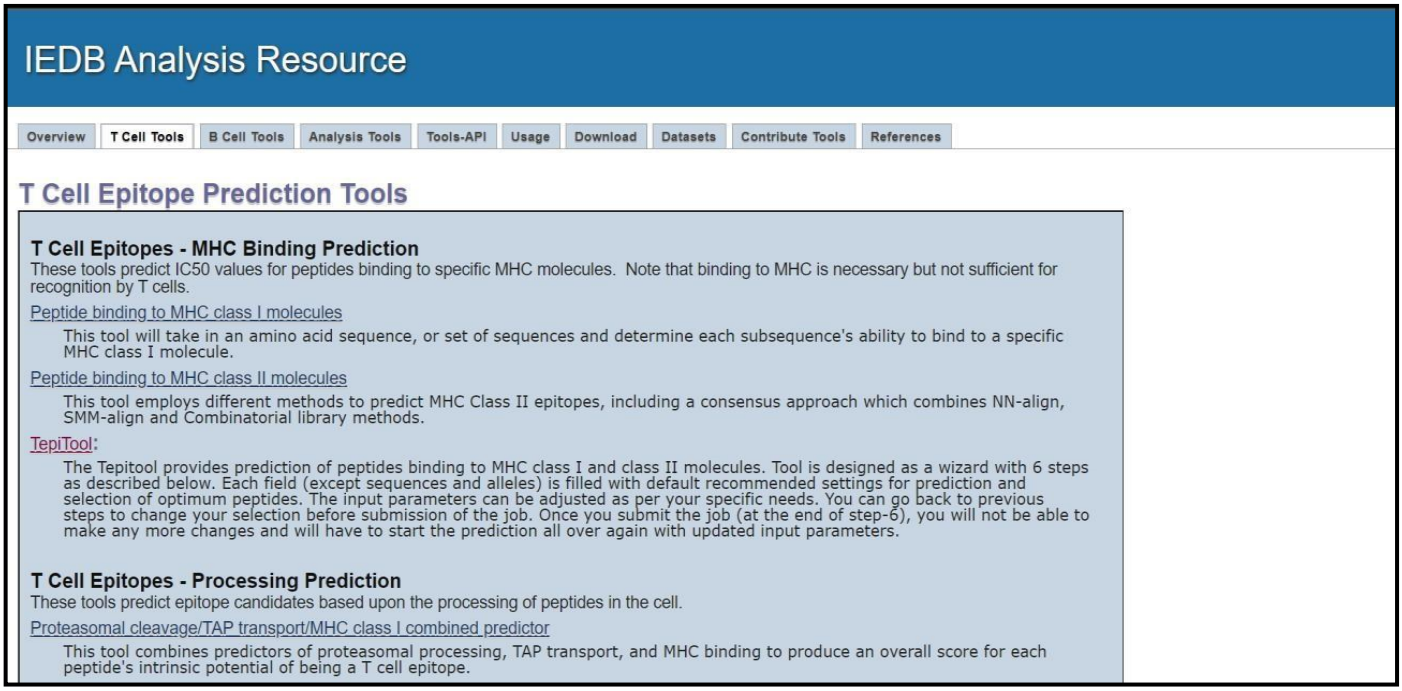


Fig 9: Homepage of TCE Prediction

CONCLUSION:

After catching up on the curation of in-scope journal articles from the past, the focus of IEDB development for the 3.0 release has shifted toward improving query and reporting interfaces. The goal of this release was to provide intuitive ways to extract biologically accurate information from the large amounts of data now stored in the IEDB. We have here described the main new elements of the 3.0 release, all of which were motivated by user feedback gathered over the years. We believe that such development focusing on the usability of the web site is equally important to the introduction of new capabilities which—while often more exciting to implement from a web site developer's perspective— have little value if they are not actually utilized by the user community.

REFERENCES:

1. IEDB Analysis Resource. (n.d.). NCBI. Retrieved October 10, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4384014/
2. Tepitool. (n.d.). NCBI. Retrieved October 10, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4981331/

DATE: 22/10/2022

WEBLEM 7A
DISCOTOPE SERVER 1.1

(URL: https://www.iedb.org/)

AIM: To predict B Cell epitope for 2Z91 using Discotope server 1.1.

INTRODUCTION:

DiscoTope is the first method to focus explicitly on discontinuous epitopes. We show that the new structure-based method has a better performance for predicting residues of discontinuous epitopes than methods based solely on sequence information, and that it can successfully predict epitope residues that have been identified by different techniques. DiscoTope detects 15.5% of residues located in discontinuous epitopes with a specificity of 95%. At this level of specificity, the conventional Parker hydrophilicity scale for predicting linear B-cell epitopes identifies only 11.0% of residues located in discontinuous epitopes. Predictions by the DiscoTope method can guide experimental epitope mapping in both rational vaccine design and development of diagnostic tools, and may lead to more efficient epitope identification.
Discontinuous epitopes, B-cell epitope, antibody, vaccine design, protein structure, antigen, accessibility, hydrophilicity.

Table view lists following columns:
Chain ID: The chain id of the protein chain used in prediction (specified by the user)
Residue ID:PDB Residue id
Residue Name: Name of the residue
Contact Number: The residue contact number is the number of Cα atoms in the antigen within a distanceof 10 Å of the residue's Cα atom. A low contact number correlates with localization of the residue close to the surface or in protruding regions of the antigen's structures.
Propensity Score: This score tells you about the probability/tendency of being part of an epitope for that particular residue. The propensity is reflected in amino acid epitope log-odds ratios, which were calculated on a set of 75 antigens. The propensity score is calculated by sequentially averaging epitope log-odds ratios within a window of 9 residues. Then the scores are summed up based on the proximity inthe 3D structure of the antigen. For any given residue, the sequentially averaged log-odds scores from allresidues within 10Å are summed to give the propensity score.

Discotope Score: This score is calculated by combining the contact numbers with propensity score.
DiscoTope score above the threshold value indicates positive predictions and that below the threshold value indicates negative predictions.

Immunoglobulins (Ig) or antibodies are glycoproteins that are produced by plasma cells. B cells are instructed by specific immunogens, for, example, bacterial proteins, to differentiate into plasma cells, which are protein-making cells that participate in humoral immune responses against bacteria, viruses, fungi, parasites, cellular antigens, chemicals, and synthetic substances. The immunogen or antigen reacts with a B-cell receptor (BCR) on the cell surface of B lymphocytes, and a signal is produced that directs the activation of transcription factors to stimulate the synthesis of antibodies, which are highly specific for the immunogen that stimulated the B cell. Furthermore, one clone of B cell makes an immunoglobulin (specificity).

Using the crystal structures of anti-ciguatoxin antibody 10C9 Fab in ligand-free form and in complexes with ABCD-ring (CTX3C-ABCD) and ABCDE-ring (CTX3C-ABCDE) fragments of the antigen CTX3Cat resolutions of 2.6, 2.4, and 2.3 angstroms, respectively, we elucidated the mechanism of the interaction between the polycyclic ethers and the antibody. 10C9 Fab has an extraordinarily large and deep binding pocket at the center of the variable region, where CTX3C-ABCD or CTX3C-ABCDE binds longitudinally in the pocket via hydrogen bonds and van der Waals interactions. Upon antigen-antibody complexation, 10C9 Fab adjusts to the antigen fragments by means of rotational motion in the variable region. In addition, the antigen fragment lacking the E-ring induces a large motion in the constant region.

METHODOLOGY:

1. Open the Homepage of IEDB database from google: https://www.iedb.org/
2. On the IEDB Database homepage, under the section B cell Epitope Predictio, Click on Discotope server option.
3. In the Discotope search panel enter the PDB Id 2Z91 along with the chain ID ‘A’.
4. Submit the query and interpret the results.

OBSERVATIONS:

