Multi-epitope Peptide Vaccine Prediction for Candida albicans targeting Pyruvate Kinase Protein; an Immunoinformatics Approach

Anfal Osama Mohamed Sati[1], Abdelrahman Hamza Abdelmoneim Hamza[1], Enas Dawoud Khairi Dawoud[1], Tebyan Ameer Abdelhameed Abbas[1], Fatima Abdelrahman Bshier Abdelrahman[1], Khalid Abbas Hassan Saad[2], Rouaa Babikir Ahmed Abduallah[3] , Marwa Mohamed Elhag Saeed Mustafa[1], Tamador abdelrahman Sidahmed[4], Reham M. Elhassan[5,6], Mohamed A. Hassan[1,7]

[1] Department of applied bioinformatics, Africa City of Technology, Khartoum- Sudan

[2] Faculty of medicine and health sciences, Omdurman Islamic University, Khartoum- Sudan

[3] Faculty of medical laboratory sciences, Omdurman Islamic university, Khartoum – Sudan

[4] Faculty of microbiology and molecular biology, Alnailain University, Khartoum- Sudan

[5] Department of pharmaceutical science, Shandong University, Jinan- Republic of China

[6] Department of pharmaceutical chemistry, Faculty of pharmacy, Sudan International University, Khartoum- Sudan

[7] Research and development, Detagen, kayseri- Turkey

Abstract:

The fungus *Candida albicans* is an opportunistic pathogen that causes a wide range of infections. It's the primary cause of candidiasis and the fourth most common cause of nosocomial infection. In addition, disseminated invasive candidiasis which is a major complication of the disease has an estimated mortality rate of 40%-60% even with the use of antifungal drugs. Over the last decades, several different anti-Candida vaccines have been suggested with different strategies for immunization against candidiasis such as, live-attenuated fungi, recombinant proteins, and glycoconjugates but none has been approved by the FDA, yet. This study aims to introduce a new possible vaccine for *C. albicans* through analyzing peptides of its pyruvate kinase (PK) protein as an immunogenic stimulant computationally.

 A total number of 28 *C. albicans*, pyruvate kinase proteins were obtained from NCBI on the 9th of February 2019 and were subjected to multiple sequence alignment using Bioedit for conservancy. The main analytical tool was IEDB, Chimera for homology modelling, and MOE for docking.

Among the tested peptides, fifteen promising T-cell peptides were predicted. Five peptides were more important than the others (HMIFASFIR, YRGVYPFIY, AVAAVSAAY, LRWAVSEAV, and IFASFIRTA) They show high Binding Affinity to MHC molecules, low binding energy required indicating more stable bonds, and their ideal length of nine peptides. (PTRAEVSDV) peptide is the most promising linear B-cell peptide due to its physiochemical parameters and optimal length (nine amino acids). It's highly recommended to have these five strong candidates in future in vivo and in vitro analysis studies.

1. Introduction:

The fungus *Candida albicans* is an opportunistic pathogen that causes a wide range of infections. *Candida albicans* is the primary cause of candidiasis and is the fourth most common cause of nosocomial infection. In addition, Disseminated invasive candidiasis that is a major complication of the disease has an estimated mortality rate of 40%-60% even with the use of antifungal drugs.[1–5] it occurs with various other manifestations ranging from vaginal infections, which affect up to 75% of the women at least once in their lifetime, to deep systemic infections in hospitalized patients leading to high morbidity and mortality rates specially, in immunocompromised individuals, including organ recipients and HIV patients [1,2,6–13]

Three proteins are most likely to be immunogenic in candida albicans which are the fructose-bisphosphate aldolase (Fba1), the phosphoglycerate kinase (Pgk) and the pyruvate kinase (Pk) with (pk) and (Fba1) being more immunodominant .[11,14] PK a 504 bp long, highly conservative protein which catalyzes the final step in glycolysis, converting phosphoenolpyruvate (PEP) and adenosine diphosphate (ADP) to pyruvate and adenosine triphosphate (ATP), once pyruvate is synthesized; it is either fed into the tri-carboxylic acid (TCA) cycle for oxidative phosphorylation or reduced to lactate under anaerobic conditions by lactate dehydrogenase enzyme LDH. [15–17]

Vaccination is a significant approach to improve the standard of public health care and provide cost effective way to control the increasing numbers of infections. Many previous studies reached to the conclusion that C. albicans has become a multidrug resistant fungi. C. albicans growing in low oxygenated environments such as intravenous lines, renal catheters, bio prosthetic devices and self-organizing into biofilms which are resistant to a wide range of antifungal drugs, making conventional antifungal agents ineffective for the treatment of C. albicans infections.[1,18–23] Adding to that, the prolonged use of antifungal agents such as azoles, polyenes, and echinocandinsin in immuno-compromised individuals may lead to an increase in C. albicans resistance to many of the drugs used currently. [4,5,11,24]

Over the last decades, several different anti-Candida vaccines have been suggested with Different strategies for immunization against candidiasis such as, live-attenuated fungi, recombinant proteins, and glycoconjugates25[26,27].Among many vaccines candidates, two univalent subunit vaccines (Als3 and the Sap2 v) and one recombinant C. albicans adhesin/invasin protein (NDV-3A) have been under development in recent years but none has been approved by the FDA, yet. The difficulties originates from the ability of C. albicans to evolve rapidly and lose host immune system's recognition and the problem of how to trigger an immune responses in immunocompromised individuals. This calls for a new and a more effective version of these vaccines. [24,27–32]

This study aims to introduce a new possible vaccine for *C. albicans* through analyzing peptides of its pyruvate kinase (PK) protein as an immunogenic stimulant computationally. Vaccine production that depends on biochemical experiments can be

expensive, time consuming, hazardous and more likely to give unsatisfactory results, that's why bioinformatics approaches provide more affordable, efficient and faster techniques for vaccine development.[33–35] Using immunoinformatics approaches, otherwise known as computational immunology which is an emerging specialization of bioinformatics that has recently been used to predict possible peptide based vaccines using massive online databases. This study is different than other studies for being the first in scilico analysis study of its kind attempting to design a peptide- based vaccine for *C. albicans* using Pyruvate kinase (PK) as an immunogenic stimulant.
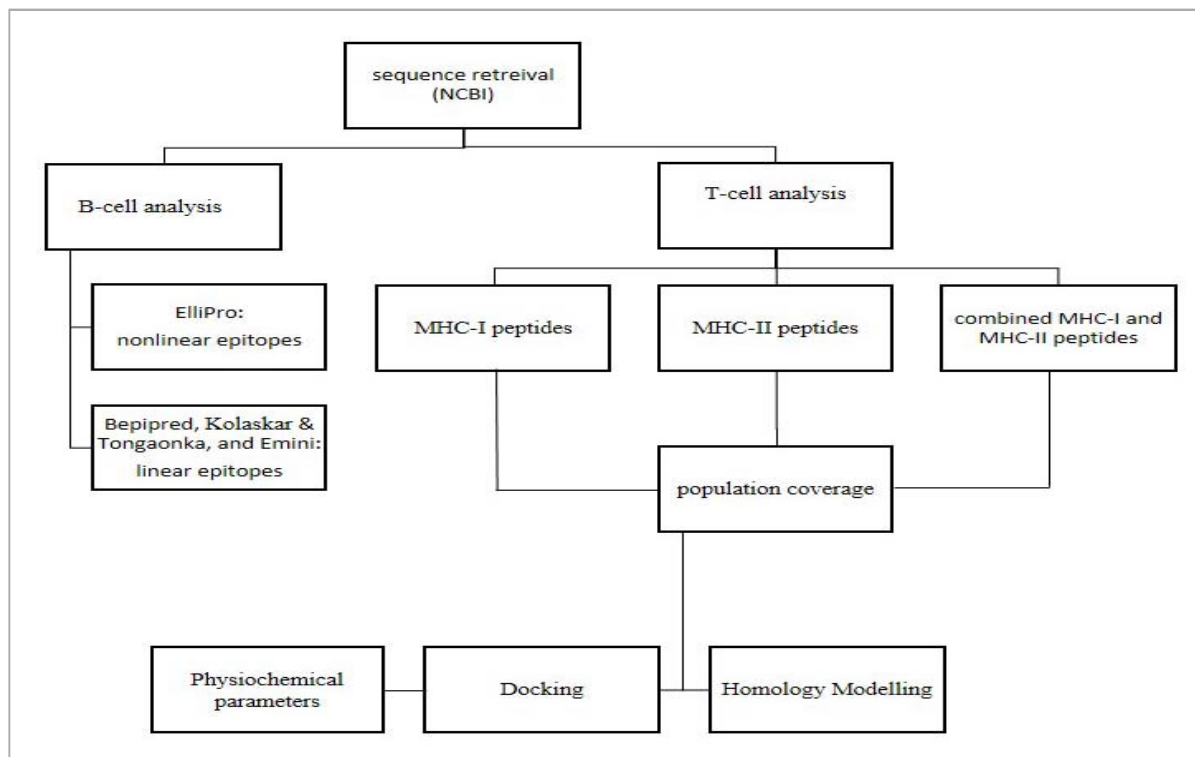
METHODOLOGY:

Figure (1): a hierarchy template to visualize the work flow followed in the study

1.1. Sequence retrieval:

A total number of 28 *candida albicans*, pyruvate kinase proteins were obtained from national center of biotechnology information NCBI on the 9th of February 2019. (Available at: https://www.ncbi.nlm.nih.gov/)

1.2. Multiple sequence Alignment:

A multiple sequence alignment was made using the Clustal w package, Bioedit tool for the sequences by blasting them against reference sequence (accession: XP_714934.1) to show areas of protein conservation in the different species and exclude peptides located in areas of low conservation. Peptides located at highly conserved regions are more likely

to give stronger vaccine that covers more populations than peptides at low areas of conservation.[36,37]

1.3. B-cell epitopes prediction:
1.3.1. Linear epitopes:

Was made through Immune Epitope Database Analysis Resource IEDB tool for Linear b-cell epitopes prediction. The tool has different methods to predict epitopes antigenicity, flexibility, immunogenicity, and surface accessibility. This this done through the submission of the Swiss-prot ID of the reference protein or FASTA. Then peptides larger than six peptides were spliced to increase the possibility of obtaining peptides with higher scores in Bepipred, Kolaskar & Tongaonka, and Emini. Finally, a total number of 144 peptides were obtained, 129 of which were conserved and were put to the Bepipred surface test and fifteen peptides were located in non-conserved regions and sequentially were removed. (Available at: http://tools.iedb.org/main/bcell/)

1.3.1.1 Bepipred linear epitope prediction:

Peptides were introduced to this test to check their linearity, linear peptides are simple, unbranched, where non-linear peptides are found in clusters. Linear peptides that are located on surface have higher probability of provoking the immune system to induce a response than non-linear or peptides which are not on surface. Peptides that passed the linearity test scored above the threshold of one, were 106 peptides, while the 23 other peptides were excluded for scoring less than one.[38–40]

1.3.1.2. Emini surface accessibility test:

It's a test that calculates surface probability using a specific formula. It was performed to linear peptides composed of six amino acids and had a threshold of one, 81 peptides passed while 25 peptides were not on the surface and were not further analyzed.[37]

1.3.1.2 Kolaskar & Tongaonka Antegenicity test:

It's a test that uses physiochemical properties of amino acids residues and the frequency of their appearance in peptides under study to predict protein antigenicity. The 81 peptides were submitted to the Antigenicity test with a threshold of 1.033, and five peptides passed indicating they are capable of inducing an immune response. [41]

1.3.2. Non-linear B-cell epitopes
1.3.2.1 ElliPro Antibody epitopes prediction:

It's a tool that works using the PDB ID of the protein, providing minimum score of (0.5) and maximum distance of (6) to predict non-linear peptides. It provides 3D model of the clustered peptides with the result. [42] (available at: http://tools.iedb.org/ellipro/)

1.4 T-cell epitopes prediction:
1.4.2 Binding to MHC class I prediction tool:

Is a tool which predicts peptides binding to MHC-I molecules following the submission of sequence, prediction method (Artificial neural network ANN method was used in this study), source of MHC molecules (Homo sapiens) and frequently occurring alleles, and HLA sets of twenty-seven alleles that are desired in the study. The result is then sorted by percentile rank and cutoff below IC50 of 500.[43–48] (Available at: http://tools.iedb.org/mhci/)

1.4.3 Binding to MHC class II prediction tool:

It's a similar tool to MHC I tool, after submitting the sequence, prediction method (NN-align algorithm used), selecting the source and locus, and MHC alleles. The result is then sorted by percentile rank and a cutoff of IC50 of 100.[49] (Available at: http://tools.iedb.org/mhcii/)

1.4.4 Population coverage:

T-cell recognizes a complex of major histocompatibility complex (MHC) molecules. This tool is used to test the ability of a peptide to elicit a response in immune system in specific populations carrying the desired MHC allele or worldwide. This tool calculates the approximate percentage of individuals' response to an epitope with regard to ethnicity and HLA genotypic frequencies and T-cell restriction. Population coverage is tested throughout the whole protein to MHC molecules class I, II and combined binding.[50] (Available at: http://tools.iedb.org/population/)

2.5. Molecular docking analysis

Molecular docking was performed using Moe 2007 Accessed July 22, 2019. The 3D structures of the promiscuous epitopes were predicted by PEP-FOLD.[51,52] The crystal structure of HLA-A*68:01 (**PDB ID 4hwz**) and HLA-DRB1*01:01(**PDB ID 5jlz**) were chosen as a model for molecular docking and were downloaded in a PDB format from the RCSB PDB resource. However, the selected crystal structures were in a complex form with ligands. Thus, to simplify the complex structure all water molecules, hetero groups and ligands were removed by Discovery Studio Visualizer 2.5.[53] Partial charge and energy minimization were applied for ligands and targets. In terms of the identification of the binding groove, the potential binding sites in the crystal structure were recognized using the Alpha Site Finder. Finally, ten independent docking runs were carried out for each Peptide. The results were retrieved as binding energies. Best poses for each epitope that displayed lowest binding energies were visualized using UCSF chimera 1.13.1 software.[54] (available at: https://www.chemcomp.com/, and http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/)

2.6.Physiochemical parameters:

Hence, the main function of a vaccine is to induce an immunogenic response once introduced to the immune system, it is essential to recognize the physiochemical parameters of the protein it's made of. This has been done using protoparam which has been used to give information about the vaccine of PK (molecular weight, theoretical pI, amino acid composition, estimated half-life, instability index, aliphatic index and grand

average of hydropathicity (GRAVY)), Protscale which provided hydrophobicity information and graphs (Both softwares are found in ExPasy server) And Bioedit software that provided the amino acid composition graph of PK protein [36,37,55] (Available at: https://web.expasy.org/protparam/, and https://web.expasy.org/protscale/)

2.7. Homology modelling:

Using **Chimera** (version 1.13.1rc) a tool powered by University of California, San Francisco, with support from NIH P41-GM103311. Chimera is used to visualize the 3-dimentional structure of the pyruvate kinase (PK) protein of C. albicans and the most promising peptides binding to MHC class I, MHC class II and peptides binding to MHC class I and II altogether in the study. (Figures 11,12 and 13)[56,57,58]

3. Results:

3.1. B-cell epitope prediction:-

The reference sequence of Pyruvate kinase *C. albicans* was analyzed using Bepipred Linear Epitope Prediction first, the average binder's score of the protein to B cell was 0.076, minimum was -0.002 and 1.916 for a maximum score, all values equal or greater than the default threshold 0.350 were potentially linear epitopes (table 1 and figure 2).