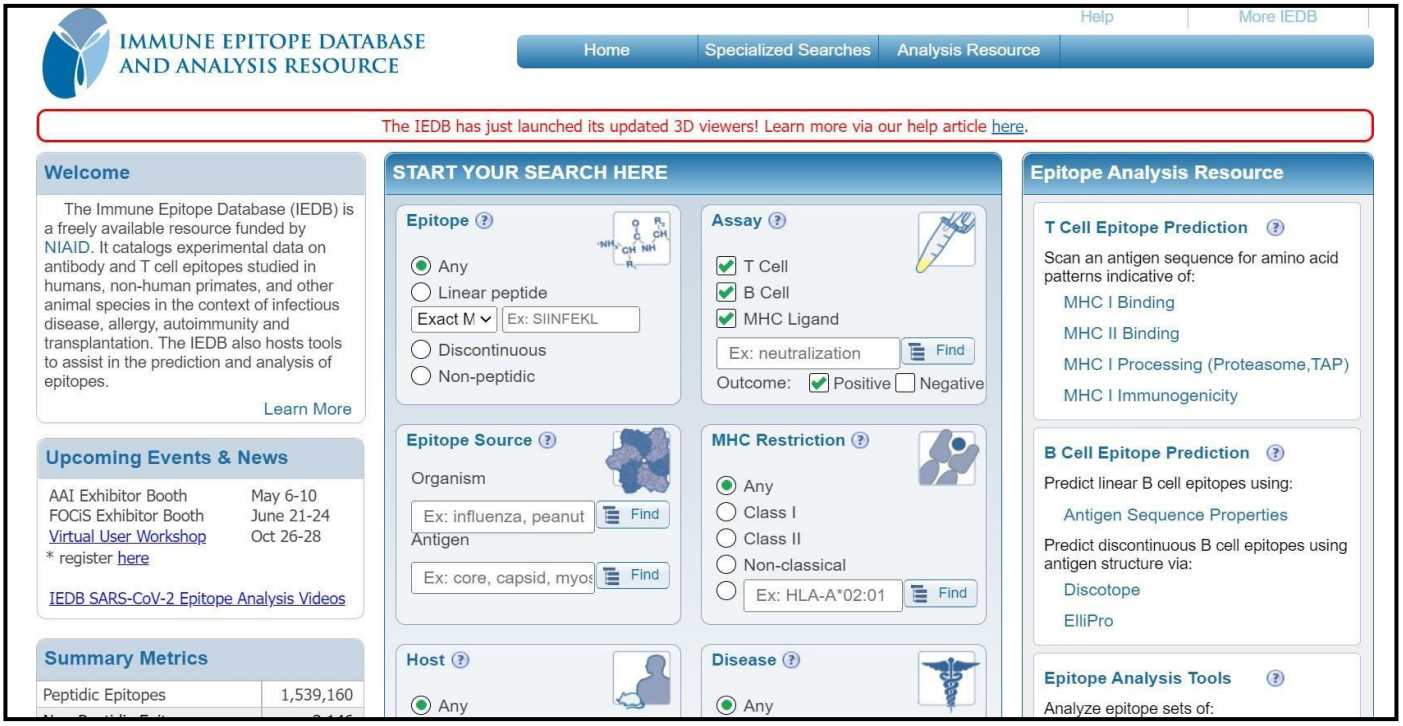


Fig 1: Homepage of IEDB Database

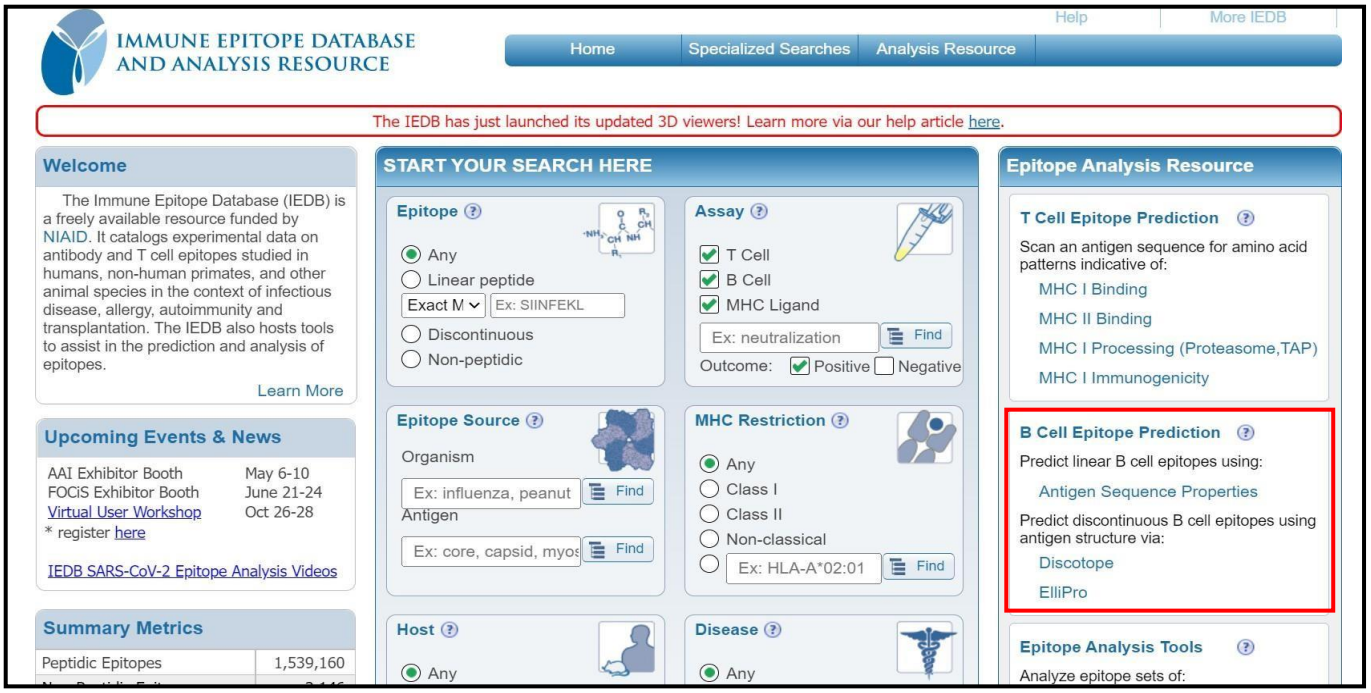


Fig 2: B Cell Epitope Prediction section: Discotope server option

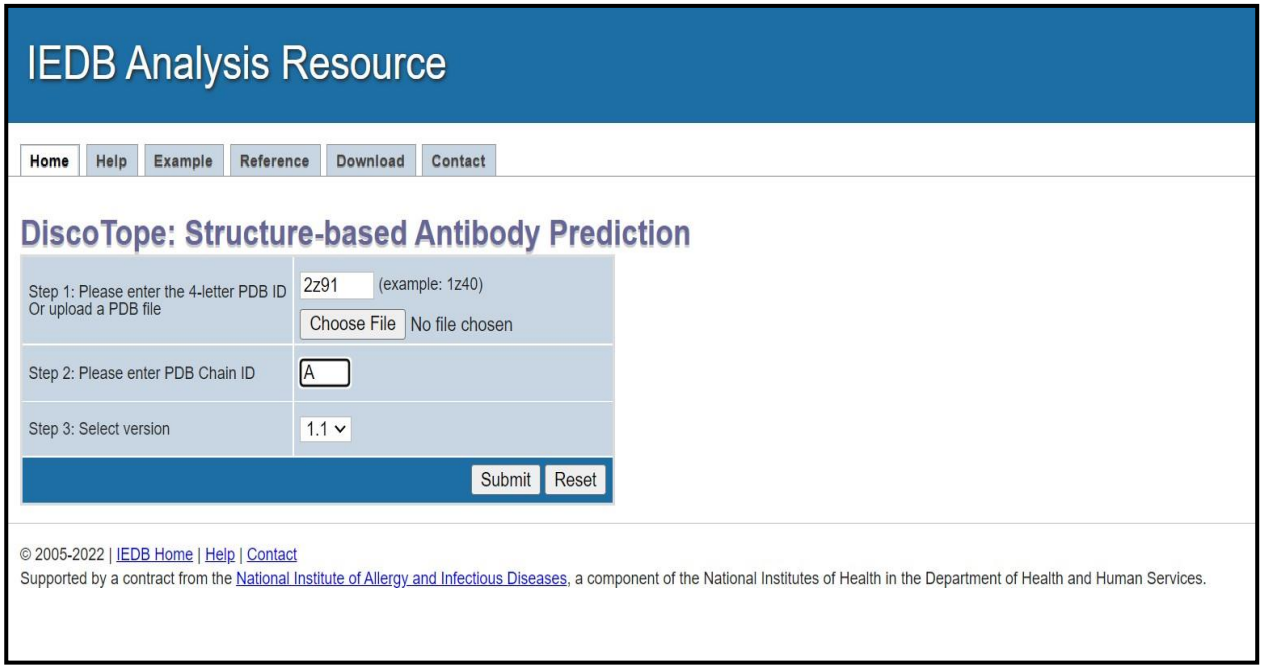


Fig 3: Search for query (2Z91)in DiscoTope server

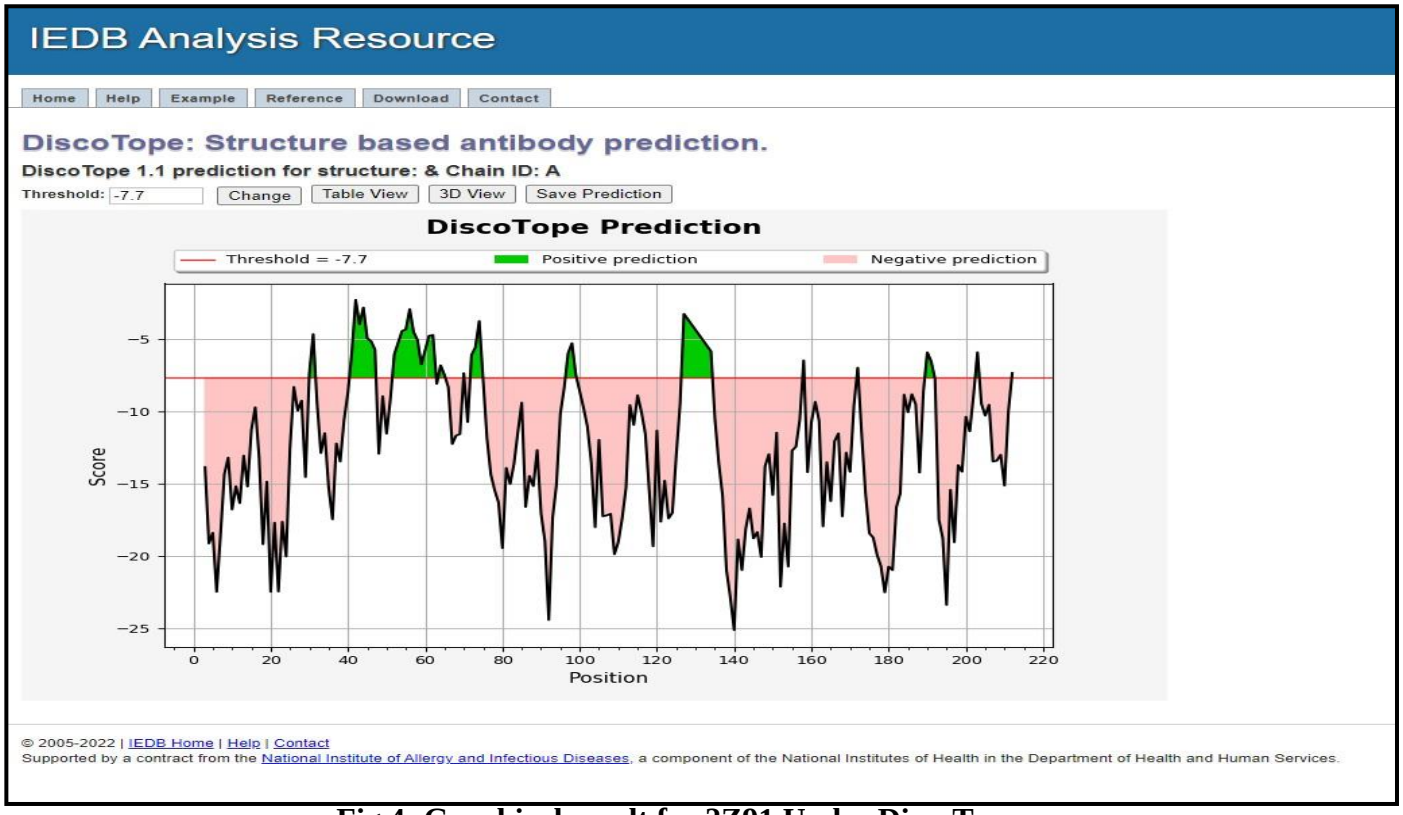
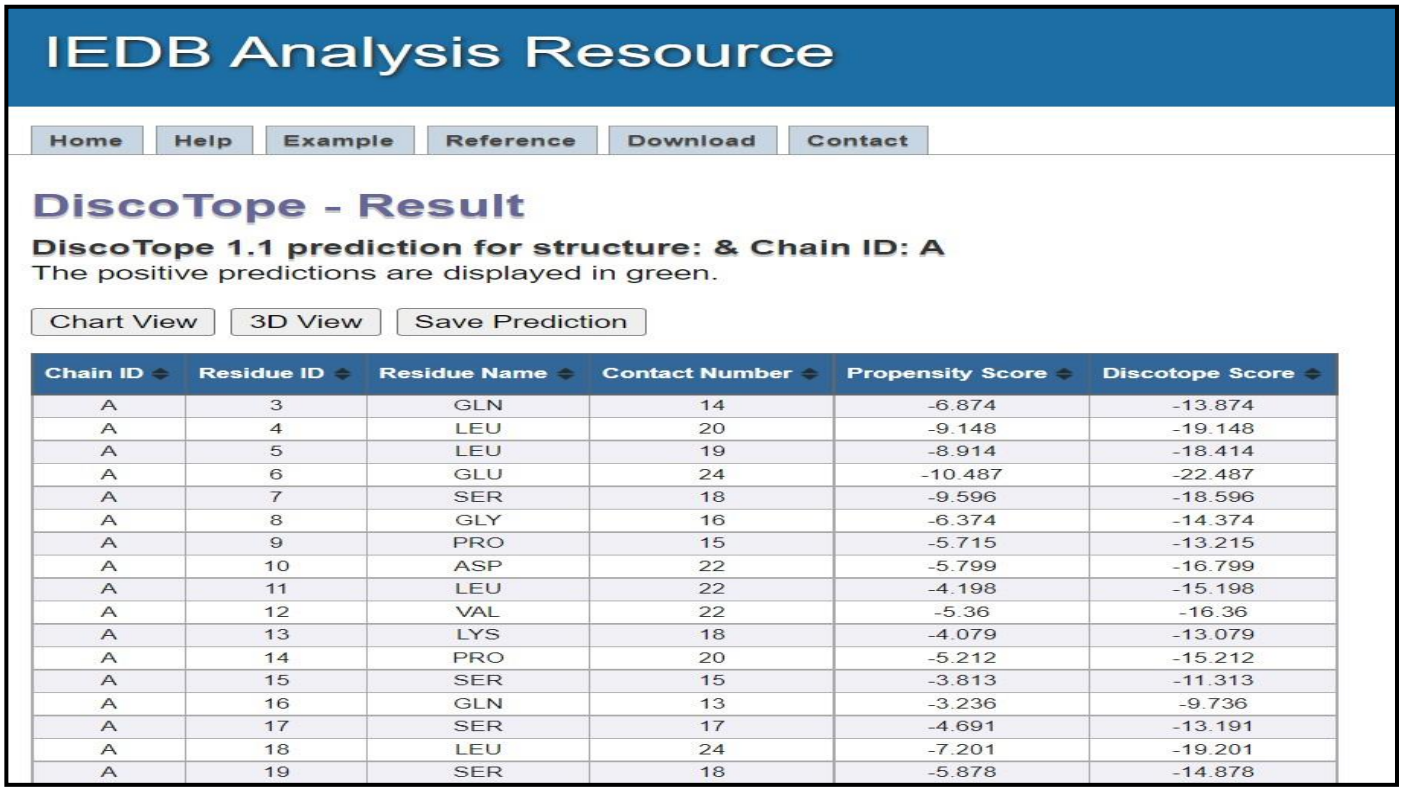


Fig 4: Graphical result for 2Z91 Under DiscoTope server



Chain ID	Residue ID	Residue Name	Contact Number	Propensity Score	Discotope Score
A	3	GLN	14	-6.874	-13.874
A	4	LEU	20	-9.148	-19.148
A	5	LEU	19	-8.914	-18.414
A	6	GLU	24	-10.487	-22.487
A	7	SER	18	-9.596	-18.596
A	8	GLY	16	-6.374	-14.374
A	9	PRO	15	-5.715	-13.215
A	10	ASP	22	-5.799	-16.799
A	11	LEU	22	-4.198	-15.198
A	12	VAL	22	-5.36	-16.36
A	13	LYS	18	-4.079	-13.079
A	14	PRO	20	-5.212	-15.212
A	15	SER	15	-3.813	-11.313
A	16	GLN	13	-3.236	-9.736
A	17	SER	17	-4.691	-13.191
A	18	LEU	24	-7.201	-19.201
A	19	SER	18	-5.878	-14.878

Fig 5: Table view for 2Z91 Under DiscoTope server

A	40	PHE	16	-0.711	-8.711
A	**41**	**PRO**	**13**	**0.523**	**-5.977**
A	**42**	**GLY**	**8**	**1.685**	**-2.315**
A	**43**	**ASN**	**11**	**1.496**	**-4.004**
A	**44**	**LYS**	**10**	**2.155**	**-2.845**
A	**45**	**LEU**	**14**	**2.034**	**-4.966**
A	**46**	**GLU**	**14**	**1.822**	**-5.178**
A	**47**	**TRP**	**15**	**1.755**	**-5.745**
A	48	MET	25	-0.449	-12.949
A	49	GLY	23	2.525	-8.975
A	50	TYR	26	1.441	-11.559
A	51	ILE	25	3.34	-9.16
A	**52**	**HIS**	**18**	**2.909**	**-6.091**
A	**53**	**TYR**	**17**	**3.217**	**-5.283**
A	**54**	**ARG**	**15**	**3.031**	**-4.469**
A	**55**	**GLY**	**16**	**3.645**	**-4.355**
A	**56**	**THR**	**12**	**3.044**	**-2.956**
A	**57**	**THR**	**17**	**3.971**	**-4.529**
A	**58**	**ASN**	**17**	**3.415**	**-5.085**
A	**59**	**TYR**	**19**	**2.727**	**-6.773**
A	**60**	**ASN**	**14**	**1.2**	**-5.8**
A	**61**	**THR**	**11**	**0.703**	**-4.797**
A	**62**	**SER**	**10**	**0.258**	**-4.742**

Fig 5.1: Table view for 2Z91 position prediction (Green Colour) Under DiscoTope server