Table 1: Predicted antigenic B-cell epitopes, 12 antigenic sites were identified from PK *C. albicans*

Start	End	Peptide	Length
161	166	VDDEQT	6
393	398	AAYEQD	6
31	37	IGPKTNN	7
119	125	DAYKTKC	7
246	252	ENQQGVN	7
338	346	ETAKGNYPV	9
377	385	KKPTATTET	9
488	497	GWTRGSGHSN	10
450	461	KPSIENWQEDVE	12
182	195	GVNLPGTDVDLPAL	14
309	322	TYNPRPTRAEVSDV	14
88	110	DTKGPEIRTGTTIGDKDYPIPPN	23

In Emini surface accessibility prediction, for a potent B-cell epitope the average surface accessibility areas of the protein was scored as 1.000, with a maximum of 5.174 and a minimum of 0.073, all values equal or greater than the default threshold 1.000 were potentially in the surface. In addition, Kolaskar and Tongaonkar antigenicity prediction's average of antigenicity was 1.033, with a maximum of 1.206 and minimum of 0.897; all values equal to or greater than the default threshold 1.033 are potential antigenic determinants. The results of all conserved predicted B cell epitopes are shown in (table 2 and figure 2).

Table 2: B-cell epitopes prediction with their surface accessibility score and antigenicity score:

Peptide	Start	End	Length	EmInI surface score Th = 1.00	Kolaskar & And Tongaonkar Antigenicity Score Th = 1.033
DAYKTKC	119	125	7	1.545	1.039
AYKTKC	120	125	6	1.169	1.068
AKGNYPV	340	346	7	1.092	1.036
PTRAEVSDV*	314	322	9	1.126	1.045
YPIPPN	105	110	6	1.558	1.047

*Refer to the conserved peptides that passed the EminI surface accessibility and Kolaskar and Tongaonkar antigenicity test with their ideal length.
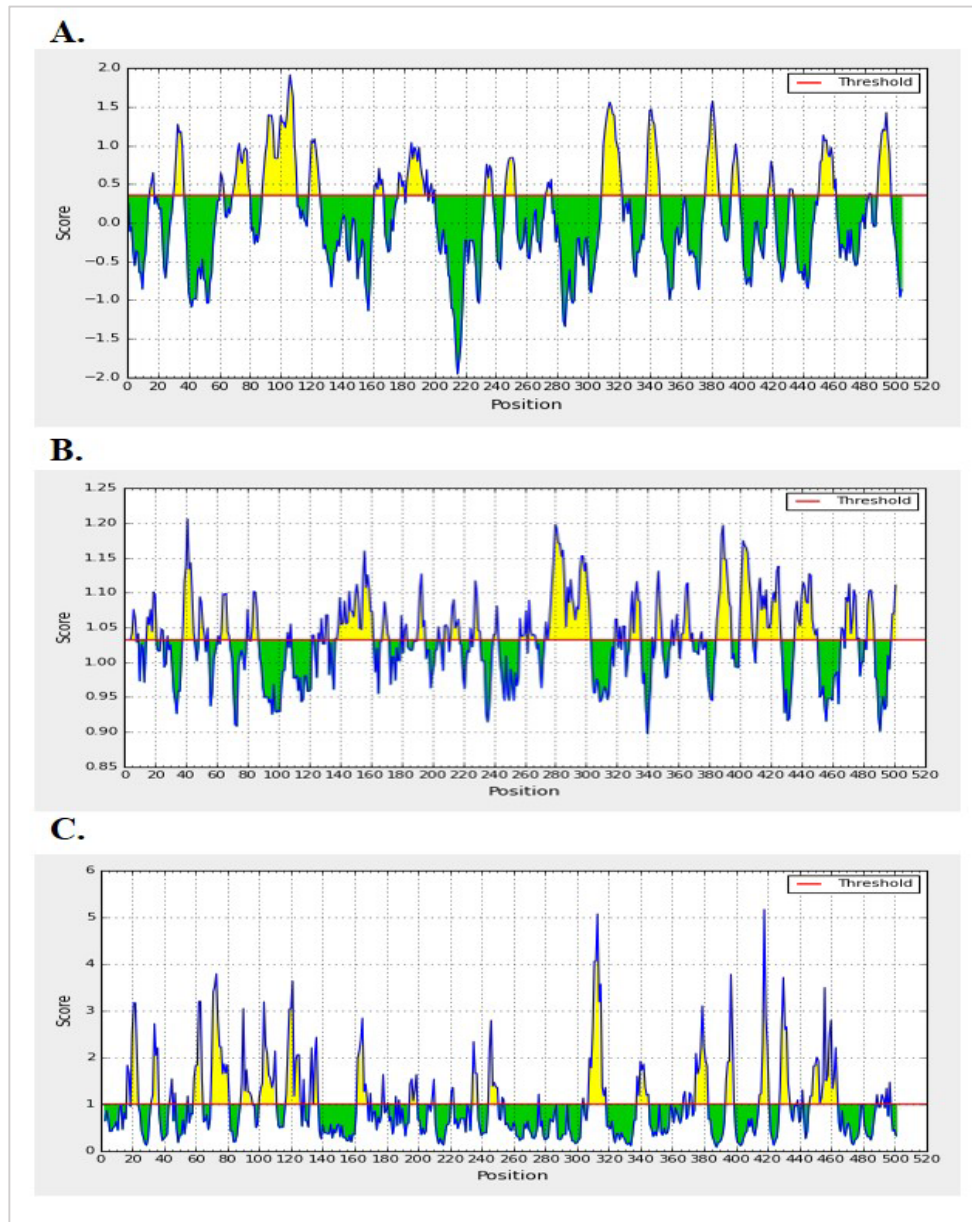
Figure 2: A. Bepipred Linear Epitope Prediction of PK *C. albicans*, the x-axis and y-axis represent the sequence position and linear probability, respectively. **B.** EminI surface accessibility

prediction of PK *C. albicans*, the x-axis and y-axis represent the sequence position and surface probability, respectively. **C.** Kolashkar and Tongaonkar antigenicity prediction of PK Candida albicans, the x-axis and y-axis represent the sequence position and antigenic propensity, respectively. Yellow peaks above the threshold (Red line) passed the test, Green peaks fall below threshold and did not pass the test.

Though five linear conserved surface antigenic epitopes passed all of the above tests, one epitope out of all was thought to be the promising B-cell epitope that is able to evoke B-lymphocyte for their efficient physiochemical properties and length (PTRAEVSDV). In addition, eighteen promising discontinuous epitopes were defined from the modeled protein after submission to ElliPro prediction tool. Epitopes were predicted to be located on the surface of the protein indicating quick recognition by host immune system. (Table 3, figure 3)

Table3: List of four predicted non-linear B-cell epitopes with highest number of Residues and their scores using ElliPro tool

| No.	Residues	No. of residues	Score
1	B:S4, B:S5, B:L6, B:S7, B:W8, B:L9, B:S10, B:N11, B:F12, B:N13, B:V14, B:E15, B:T16, B:V17, B:P18, B:S19, B:K20, B:K361, B:A362, B:I363, B:A364, B:P366, B:Q367, B:F369, B:N370, B:E371, B:R373, B:S374, B:L375, B:A376, B:K377, B:K378, B:P379, B:T380, B:A381, B:T382, B:T385, B:C386, B:V388, B:A389, B:S392, B:A393, B:A394, B:Y395, B:E396, B:Q397, B:D398, B:A399, B:K400, B:Y418, B:V422, B:R429, B:Y448, B:D449, B:K450, B:P451, B:S452, B:I453, B:E454, B:N455, B:Q457, B:E458, B:D459, B:V460, B:E461, B:N462, B:R463, B:L464, B:R465, B:W466, B:V468, B:S469, B:E470, B:V472, B:E473, B:L474, B:G475, B:I476, B:I477, B:S478, B:K479, B:G480, B:D481, B:S482, B:I483, B:T485, B:Q487, B:G488, B:W489, B:T490, B:R491, B:G492, B:S493, B:G494, B:H495, B:S496, B:N497, B:T498, B:V499, B:R500, B:I501	101	0.758
2	B:T96, B:G97, B:T98, B:T99, B:I100, B:G101, B:D102, B:K103, B:D104, B:Y105, B:P106, B:I107, B:P108, B:P109, B:N110, B:H111, B:E112, B:M113, B:I114, B:F115, B:T116, B:T117, B:D118, B:D119, B:A120, B:Y121, B:K122, B:T123, B:K124, B:C125, B:D126, B:D127, B:K128, B:V129, B:M130, B:Y131, B:I132, B:Y134, B:K135, B:N136, B:I137, B:T138, B:K139, B:V140, B:I141, B:A142, B:P143, B:G144, B:K145, B:I146, B:V153, B:L154, B:S155, B:F156, B:E157, B:V158, B:I159, B:S160, B:V161, B:D162, B:D163, B:E164, B:Q165, B:T166, B:L167, B:K168, B:V169, B:R170, B:S171, B:L172, B:N173, B:A174, B:G175, B:K176, B:I177, B:S178, B:S179, B:H180, B:K181, B:L185, B:P186, B:G187, B:T188, B:D189, B:V190, B:D191, B:L192	87	0.755
3	D:S4, D:S5, D:L6, D:S7, D:W8, D:L9, D:S10, D:N11, D:F12, D:N13, D:V14, D:E15, D:T16, D:V17, D:P18, D:S19, D:K20, D:K361, D:A362, D:I363, D:A364, D:P366, D:Q367, D:F369, D:N370, D:E371, D:R373, D:S374, D:L375, D:A376, D:K377, D:K378, D:P379, D:T380, D:A381, D:T382, D:T385, D:C386, D:V388, D:A389, D:S392, D:A393, D:A394, D:Y395, D:E396, D:Q397, D:D398, D:A399, D:K400, D:Y418, D:V422, D:R429, D:Y448, D:D449, D:K450, D:P451, D:S452, D:I453, D:E454, D:N455, D:Q457, D:E458, D:D459, D:V460, D:E461, D:N462, D:R463, D:L464, D:R465, D:W466, D:V468, D:S469, D:E470, D:V472, D:E473, D:L474, D:G475, D:I476, D:I477, D:S478, D:K479, D:G480, D:D481, D:S482, D:I483, D:T485, D:Q487, D:G488, D:W489, D:T490, D:R491, D:G492, D:S493, D:G494, D:H495, D:S496, D:N497, D:T498, D:V499, D:R500, D:I501	101	0.758
4	D:T96, D:G97, D:T98, D:T99, D:I100, D:G101, D:D102, D:K103, D:D104, D:Y105, D:P106, D:I107, D:P108, D:P109, D:N110, D:H111, D:E112, D:M113, D:I114, D:F115, D:T116, D:T117, D:D118, D:D119, D:A120, D:Y121, D:K122, D:T123, D:K124, D:C125, D:D126, D:D127, D:K128, D:V129, D:M130, D:Y131, D:I132, D:Y134, D:K135, D:N136, D:I137, D:T138, D:K139, D:V140, D:I141, D:A142, D:P143, D:G144, D:K145, D:I146, D:V153, D:L154, D:S155, D:F156, D:E157, D:V158, D:I159, D:S160, D:V161, D:D162, D:D163, D:E164, D:Q165, D:T166, D:L167, D:K168, D:V169, D:R170, D:S171, D:L172, D:N173, D:A174, D:G175, D:K176, D:I177, D:S178, D:S179, D:H180, D:K181, D:L185, D:P186, D:G187, D:T188, D:D189, D:V190, D:D191, D:L192	87	0.755
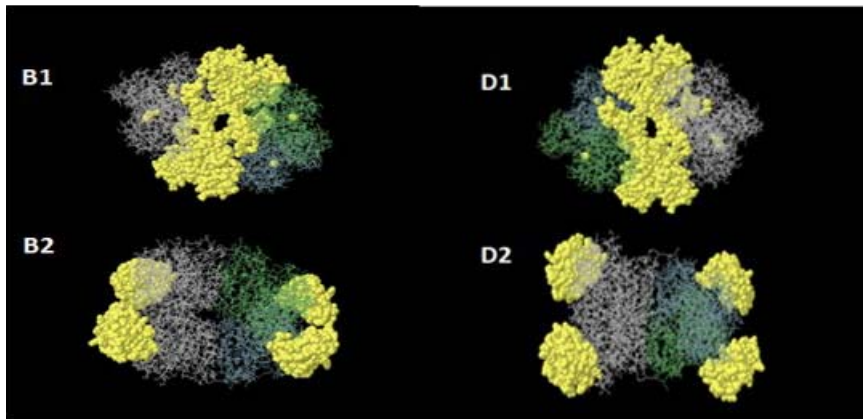
Figure 3: Three-dimensional representation of four non-linear epitopes (B1–B2, D1-D2) of the highly immunogenic pyruvate Kinase PK protein of *C. albicans* using ElliPro prediction tool. The clustered epitopes are highlighted in yellow, and the bulk of PK protein is depicted in grey.

3.2. T-cell peptide predictions

3.2.1. Prediction of MHC-I binding profile for T cytotoxic cell conserved epitopes:

167 epitopes were anticipated to interact with different MHC-1 alleles. The core epitopes (HLYRGVYPF/ RAAKFSHLY) were noticed to be the dominant binders with 10 alleles for each (HLA-A*02:06,HLA-A*23:01, HLA-A*29:02,HLA-A*32:01, HLA-B*15:01,HLA-B*15:02, HLA-B*35:01, HLA-C*07:02, HLA-C*12:03, HLA-C*14:02, HLA-A*11:01, HLA-A*29:02, HLA-A*30:02, HLA-A*68:01, HLA-B*15:01, HLA-B*35:01, HLA-B*57:01, HLA-B*58:01, HLA-C*12:03, HLA-C*15:02), followed by YVDDGVLSF which binds with nine alleles, AVAAVSAAY binding with eight alleles and YRGVYPFIY that's believed to bind with five alleles. (Table 4).

Table 4: List of most promising epitopes that had a good binding affinity with MHC-I alleles in terms of IC50 and Percentile rank.

Peptide	Start	End	No. of MHC-I allele hits
AVAAVSAAY*	387	395	8
HLYRGVYPF*	438	446	10
RAAKFSHLY*	432	440	10
YRGVYPFIY*	440	448	5
YVDDGVLSF*	148	156	9

*Top promising epitopes with efficient binding affinity and massive global population coverage.

3.2.2. Prediction of MHC-II binding profile for T helper cell conserved epitopes:

125 conserved predicted epitopes were found to interact with MHC-II alleles. The core epitopes (HMIFASFIR) is thought to be the top binder as it interacts with nine alleles; (HLA-DPA1*01, HLA-DPB1*04:01, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPA1*02:01, HLA-DPB1*01:01, HLA-DPA1*02:01, HLA-DPB1*05:01, HLA-DRB5*01:01). Followed by

LRWAVSEAV which binds to seven alleles and (IAYPQLFNE, IFASFIRTA, VFVVQKQLI) binding to six alleles for each. (Table 5).

Table 5: List of the five promising core sequence epitopes that had a strong binding affinity with MHC-II in terms of IC50 and Percentile Ranks

peptide Sequence	No. of MHC-II allele hits
HMIFASFIR	9
IAYPQLFNE	6
IFASFIRTA	6
VFVVQKQLI	6
LRWAVSEAV	7

3.2.3. Molecular docking analysis:

T-cell epitopes were tested to visualize the binding affinity between promising epitopes and HLA molecules. Global energy is the energy required to estimate the strength of association between the epitope within the active cleft of MHC molecule; more negative value indicates favored and stable binding of the complex. (Table 6, Figure 4, and 5)

Table 6. Docking results of the most promiscuous epitopes that show the best binding affinity