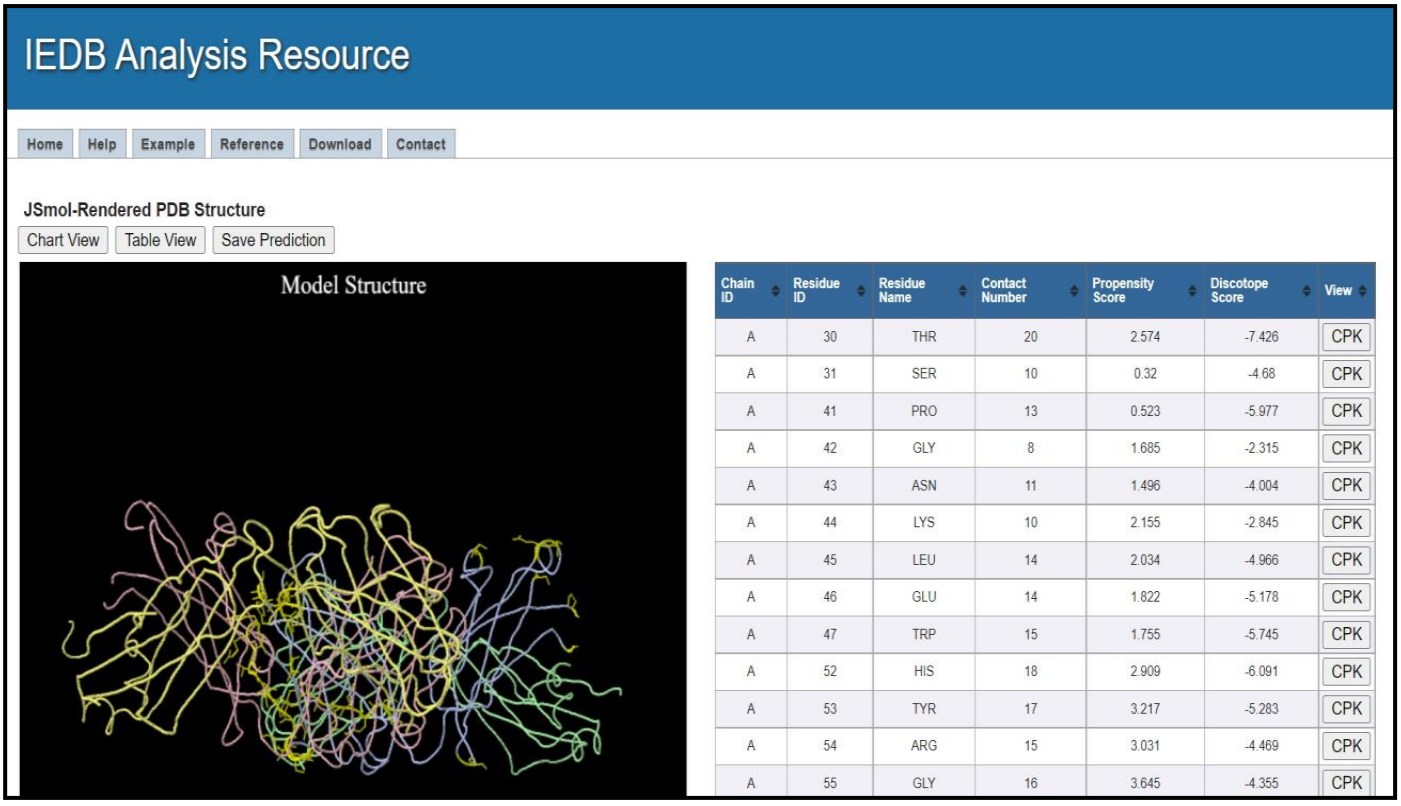
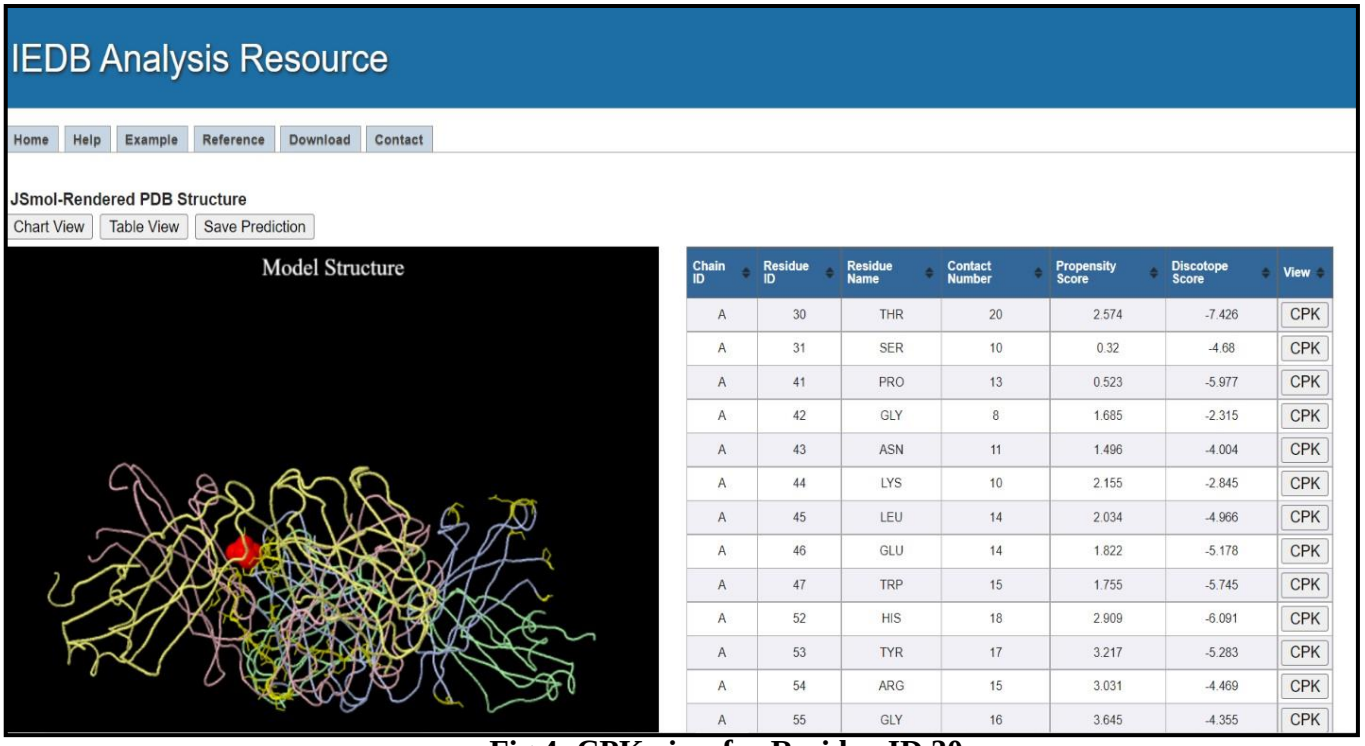


Fig 6: 3D view for 2Z91 Under DiscoTope server



Chain ID	Residue ID	Residue Name	Contact Number	Propensity Score	Discotope Score	View
A	30	THR	20	2.574	-7.426	CPK
A	31	SER	10	0.32	-4.68	CPK
A	41	PRO	13	0.523	-5.977	CPK
A	42	GLY	8	1.685	-2.315	CPK
A	43	ASN	11	1.496	-4.004	CPK
A	44	LYS	10	2.155	-2.845	CPK
A	45	LEU	14	2.034	-4.966	CPK
A	46	GLU	14	1.822	-5.178	CPK
A	47	TRP	15	1.755	-5.745	CPK
A	52	HIS	18	2.909	-6.091	CPK
A	53	TYR	17	3.217	-5.283	CPK
A	54	ARG	15	3.031	-4.469	CPK
A	55	GLY	16	3.645	-4.355	CPK

Fig 4: CPK view for Residue ID 30

RESULTS:

The results are divided into 3 sections: Graphical View/Chart view, Table view & 3D view. Details are discussed below:

1. **Graphical View/Chart view:**

In this, the default threshold value for version 1.1 is -7.7 based on which it states that the predictions above the threshold (red line) are positive predictions (displayed in green) and predictions below the threshold are negative predictions (displayed in orange) indicating which amino acids from 3D structure should be considered as epitope.

2. **Table View:**

The Table view lists the following columns such as:
Chain ID: The chain id of the protein chain used in prediction (specified by the user)
Residue ID:PDB Residue id
Residue Name: Name of the residue
Contact Number: The residue contact number is the number of Cα atoms in the antigen within a distance of 10 Å of the residue's Cα atom. A low contact number correlates with localization of the residue close to the surface or in protruding regions of the antigen's structures.
Propensity Score: This score tells you about the probability/tendency of being part of an epitope for that particular residue. The propensity is reflected in amino acid epitope log- odds ratios, which were calculated on a set of 75 antigens. The propensity score is calculated by sequentially averaging epitope log-odds ratios within a window of 9 residues. Then the scores are summed up based on the proximity in the 3D structure of the antigen.

For any given residue, the sequentially averaged log-odds scores from all residues within 10Å are summed to give the propensity score.

- **Discotope Score:**This score is calculated by combining the contact numbers with propensity score. DiscoTope score above the threshold value indicates positive predictions and that below the threshold value indicates negative predictions.
- Whereas, the positive predictions are displayed in green.

3. 3D view:

The 3d view uses Jmol to display the structure with positive predictions highlighted in yellow. The side chain of each predicted residue is shown. You can rotate, zoom and manipulate the structure by using different buttons on the mouse. The table lists the predicted epitope residues along with their chain id, residue id, contact number, propensity score and DiscoTope score. Clicking on the CPK button in each residue will highlight this residue in CPK on the 3D viewer.

CONCLUSION:

The Discotope server helps to predict discontinuous epitopes from 3D structures of proteins in PDB format. These methods achieve highly significant predictive performances suggesting these tools to be a powerful asset in rational epitope discovery.

REFERENCES:

1. BCE Prediction Discotope. (n.d.). NCBI. Retrieved October 22, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2242418/
2. Immunoglobulin. (n.d.). NCBI. Retrieved October 22, 2022, fromhttps://www.ncbi.nlm.nih.gov/books/NBK513460/
3. IEDB Analysis Resource. (n.d.). Retrieved October 22, 2022, from https://www.iedb.org/
4. BCE Prediction Discotope. (n.d.). IEDB Analysis Resource. Retrieved October 22, 2022, from http://tools.iedb.org/discotope/
5. 2Z91. (n.d.). IEDB Analysis Resource. Retrieved October 22, 2022, fromhttp://tools.iedb.org/discotope/result/

DATE: 22/10/2022

WEBLEM 7B
TEPITOOL

(URL: http://tools.iedb.org/tepitool/)

AIM:

To predict peptides for LILRA5 binding to MHC class I and class II molecules using Tepitool.

INTRODUCTION:

The Tepitool provides prediction of peptides binding to MHC class I and class II molecules. Tool is designed as a wizard with 6 steps as described below. Each field (except sequences and alleles) is filled with default recommended settings for prediction and selection of optimum peptides. The input parameters can be adjusted as per your specific needs. You can go back to previous steps to change your selection before submission of the job. Once you submit the job (at the end of step-6), you will not be ableto make any more changes and will have to start the prediction all over again with updated input parameters.

Immunoglobulins (Ig) or antibodies are glycoproteins that are produced by plasma cells. B cells are instructed by specific immunogens, for, example, bacterial proteins, to differentiate into plasma cells, which are protein-making cells that participate in humoral immune responses against bacteria, viruses, fungi, parasites, cellular antigens, chemicals, and synthetic substances.

Leukocyte immunoglobulin-like receptor subfamily A member 5 (LILR-A5) also known as CD85 antigen-like family member F (CD85f), immunoglobulin-like transcript 7 (ILT-7), and leukocyte immunoglobulin-like receptor 9 (LIR-9) is a protein that in humans is encoded by the LILRA5 gene. This gene is one of the leukocyte receptor genes that form a gene cluster on the chromosomal region 19q13.4

METHODOLOGY:

1. Copy the FASTA sequence for query LILRA5 from the Uniport database. (https://www.uniprot.org/)
2. Open the Homepage of IEDB database from google: https://www.iedb.org/
3. In the IEDB Database homepage, Click on T cell Epitope Prediction.
4. Under the T cell Epitope Prediction section, select the Tepitool option.
5. Follow the submission Steps:

STEP 1: SEQUENCE - Provide sequence data:
Paste the FASTA sequence of query retrieved from Uniprot database

STEP2: Select SPECIES & ALLELE CLASS:
Select the host species and MHC allele class.

STEP 3: Select ALLELES i.e. Specify alleles:
Select the specific alleles for prediction.(NOTE: For multiple alleles selection, hold CTRL button and select)

STEP 4: PEPTIDES - Select peptides to be included in prediction:
Select peptides to be included in prediction as per the option provided by the tool.

STEP 5: METHOD - Select prediction & peptide selection methods and cutoff values: Select preferred methods for binding prediction and peptide selection strategy and cutoff values.

STEP 6: REVIEW: Review selections, enter job details & submit data:
Review selections, enter job details and submit data.

OBSERVATIONS:

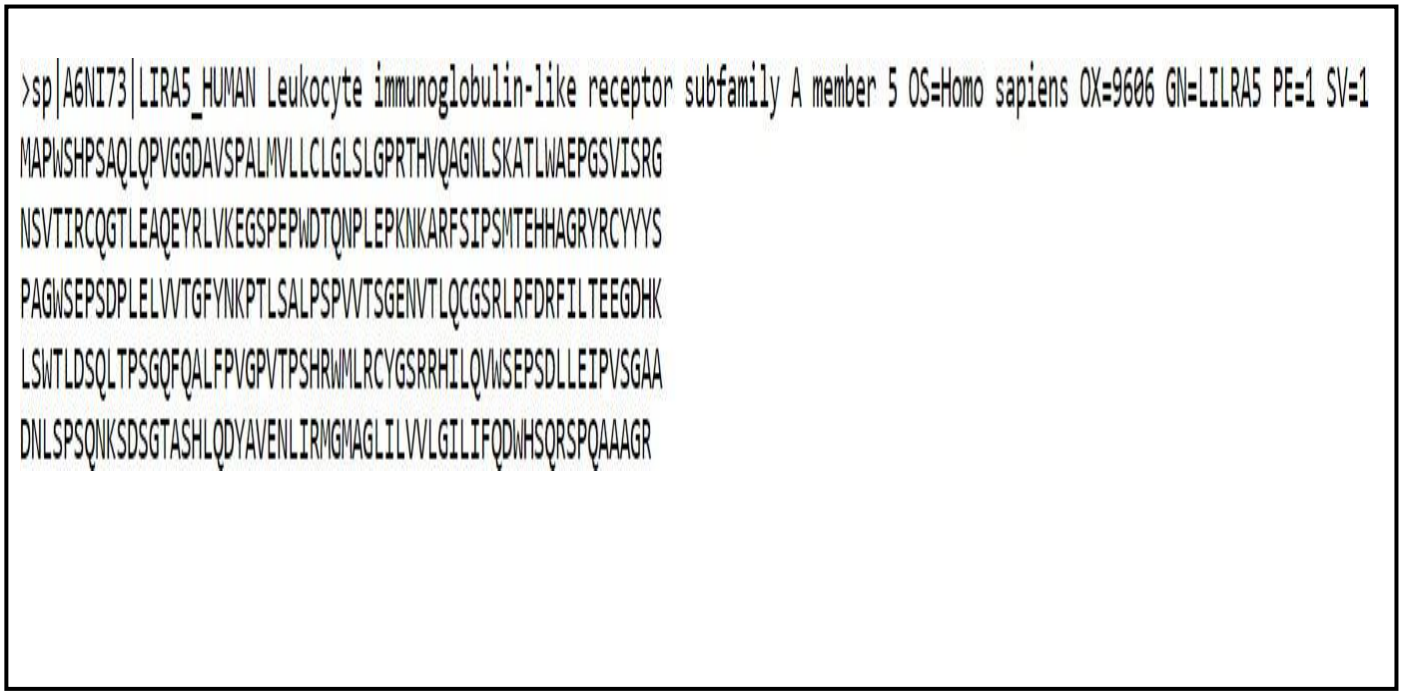


Fig 1: FASTA Sequence for LILRA5 query (UniProt Id:A6NI73)