Epitope	Binding MHC molecule	Binding Energy (*ΔG kcal/mol)
AVAAVSAAY	HLA-A*68:01	-26.2570
HLYRGVYPF	HLA-A*68:01	-14.2184
RAAKFSHLY	HLA-A*68:01	-27.2360
YRGVYPFIY	HLA-A*68:01	-25.1326
YVDDGVLSF	HLA-A*68:01	-17.2499
HMIFASFIR	HLA-DRB1*01:01	-9.3566
IAYPQLFNE	HLA-DRB1*01:01	-6.9262
IFASFIRTA	HLA-DRB1*01:01	-20.9974
LRWAVSEAV	HLA-DRB1*01:01	-18.8068
VFVVQKQLI	HLA-DRB1*01:01	-10.6033
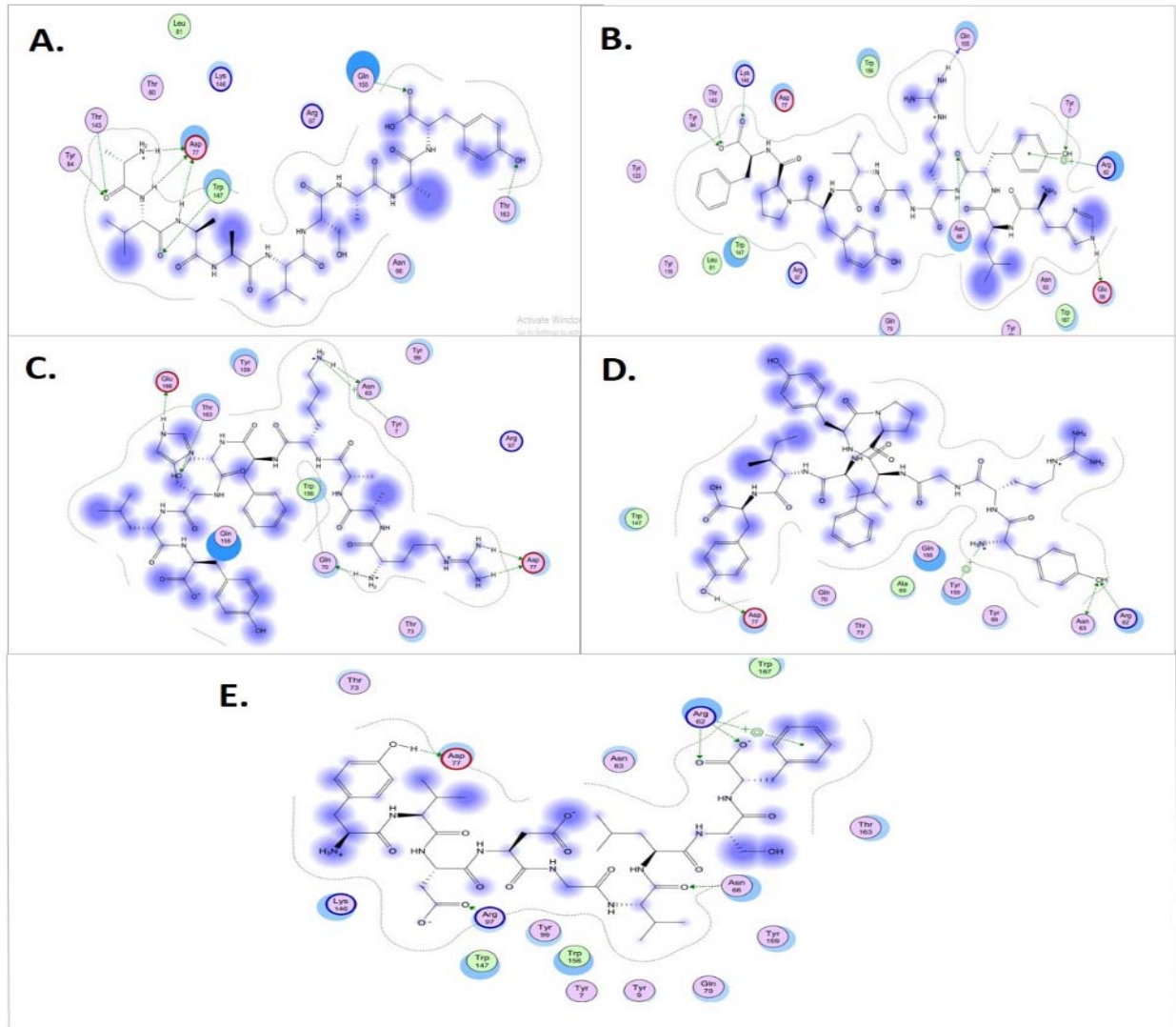
Figure 4. Illustrates 2-dimentional interaction of the best docking poses with HLA-A*68:01 MHC-I allele for five promising peptides **A.** AVAAVSAAY **B.** HLYRGVYPF **C.** RAAKFSHLY **D.** YRGVYPFIY **and E.** YVDDGVLSF.
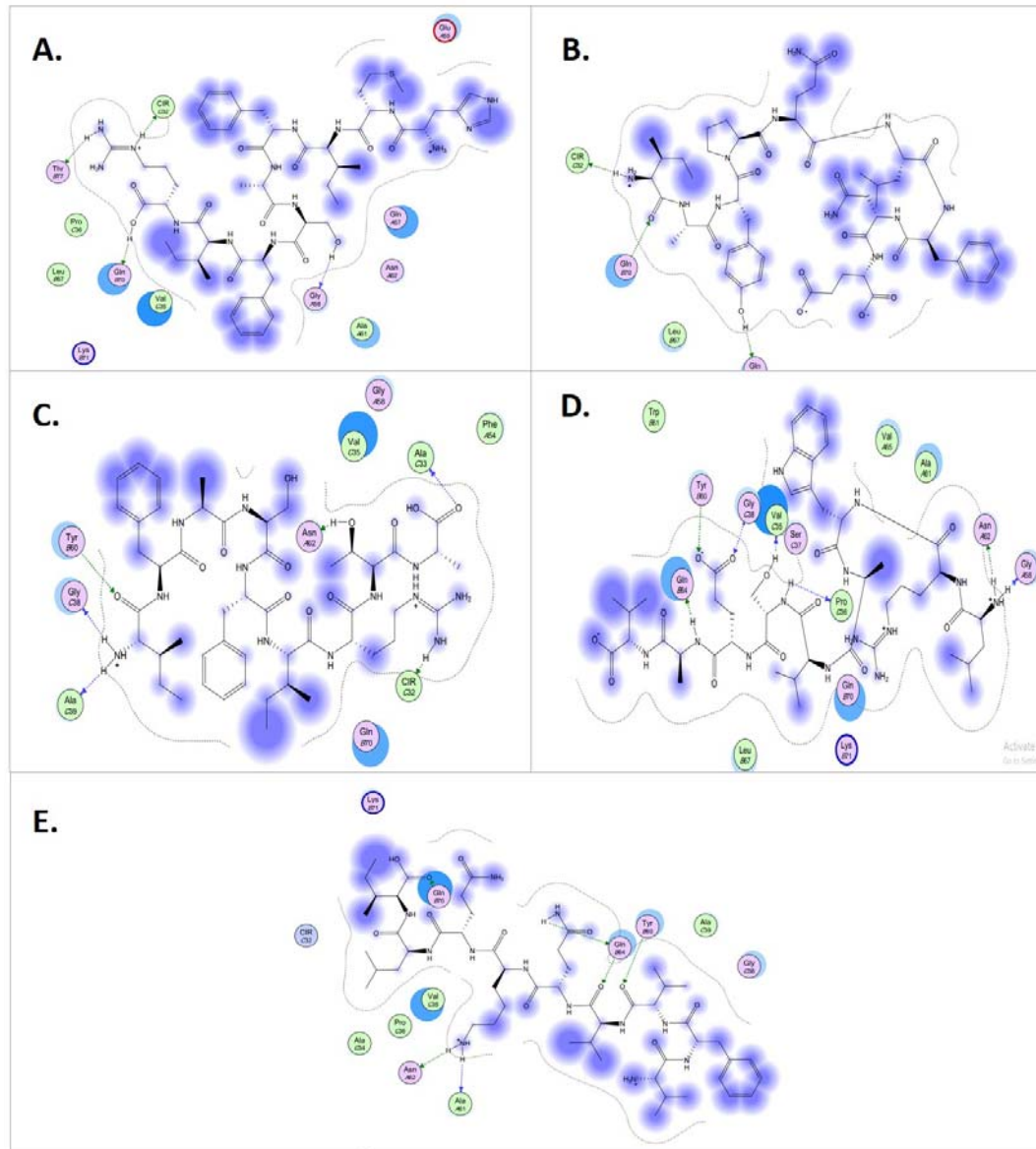
Figure 5. Illustrates 2-dimentional interaction of the best docking poses with HLA-DRB1*01:01 MHC-II allele for five promising peptides **A.** HMIFASFIR **B.** IAYPQLFNE **C.** IFASFIRTA **D.** LRWAVSEAV **and E.** VFVVQKQLI

3.4 Physiochemical parameters:

The length of the protein was found to be 504 amino acids. Molecular weight and Theoretical pI parameters were calculated as 55071.50KDa and 7.17, respectively. The pI value reflects that the protein nature is acidic. The total numbers of negatively and positively charged residues estimated as 65 and 65, correspondingly. Extinction coefficient of protein at 280 nm was measured 25035 M-1 cm-1 in water. The estimated half-life of vaccine was; 30 hours in (mammalian reticulocytes, in vitro),>20 hours in (yeast, in vivo) and >10 hours in (Escherichia coli, in vivo). The Instability index was computed to be 27.29, which indicates the thermo-stability of protein. The aliphatic index and the grand average of hydropathicity (GRAVY) value of vaccine were determined 91.19 and (-0.151), respectively. The high aliphatic index shows that

the protein is stable in a wide range of temperatures, and the negative GRAVY value shows protein hydrophilicity and a better interaction with the surrounding water molecules. (Illustrated in table 7, figure 5, 6).