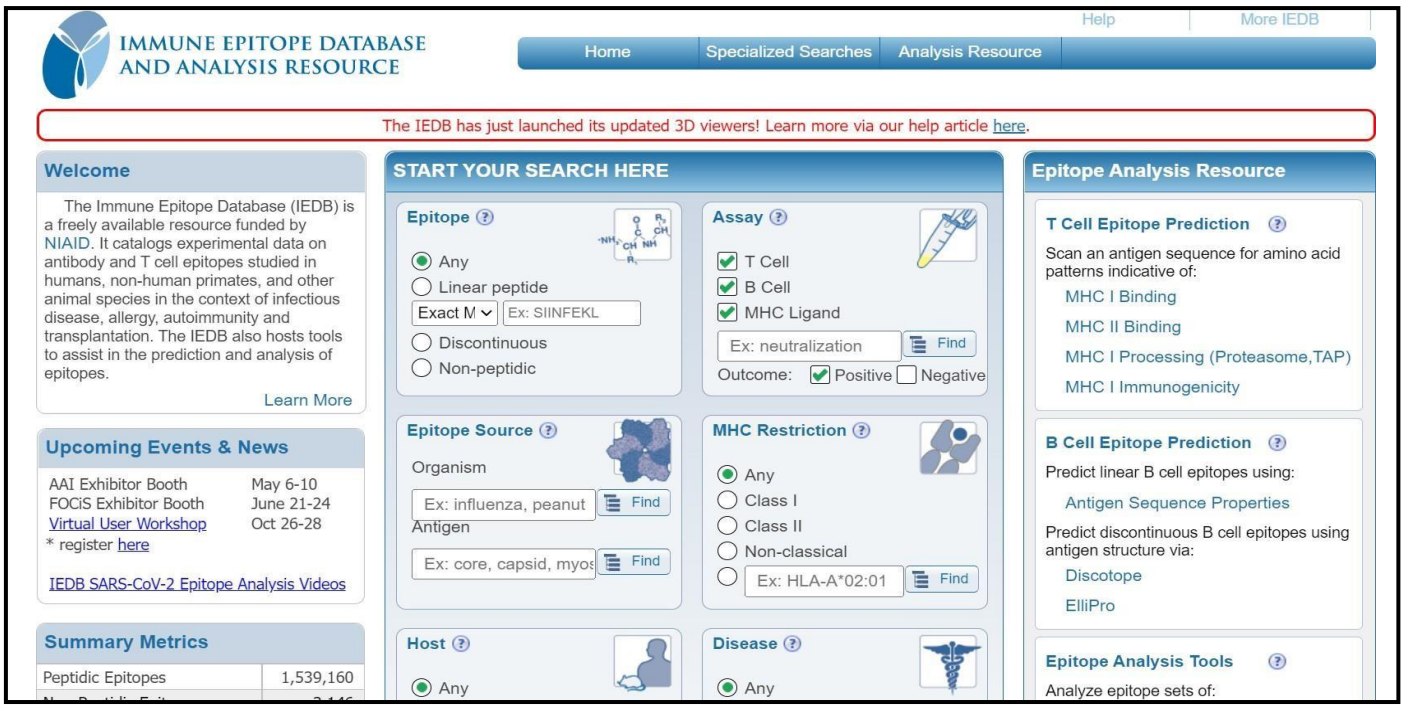


Fig 1.2: Homepage of IEDB database

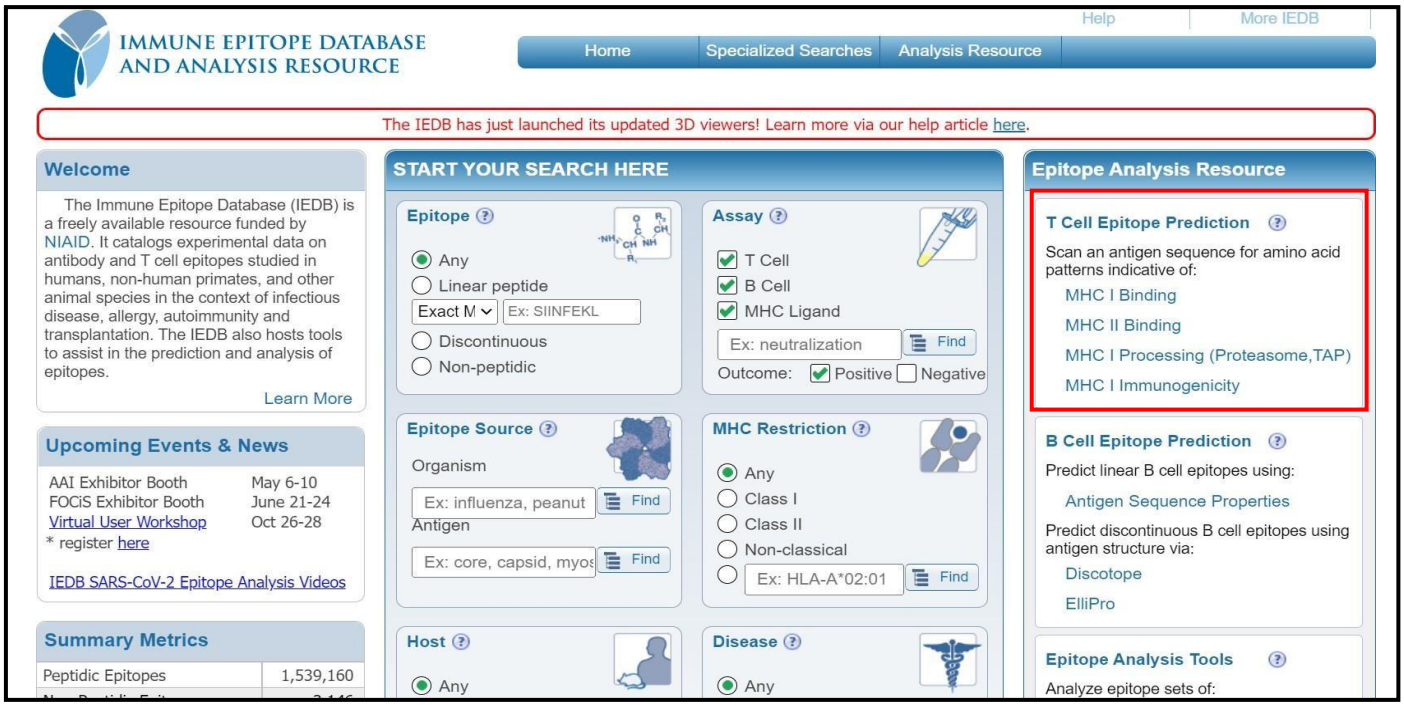


Fig 1.3: TCE Prediction under IEDB Database

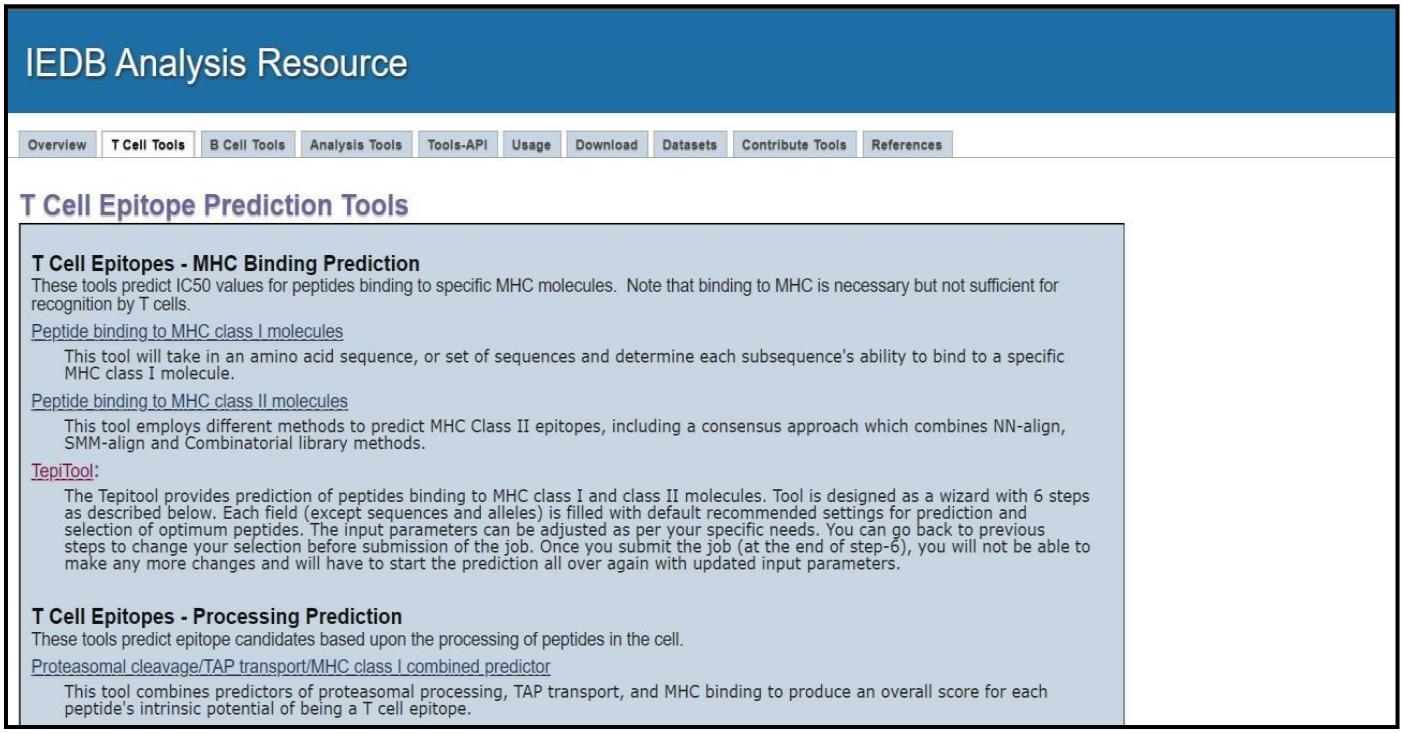


Fig 1.4: Homepage of TCE Prediction

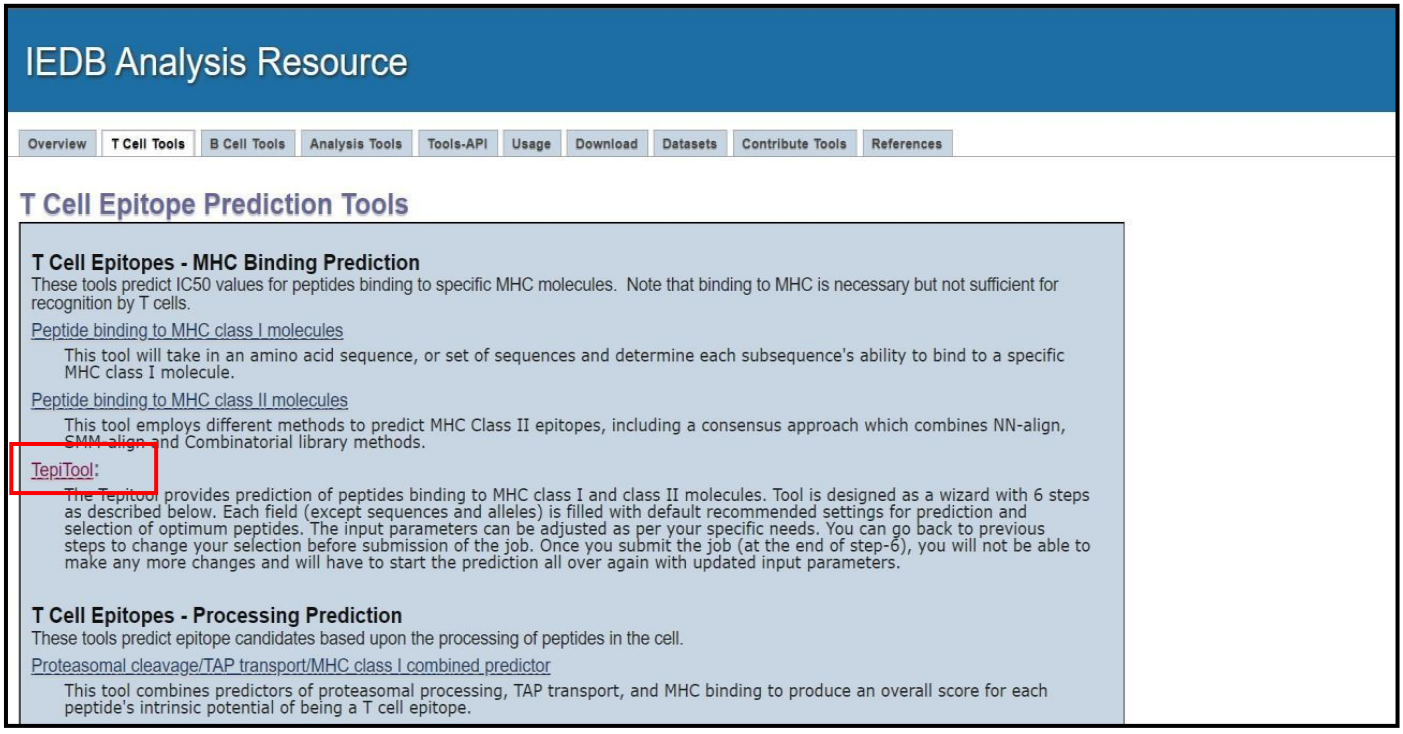


Fig 1.5: TepiTool under TCE Prediction Tools

MHC CLASS I

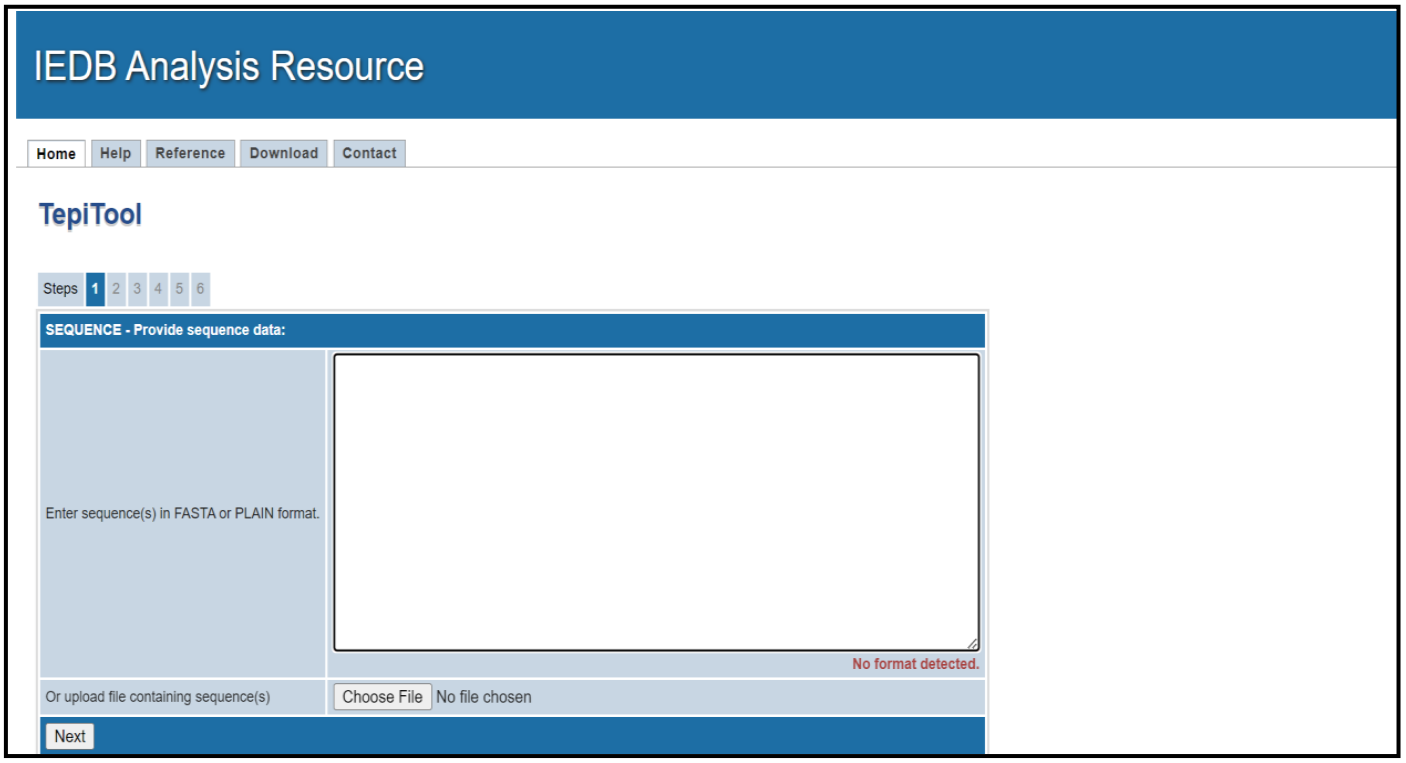


Fig 2: Search Option Under TepiTool

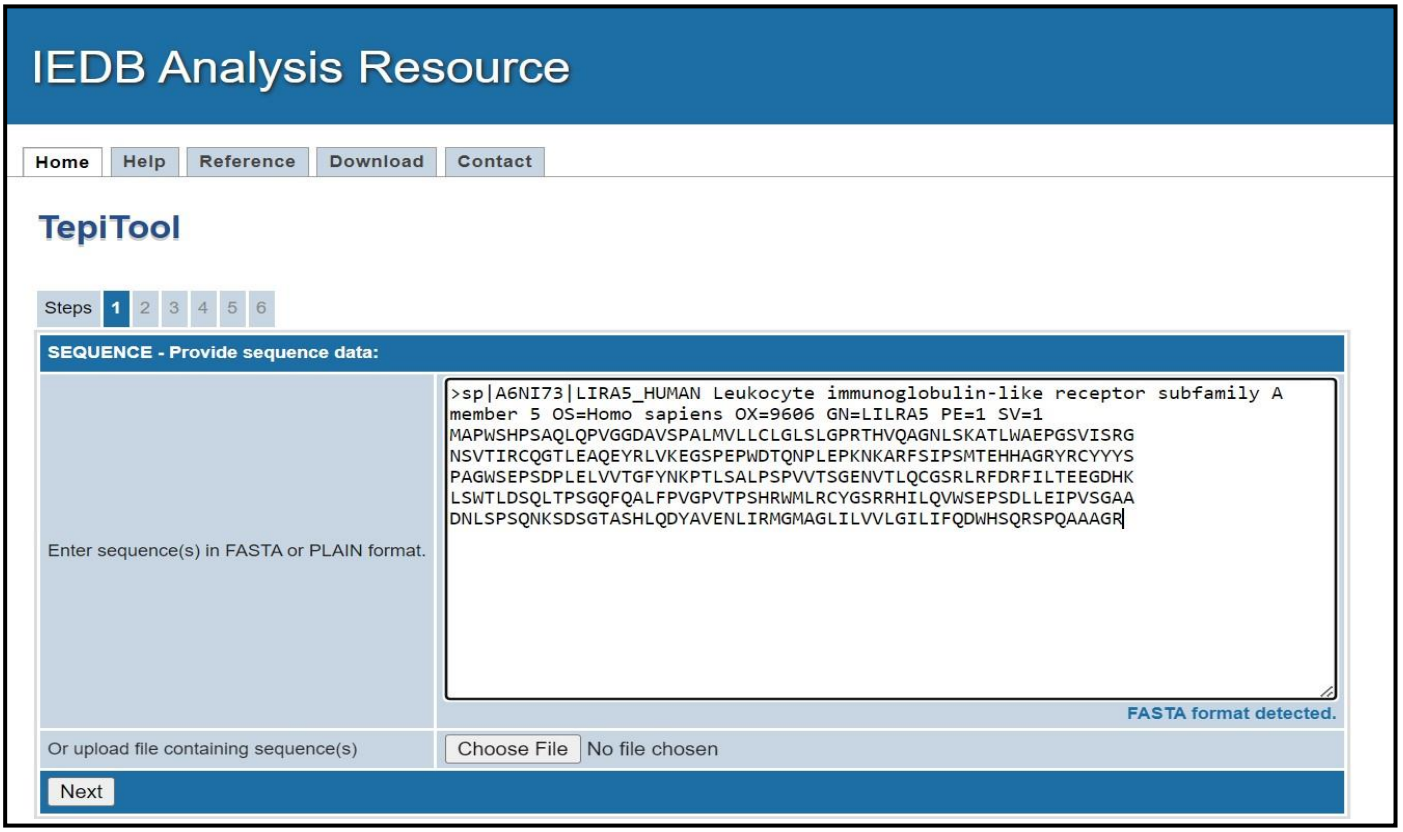


Fig 2.1: STEP 1- Provide FASTA sequence for query

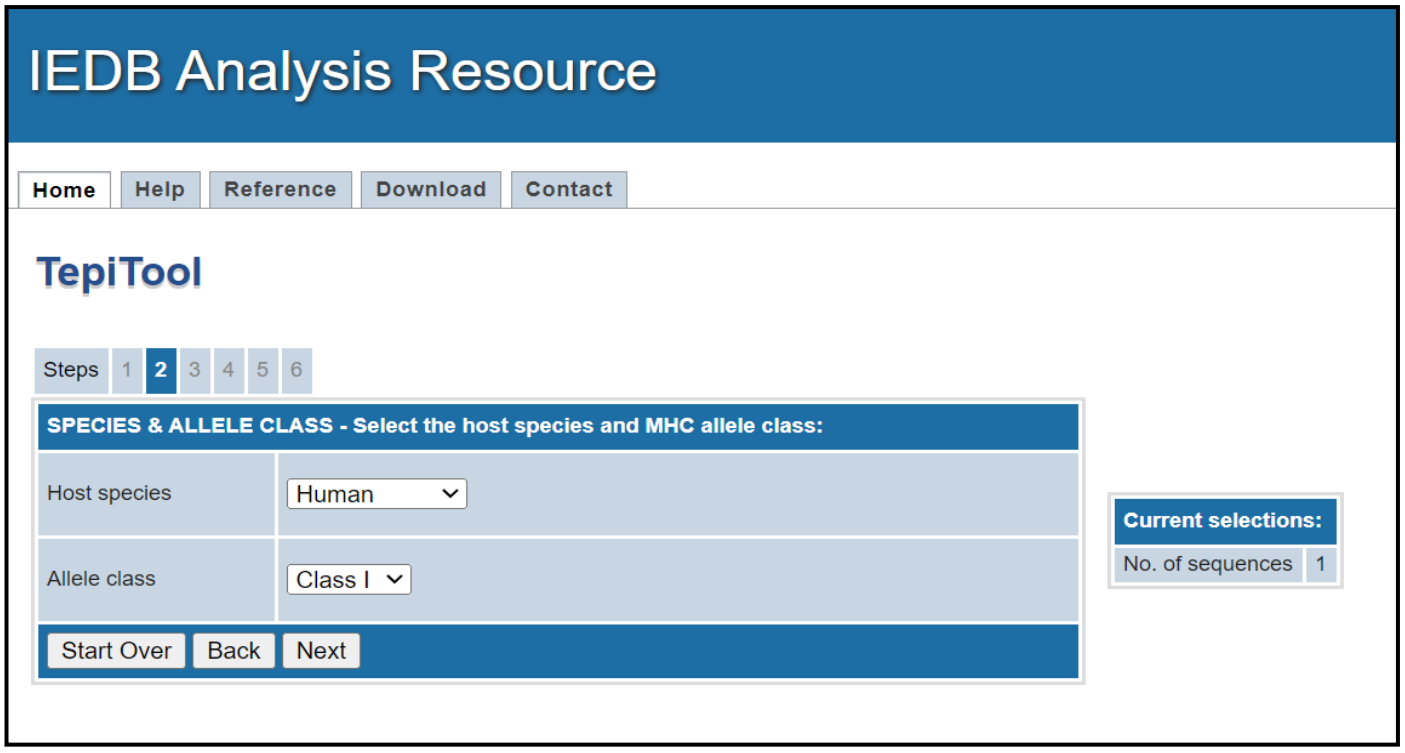


Fig 2.2: STEP 2- Select the host species and MHC allele class

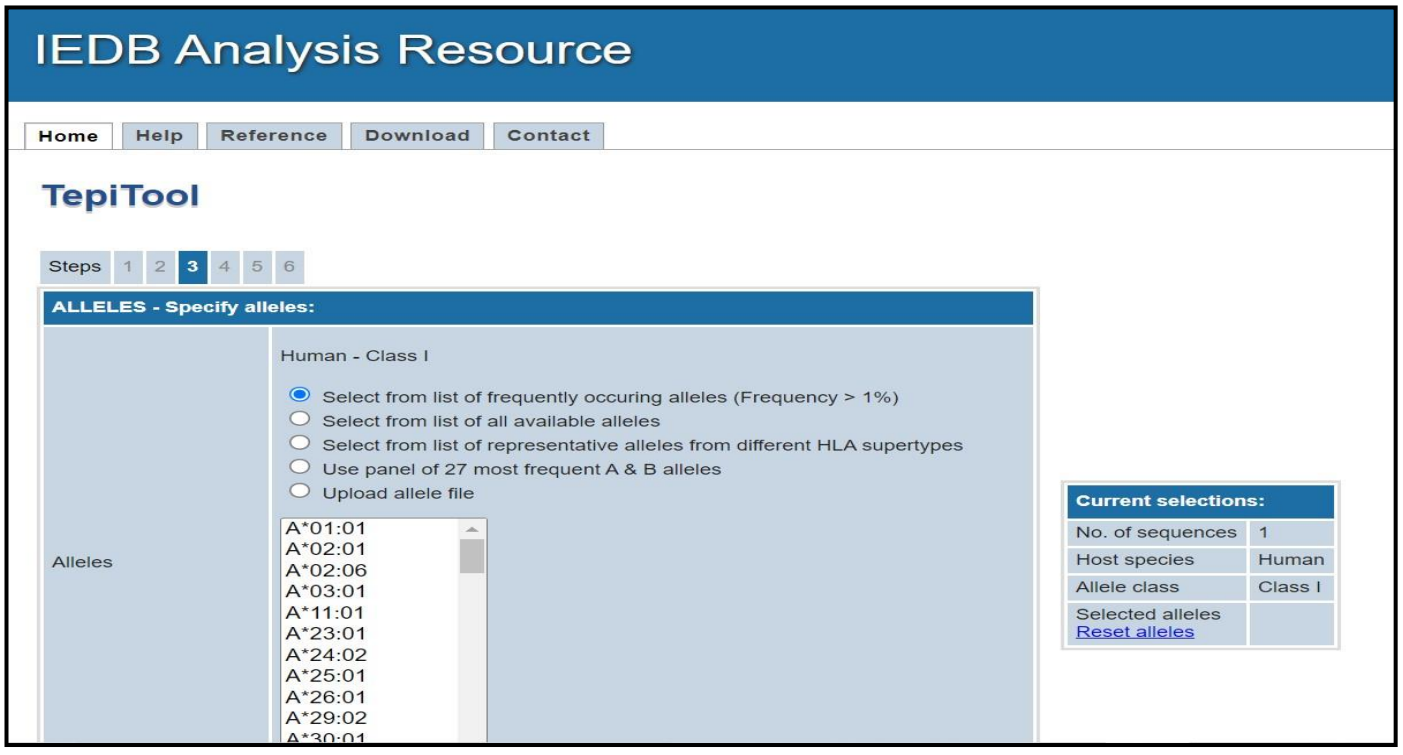


Fig 2.3: STEP 3- Mention the specific alleles

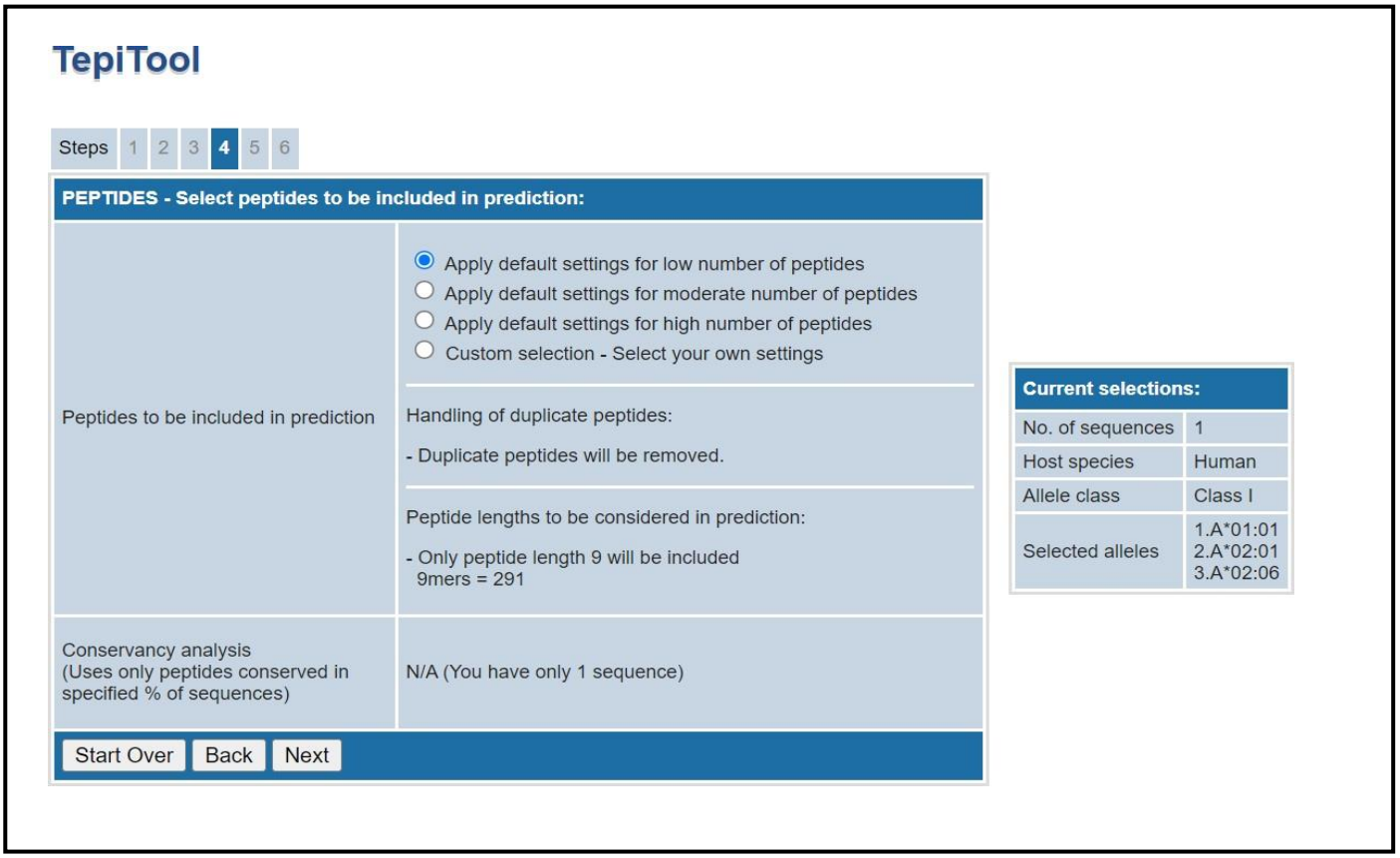


Fig 2.4: STEP 4- Select peptides to be included in prediction

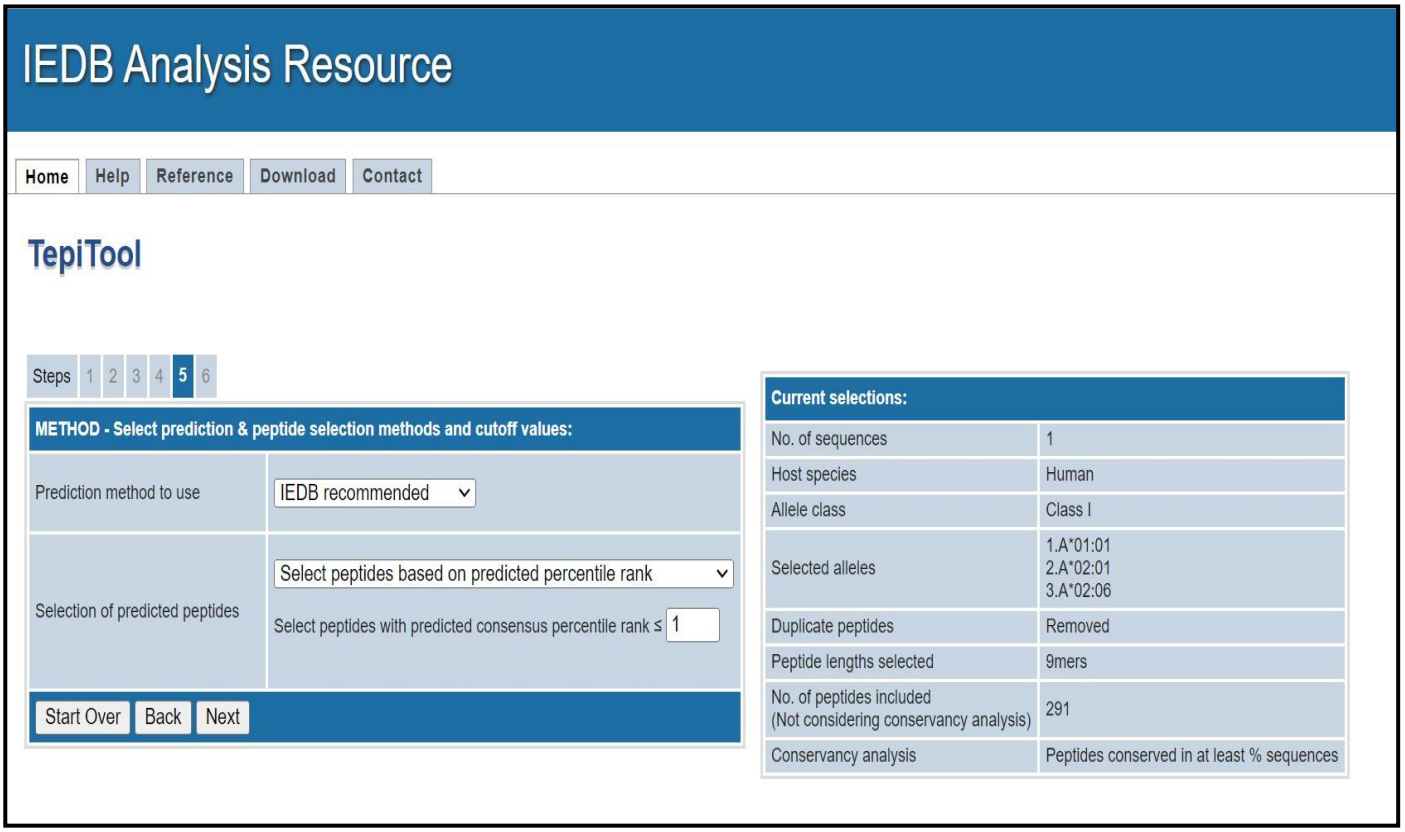


Fig 2.5: METHOD - Select prediction & peptide selection methods and cutoff values

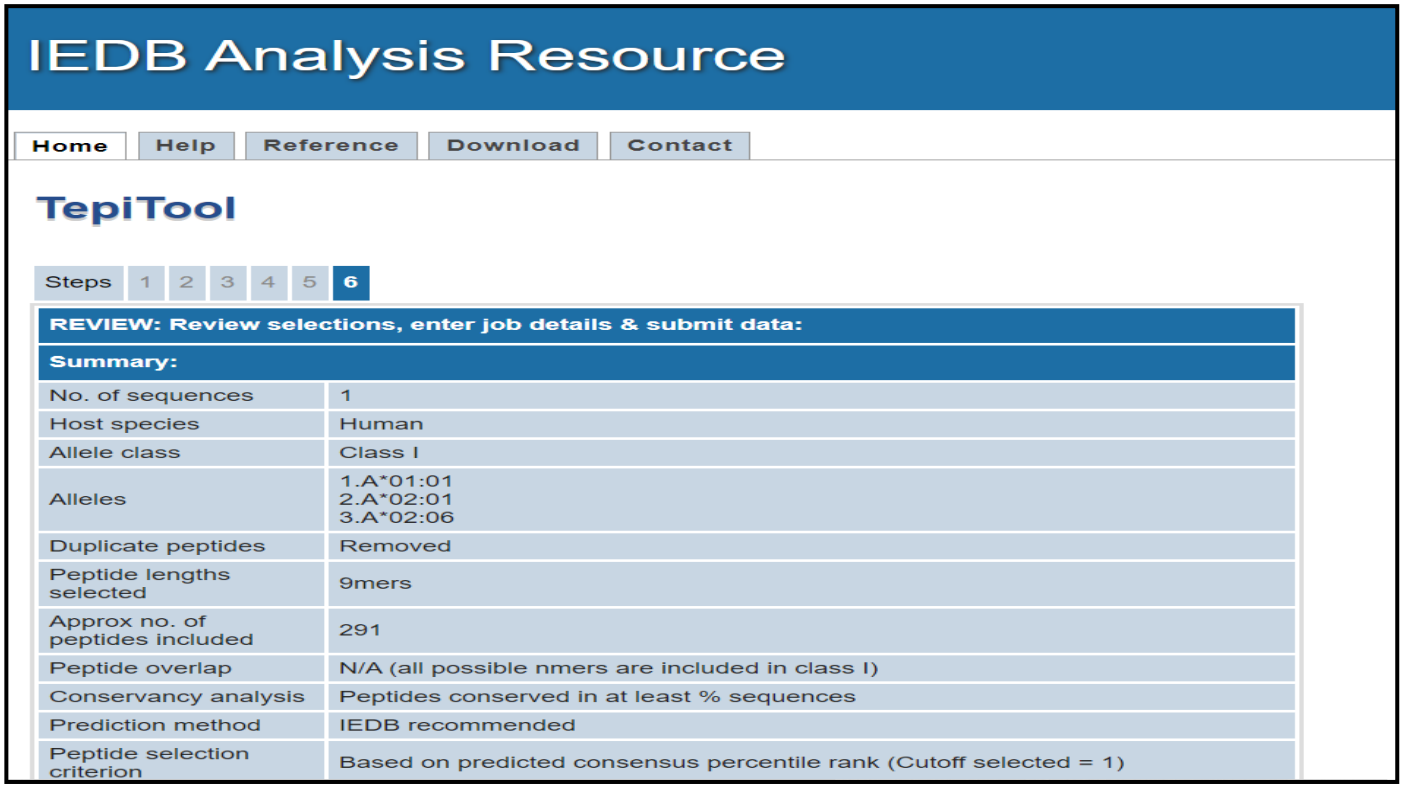
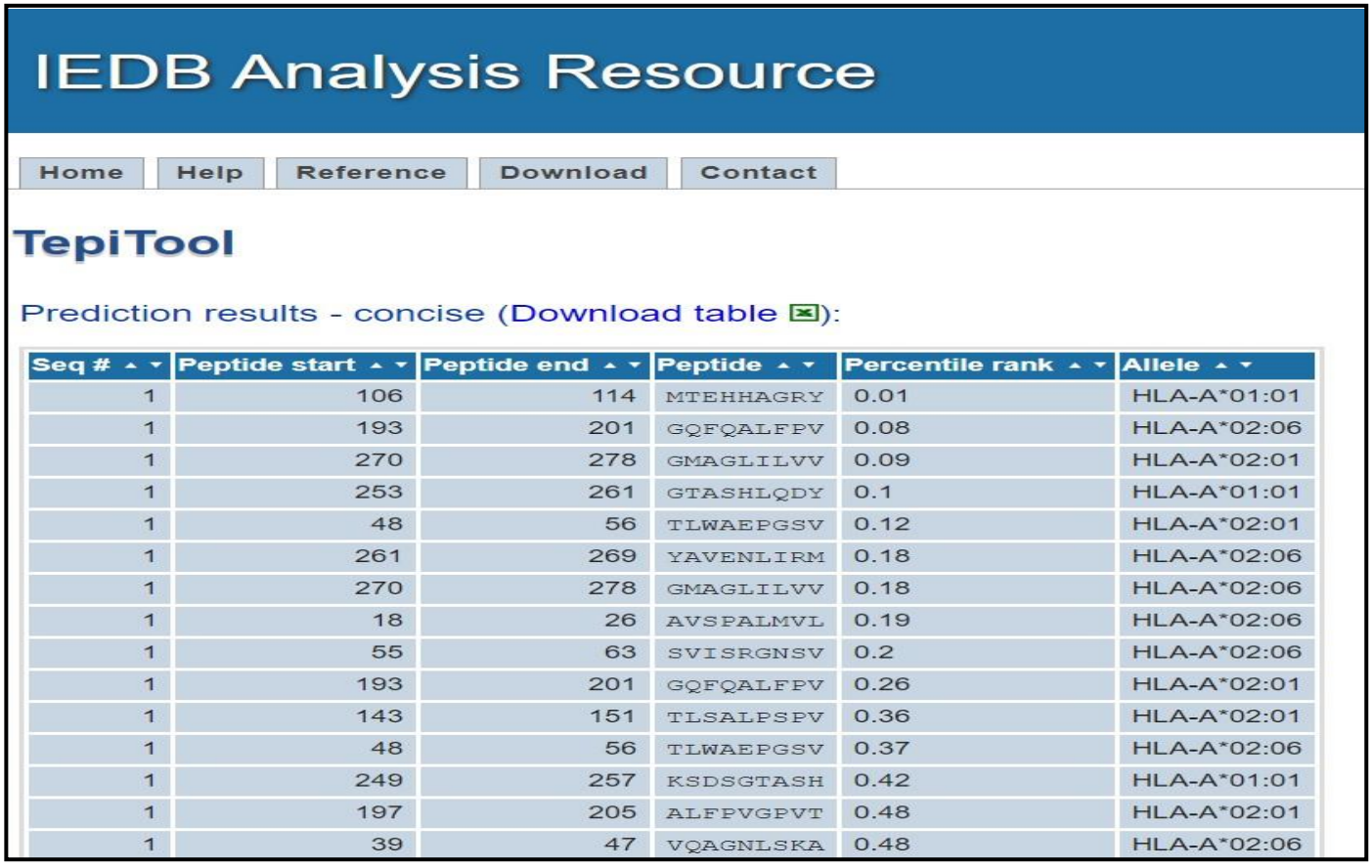


Fig 2.6: REVIEW: Review selections, enter job details & submit data



Seq #	Peptide start	Peptide end	Peptide	Percentile rank	Allele
1	106	114	MTEHHAGRY	0.01	HLA-A*01:01
1	193	201	GQFQALFPV	0.08	HLA-A*02:06
1	270	278	GMAGLILVV	0.09	HLA-A*02:01
1	253	261	GTASHLQDY	0.1	HLA-A*01:01
1	48	56	TLWAEPGSV	0.12	HLA-A*02:01
1	261	269	YAVENLIRM	0.18	HLA-A*02:06
1	270	278	GMAGLILVV	0.18	HLA-A*02:06
1	18	26	AVSPALMVL	0.19	HLA-A*02:06
1	55	63	SVISRGNSV	0.2	HLA-A*02:06
1	193	201	GQFQALFPV	0.26	HLA-A*02:01
1	143	151	TLSALPSPV	0.36	HLA-A*02:01
1	48	56	TLWAEPGSV	0.37	HLA-A*02:06
1	249	257	KSDSGTASH	0.42	HLA-A*01:01
1	197	205	ALFPVGPVT	0.48	HLA-A*02:01
1	39	47	VQAGNLSKA	0.48	HLA-A*02:06

Fig 3: Resultpage for LILRA5 under TepiTool

MHC CLASS II:

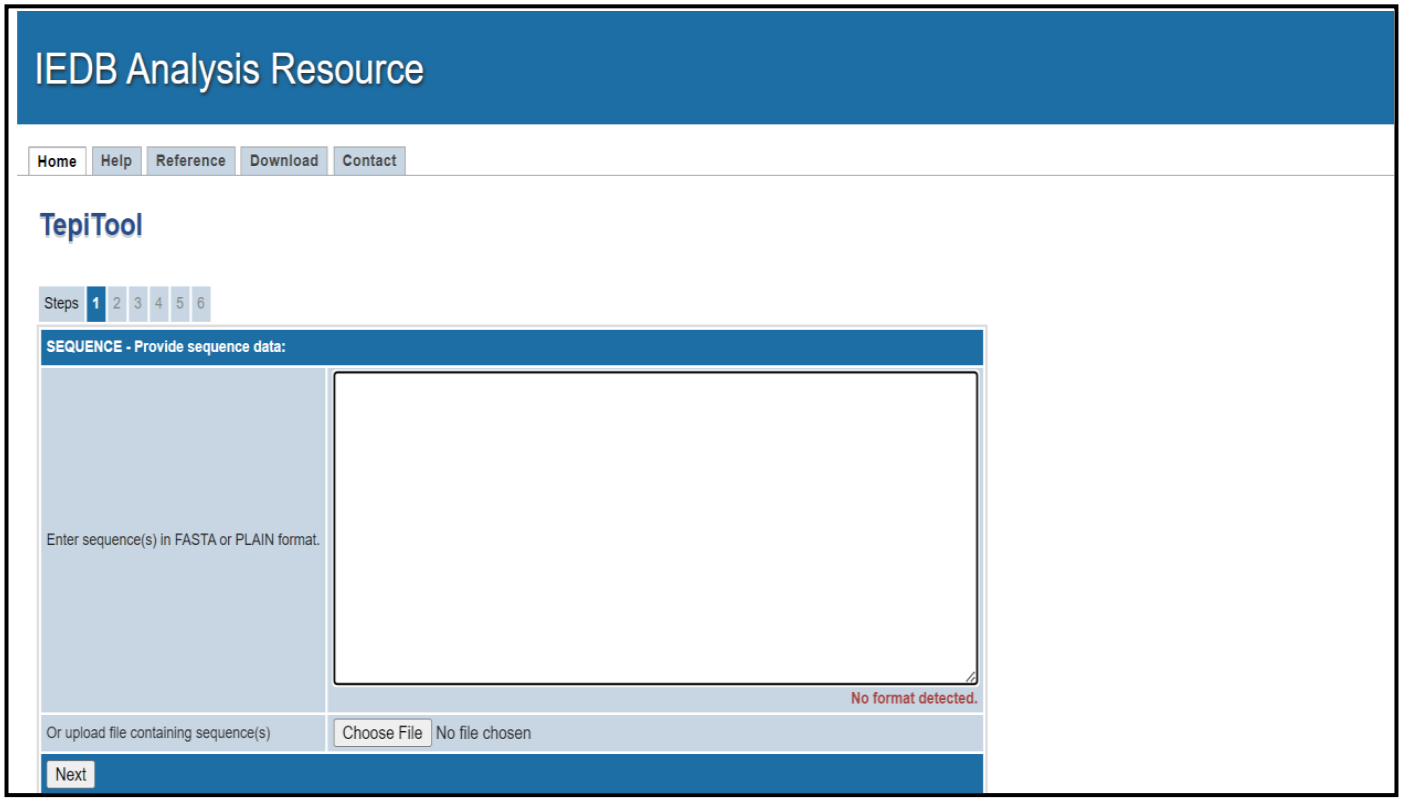


Fig 4: Search Option Under TepiTool

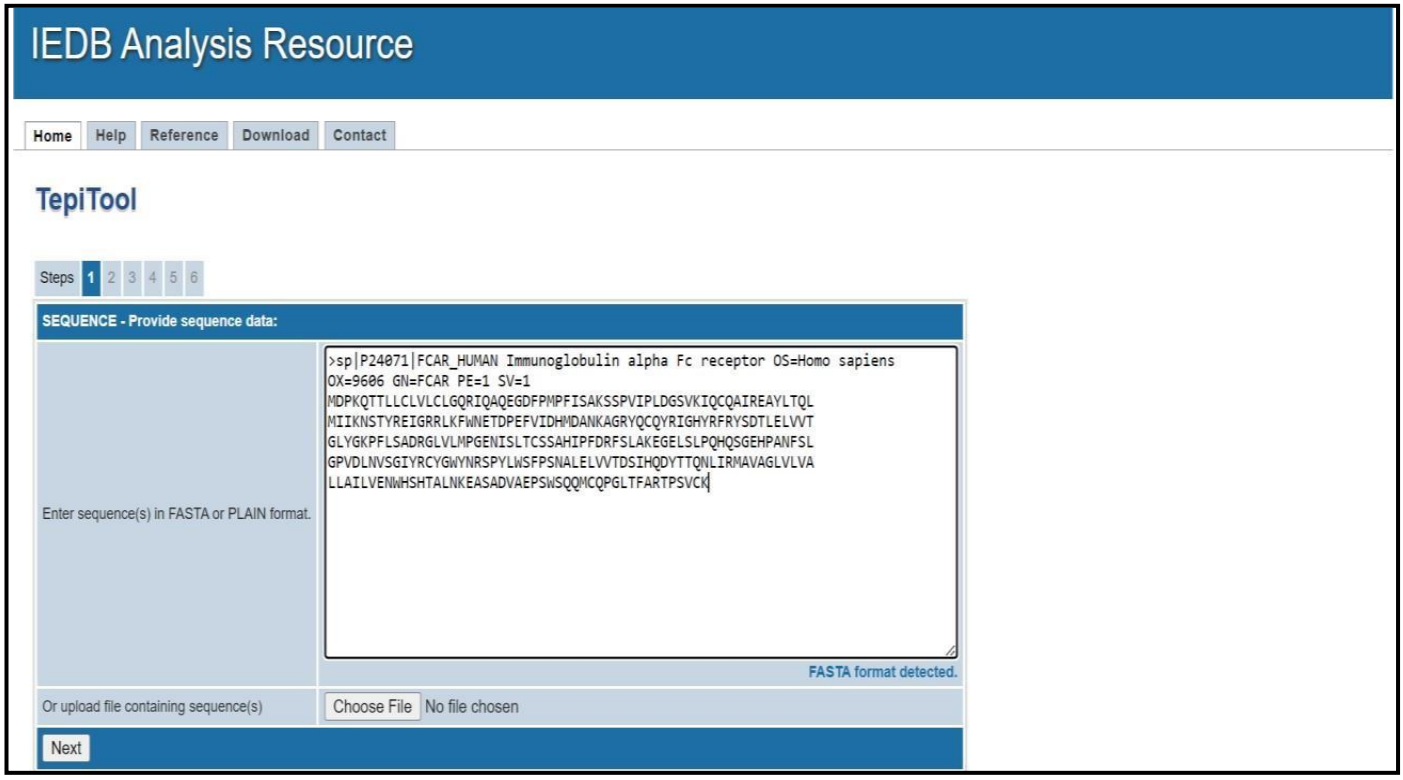


Fig 4.1: SEQUENCE - Provide sequence data

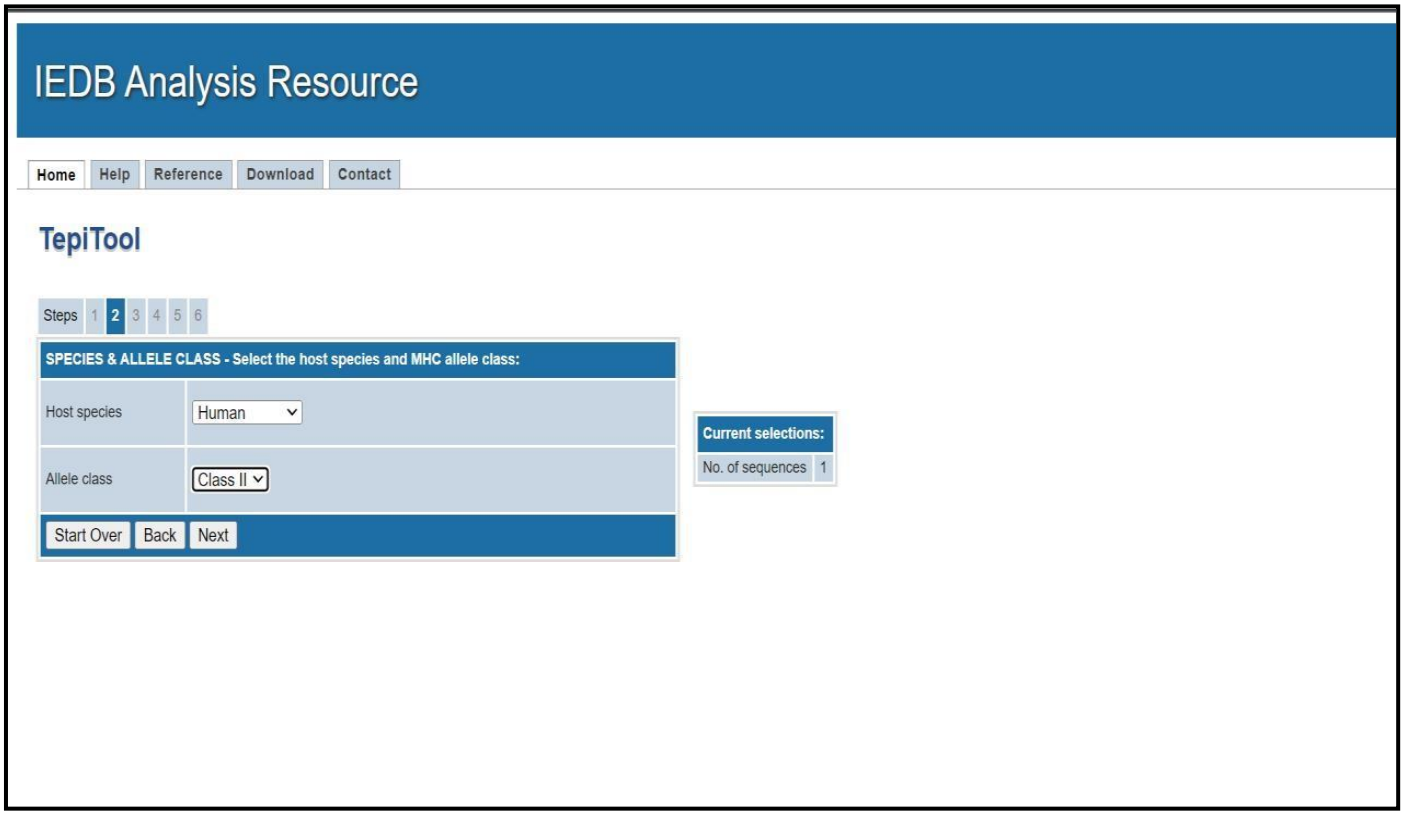


Fig 4.2: SPECIES & ALLELE CLASS - Select the host species and MHC allele class

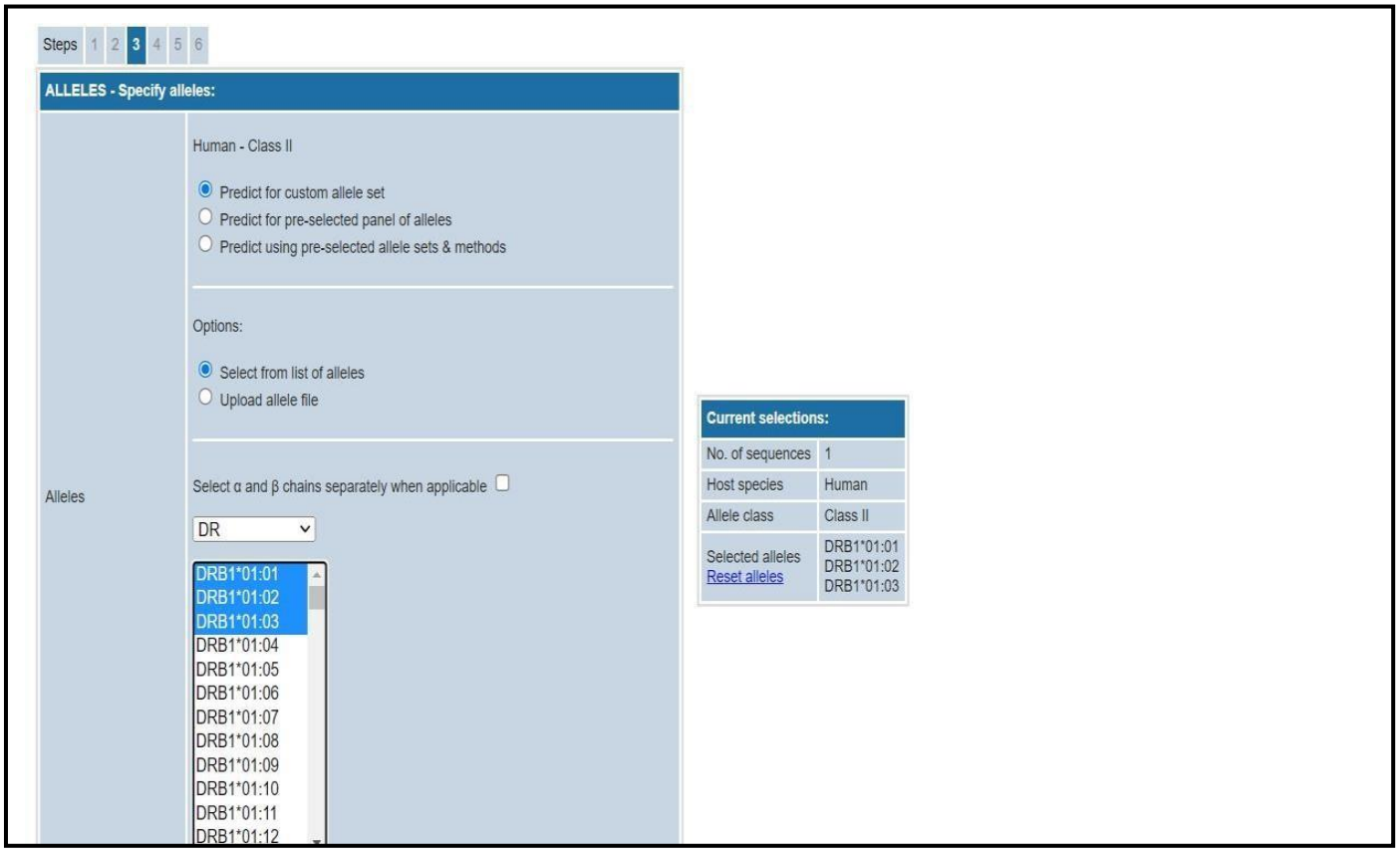


Fig 4.3: ALLELES - Specify alleles

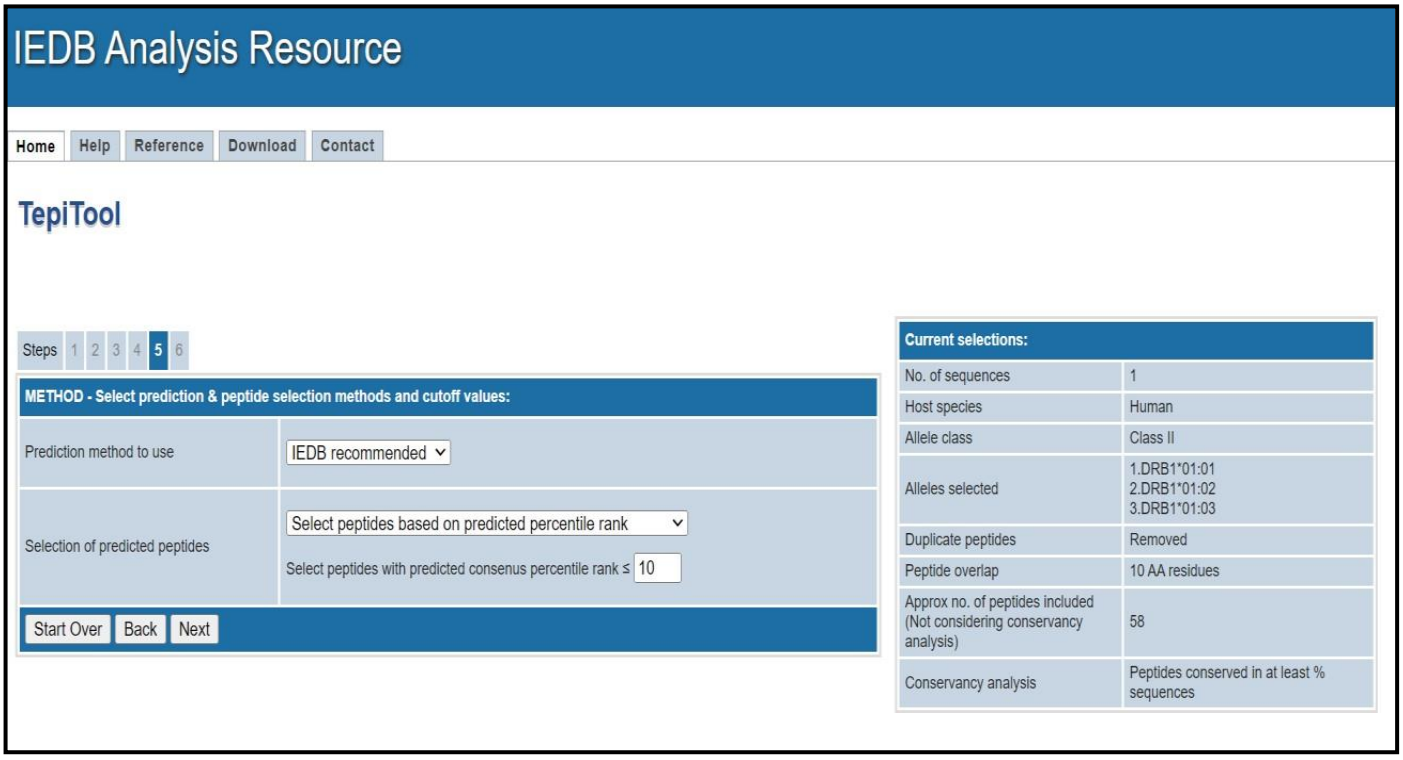


Fig 4.4: PEPTIDES - Select peptides to be included in prediction

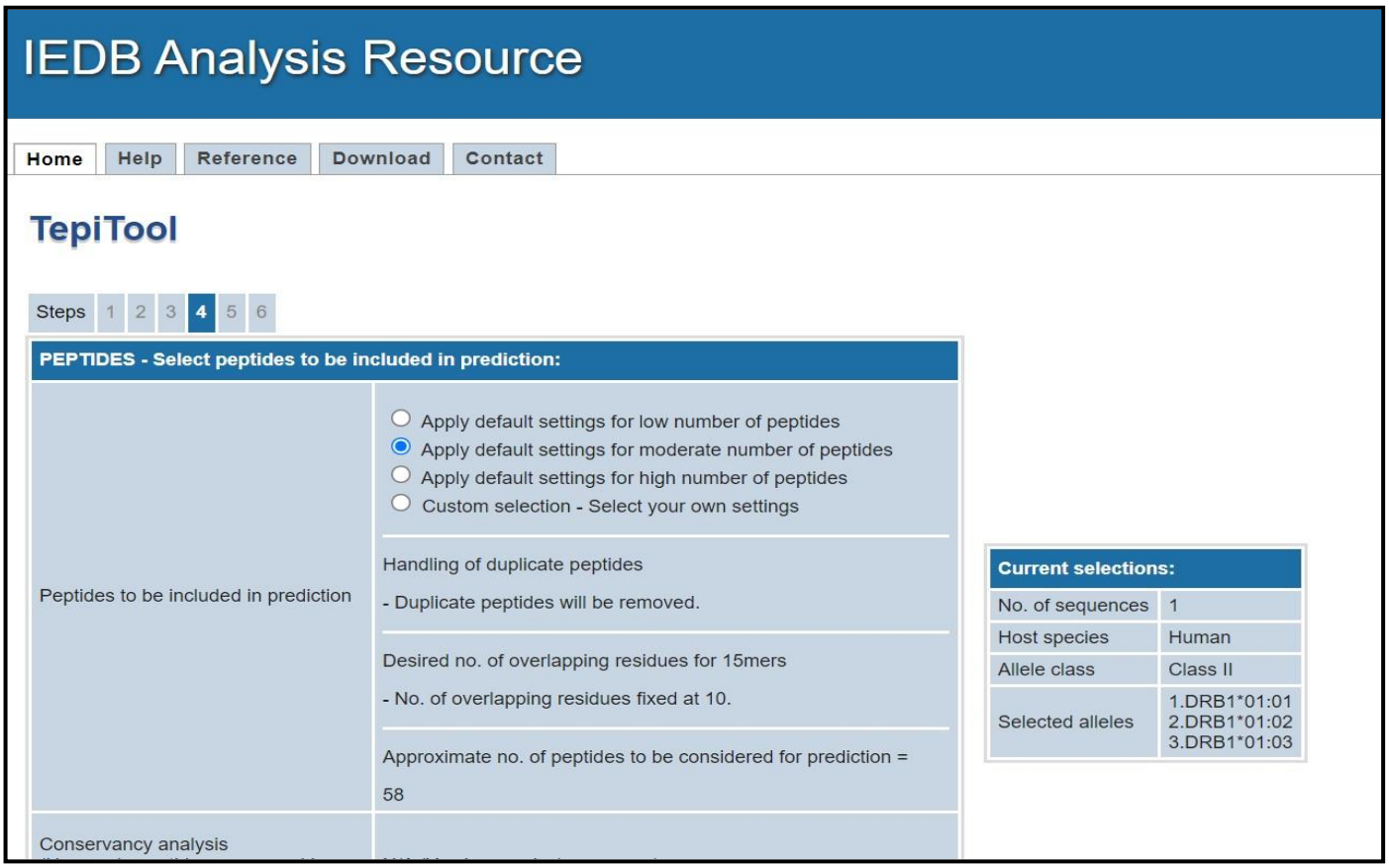


Fig 4.5: METHOD - Select prediction & peptide selection methods and cutoff values

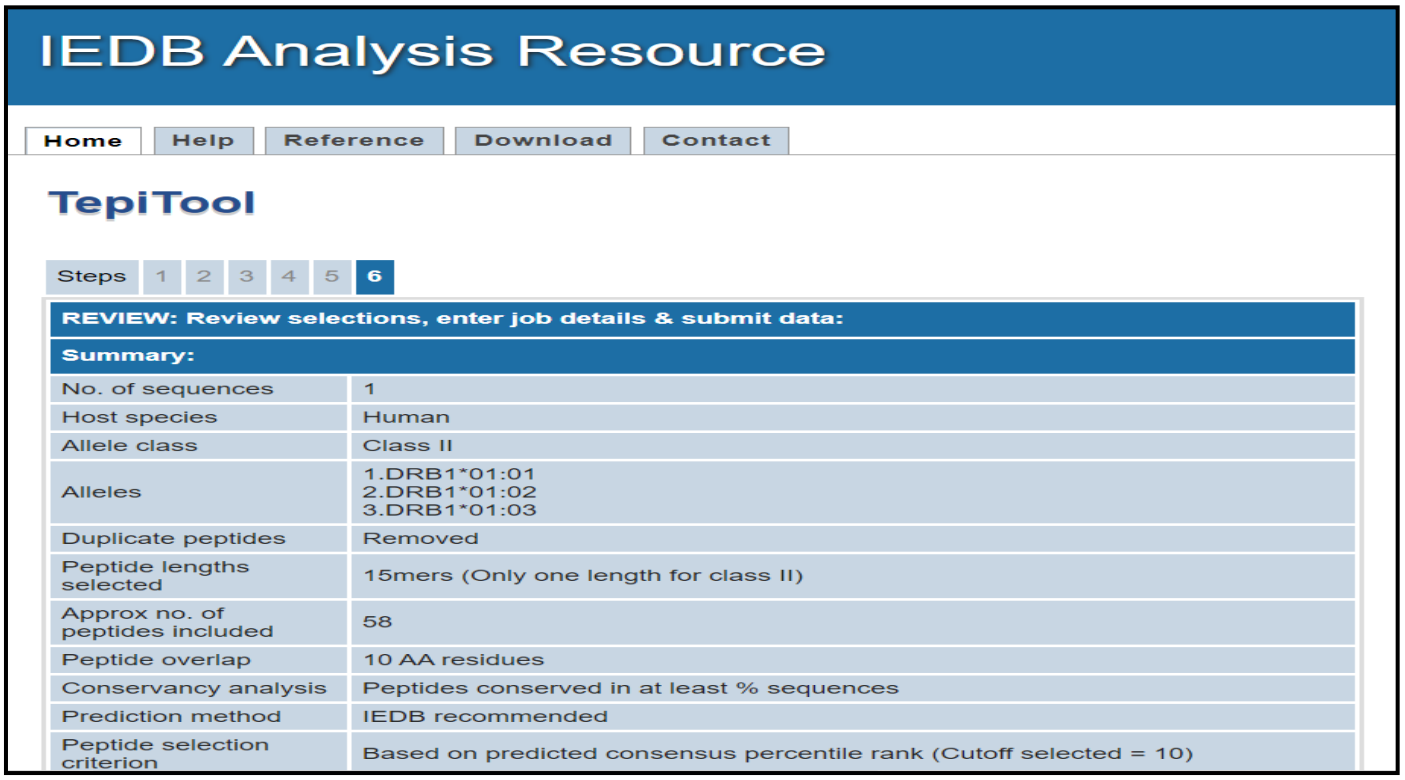


Fig 4.6: REVIEW: Review selections, enter job details & submit data

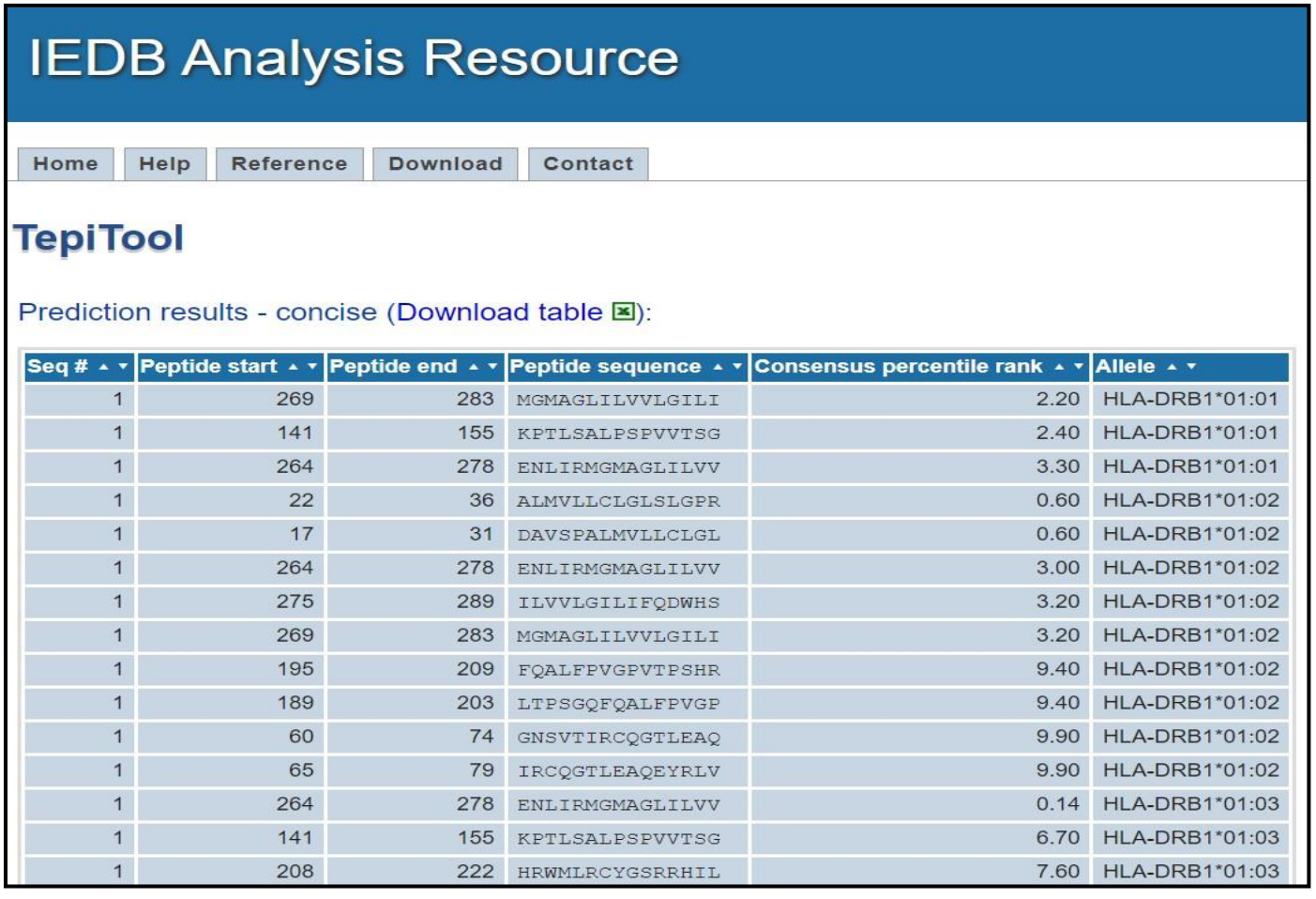


Fig 5: Resultpage for LILRA5 under TepiTool

RESULTS:

TepiTool Prediction for structure Immunoglobulin alpha Fc receptor MHC Class I and MHC Class II:

The prediction results page is divided into 2 sections. The details are given below:

1. **Concise results/ Prediction Results:**

The concise results table shows the final list of predicted peptides selected based on the input parameters provided. The table will contain the sequence # of the peptide's source protein in the input sequence set, start and end positions of the peptide within the source protein sequence, the peptide sequence, the selection criterion parameter value (consensus percentile rank), the allele (where applicable). For the Immunoglobulin alpha Fc receptor query, the lower value is 0.01 indicating better binding affinity. Also the concise results table can also be downloaded as .csv file which can be opened using any spreadsheet such as MS Excel for further analysis. This section will be included in the email.

2. **Download results details:**

This section provides links for downloading the results details as csv files. It can include the following based on the input parameters:

1. Non-redundant results: This file will contain prediction results with redundant peptides within each sequence removed. Redundant peptides means peptides that overlap with more residues than desired. This result set includes both positive and negative peptides basedon the input parameters.
2. Complete results: This file will contain binding predictions of all peptides. This will include the predicted IC50 and percentile rank or other scores depending on the prediction method chosen. In case of IEDB recommended or consensus method, the results will include details from each of the prediction methods employed.
3. Conservancy of peptides (applicable only if conservancy analysis is done): This file will contain conservancy of each peptide in the input sequence set.

CONCLUSION:

Computational prediction of T-cell epitope candidates is currently being used in several applications including vaccine discovery studies, development of diagnostics and removal of unwanted immune responses against protein therapeutics. This protocol provides step-by-step instructions with necessary recommendations for prediction of the best T-cell epitope candidates in line with the newly developed online tool called TepiTool. The TepiTool provides some of the top MHC class I and class II binding prediction algorithms for a number of species including humans, chimpanzees, bovines, gorillas, macaques, mice and pigs. The tool is designed as a user-friendly wizard with well-defined steps which helps the users to predict the best MHC binding peptides from their sequences of interest.

REFERENCES:

1. TepiTOOL. (n.d.). IEDB Analysis Resource. Retrieved October 22, 2022, from http://tools.iedb.org/tepitool/
2. LILRA5. (n.d.). UniProt Database. Retrieved October 22, 2022, from https://rest.uniprot.org/uniprotkb/A6NI73 .fasta
3. Tepitool. (n.d.). NCBI. Retrieved October 22, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4981331/