Table 7. List of amino acid that formed pyruvate kinase protein, their sequential location, and molar percentage (Mol %)

AA	Number	Mol%	AA	Number	Mol%
Ala (A)	54	10.7	Leu (L)	37	7.3
Arg (R)	29	5.8	Lys (K)	36	7.1
Asn (N)	16	3.2	Met (M)	15	3
Asp (D)	28	5.6	Phe (F)	16	3.2
Cys (C)	10	2	Pro (P)	21	4.2
Gln (Q)	15	3	Ser (S)	27	5.4
Glu (E)	37	7.3	Thr (T)	26	5.2
Gly (G)	37	7.3	Trp (W)	2	0.4
His (H)	12	2.4	Tyr (Y)	9	1.8
Ile (I)	38	7.5	Val (V)	39	7.7
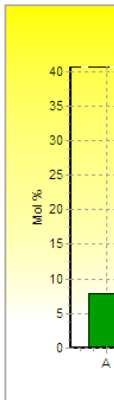
Figure 6. A graph illustrating the Amino Acid Composition (Mol %) of PK protein using BioEdit sequence alignment tool (Version 7.2.5.)
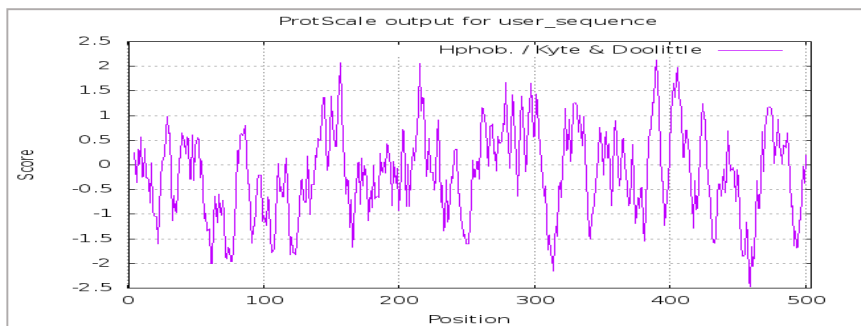
Figure 7. A graph showing the Hydrophobicity scale of PK protein using ProtScale server.

3.5. Population coverage analysis:-

The most interesting findings in this test is that the population coverage analysis result for most common binders to MHC-I and MHC-II alleles per each and combined; exhibit an exceptional coverage with percentages of 93.11%, 99.7%, and 99.7%. Respectively.

3.5.1 Population coverage for isolated MHC-I:

Five epitopes are given to interact with the most frequent MHC class I alleles (HLYRGVYPF, RAAKFSHLY, YVDDGVLSF, AVAAVSAAY and YRGVYPFIY), representing a considerable coverage against the whole world population. The maximum population coverage percentage over these epitopes worldwide was found to be 54.79% for YRGVYPFIY (table 8, figure 7).

Table 8. Global Population coverage of promising epitopes binding to MHC class I alleles

Core peptides	Coverage world class I	Total HLA alleles hits
YRGVYPFIY*	54.79%	5
HLYRGVYPF*	54.01%	10
YVDDGVLSF*	53.27%	9
AVAAVSAAY*	51.84%	8
RAAKFSHLY*	51.64%	10

MHC class	Coverage	Average hit	PC90
I	93.11%	3.35	1.19
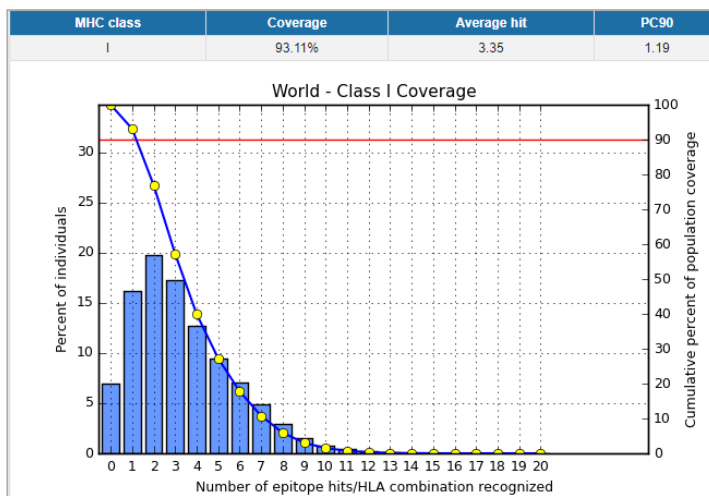
Figure 8. Illustrates the global coverage for the top five MHC-I peptides (HLYRGVYPF, RAAKFSHLY, YVDDGVLSF, AVAAVSAAY and YRGVYPFIY). Note: In the graph, the line (-o-) represents the cumulative percentage of population coverage of the epitopes; the bars represent the population coverage for each epitope.

3.5.2 Population coverage for isolated MHC-II:

In the case of MHC class II, five epitopes were assumed to interact with the most frequent MHC class II alleles (HMIFASFIR, LRWAVSEAV, IAYPQLFNE, IFASFIRTA and VFVVQKQLI), inferring a massive global coverage. The highest population coverage percentage of these epitopes worldwide was that of HMIFASFIR with percentage of 95.03% (table 9, figure 8).

Table 9. Global Population coverage of promising epitopes in isolated MHC class II

Core peptides	Coverage world MHC II	Total HLA hits
HMIFASFIR	93.53%	9
IFASFIRTA	90.79%	6
IAYPQLFNE	89.92%	6

| VFVVQKQLI | 88.99% | 6 |
| LRWAVSEAV | 79.58% | 7 |

MHC class	Coverage	Average hit	PC90
II	99.7%	7.77	4.65
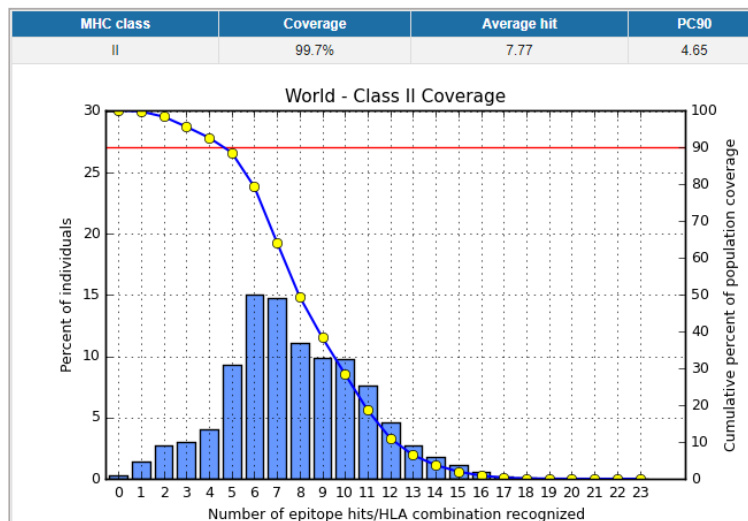
Figure 9. Illustrates the global proportion for the top five MHC-II epitopes (HMIFASFIR, LRWAVSEAV, IAYPQLFNE, IFASFIRTA and VFVVQKQLI). Notes: In the graph, the line (-o-) represents the cumulative percentage of population coverage of the epitopes; the bars represent the population coverage for each epitope.

3.5.3 Population coverage for MHC-I and MHC-II alleles combined

Regarding MHC-I and MHC-II alleles combined, three epitopes were supposed to interact with most predominant MHC class I and MHC class II alleles (HMIFASFIR, TETCAVAAV and LRWAVSEAV); represent a significant global coverage by IEDB population coverage tool. The most common population coverage percentage of these epitopes in the World was granted to HMIFASFIR with percentage of 97.02%. The remarkable finding in this test is the average for the most common binders to combined MHC-I and MHC-II alleles; reveal an outstanding coverage with percentage of 99.75%.Table 10, Figure 9

Table 10. Global population coverage of the ten promising epitopes in MHC class I and II combined mode.

Core peptides	Global Coverage MHC classes I-II Combined	Total HLA hits
HMIFASFIR	97.02%	13
TETCAVAAV	80.74%	8
LRWAVSEAV	80.56%	8
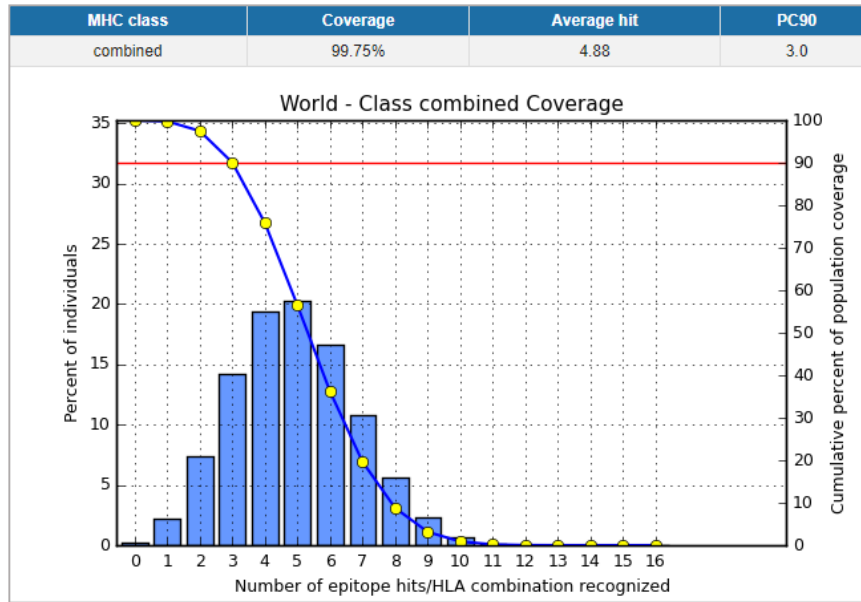
Figure 10. Illustrates the global population proportion for the top three MHC-I & II epitopes in combined mode (HMIFASFIR, TETCAVAAV and LRWAVSEAV). Notes: In the graphs, the line (-o-) represents the cumulative percentage of population coverage of the epitopes; the bars represent the population coverage for each epitope.
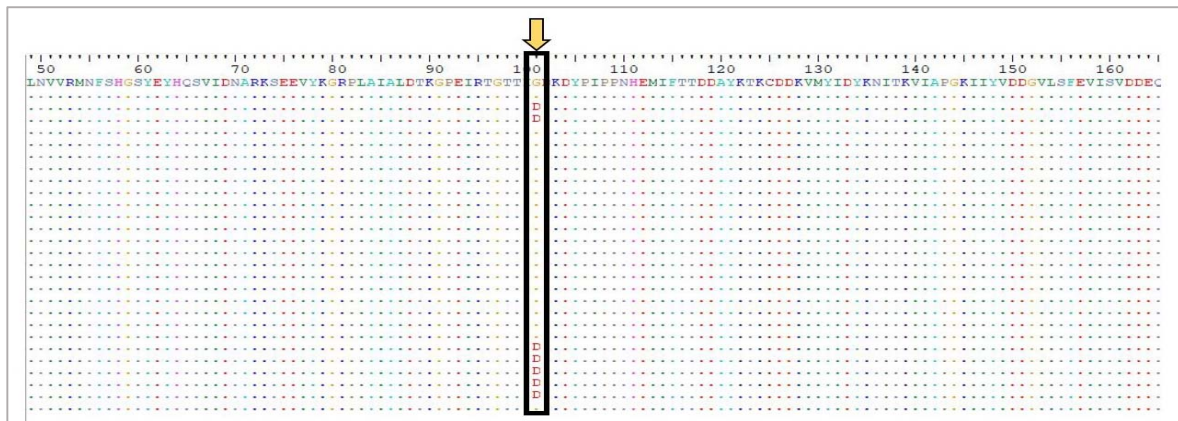
Figure 11. A fragment of the multiple sequence alignment of pyruvate kinase protein of *C. albicans* showing the only area of non-conservation in some species at position 101, where Glycine mutated into Aspartic acid, using Bioedit software.
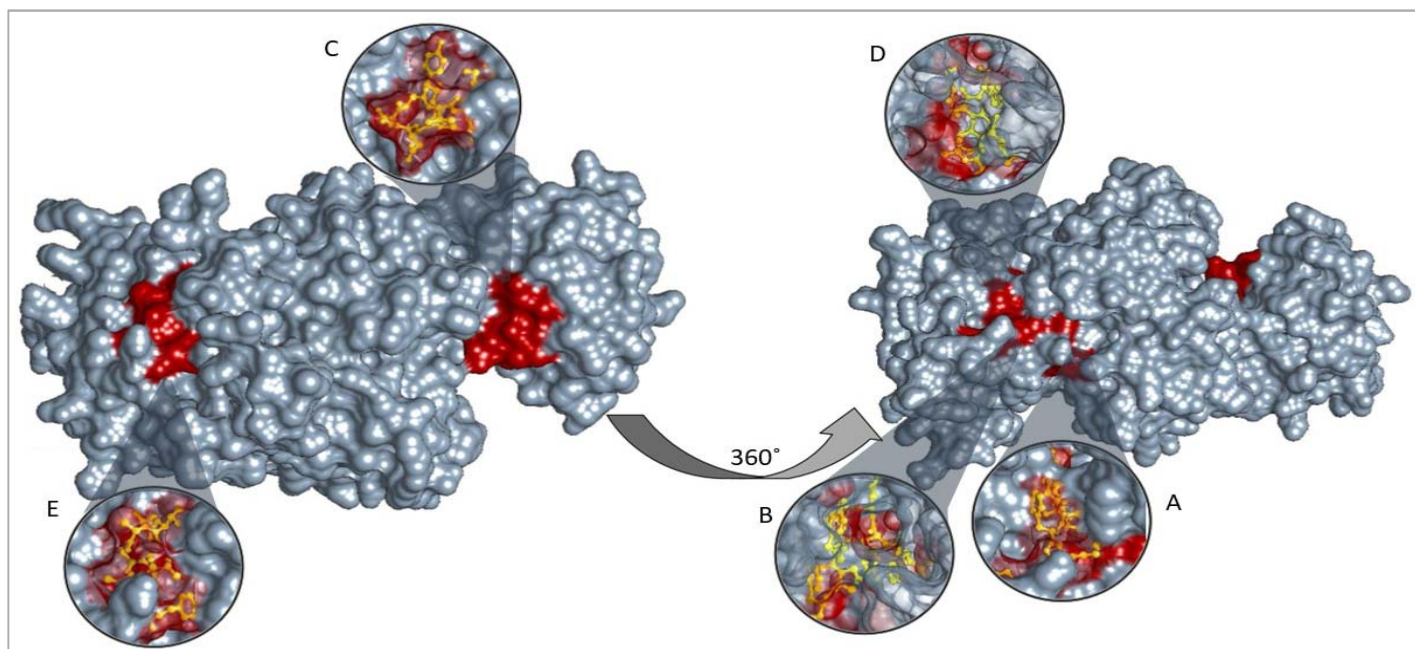
Figure 12. 3D structure of pyruvate kinase protein visualizing top five T-cell peptides binding to MHC class I **A.** YRGVYPFIY **B.** HLYRGVYPF **C.** YVDDGVLSF **D.** RAAKFSHLY and **E.** AVAAVSAAY using chimera (version 1.13.1rc)
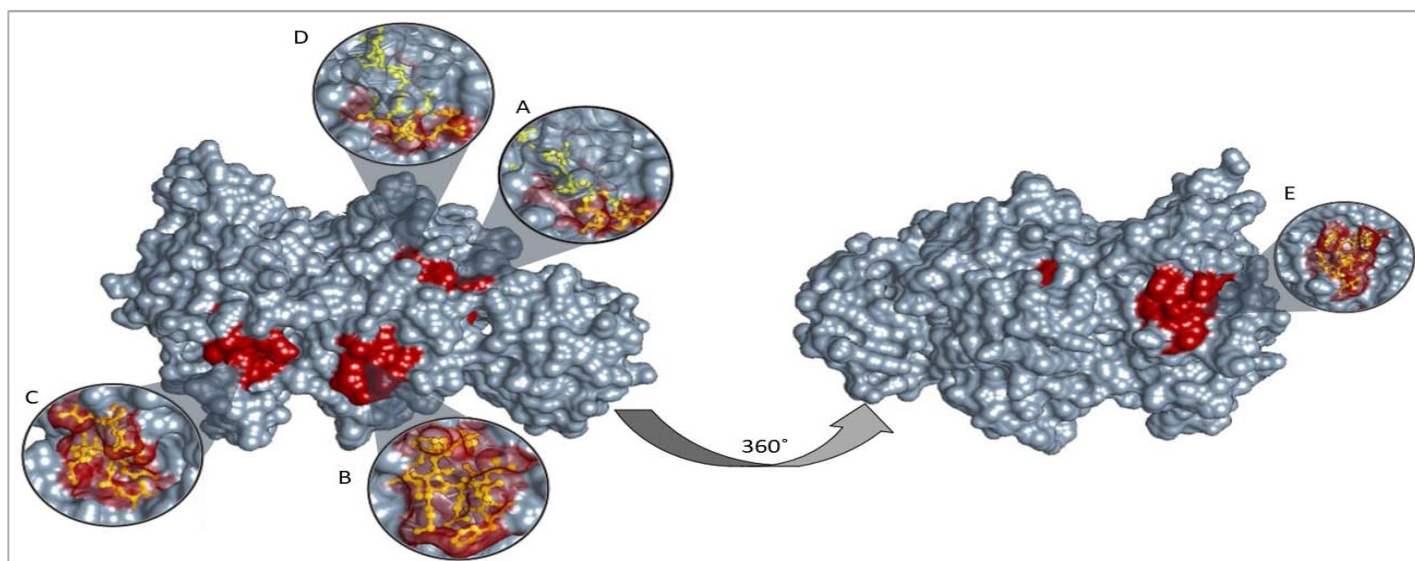
Figure 13. 3D structure of pyruvate kinase protein visualizing top five T-cell peptides binding to MHC class II **A.** HMIFASFIR **B.** VFVVQKQLI **C.** IAYPQLFNE **D.** IFASFIRTA **and E.** LRWAVSEAV, using chimera (version 1.13.1rc)
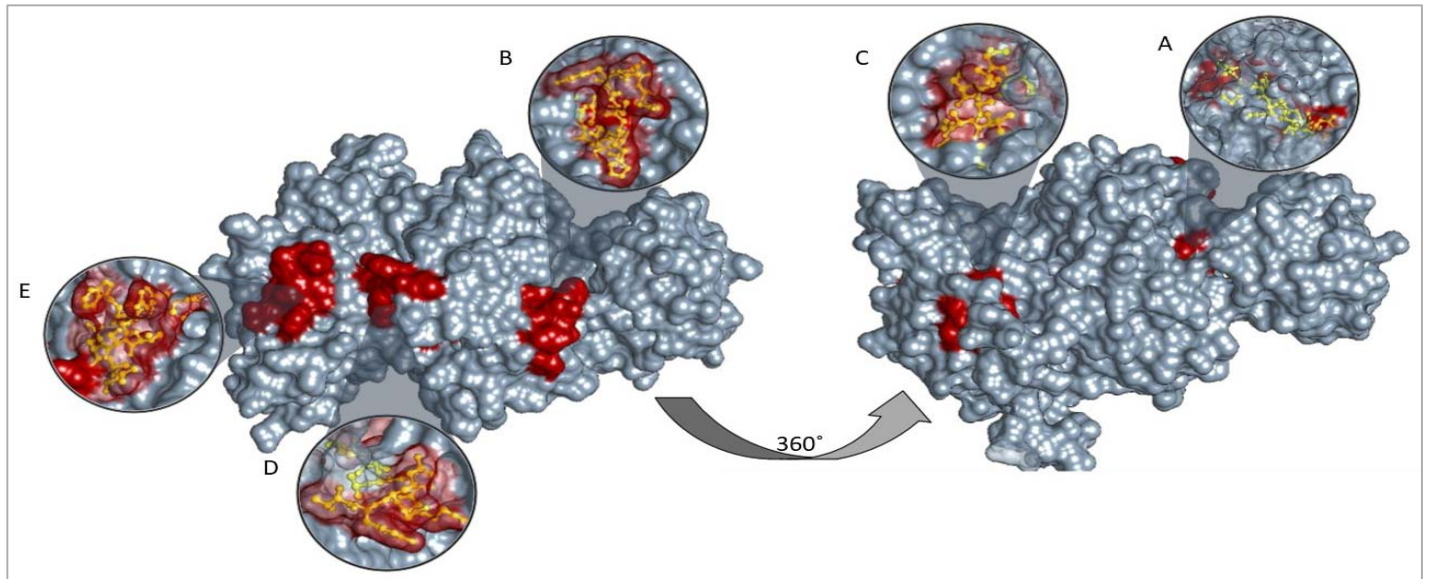
Figure 14. 3D structure of pyruvate kinase protein visualizing top five T-cell peptides binding to MHC class I and II altogether **A.** HMIFASFIR **B.** NFSHGSYEY **C.** TETCAVAAV **D.** VYKGRPLAI **and E.** LRWAVSEAV, using chimera (version 1.13.1rc)

4. Discussion:

This study with a major focus on T-cell epitopes has revealed fifteen promising T-cell peptides to make a peptide based vaccine for *C. albicans* using pyruvate kinase as a target protein to induce an immune response. The peptide HMIFASFIR is one of the most important findings of this study regarding its interaction with both MHC-II alleles and MHC-I and MHC -II alleles combined. Giving significantly high worldwide population coverage percentages of 95.03% and 97.03% respectively, ideal length of nine amino acids, an IC50 of 8.2 when binding to MHC allele HLA-DRB5*01:01 indicating strong interaction between the two and a binding energy of -9.356 (kcal/mol) when binding to HLA DRB1*01:01. In addition, We suggest PTRAEVSDV peptide is the most promising linear B-cell peptide due to its physiochemical parameters and optimal length (nine amino acids).

Five T-cell peptides were suggested to be the most promising ones in accordance to their global population coverage, binding energies to most frequent MHC-I allele in the study (HLA A*68:01) and their strength of interaction to other alleles with the IC50 scores. These epitopes could potentially induce CD8+ cytotoxic T- cell immune response when interacting strongly with MHC class I alleles. YRGVYPFIY, HLYRGVYPF, YVDDGVLSF, RAAKFSHLY, and AVAAVSAAY (having global population coverage of 54.79%, 54.01%, 53.27%, 51.84%, 41.64%, respectively) all five peptides had a massive global coverage of 93.11% when combined.

YRGVYPFIY peptide giving the highest global coverage with a percentage of 54.79% of world population, a very low binding energy to HLA-A*68:01 of -25.1326 (Kcal/mol) and a strong interaction with HLA-C*07:02 with estimated IC50 of 35.88 making it one of the strongest T-cell candidates. Another strong peptide candidate is AVAAVSAAY peptide which had the best binding affinity of all five peptides to HLA-A*68:01 which is the most frequent MHC-I allele of -26.2570 (Kcal/mol) and a strong interaction with HLA-B*35:01 with an estimated IC50 of 12.76. HLYRGVYPF, RAAKFSHLY peptides which bind most to MHC class I alleles (10 alleles each).

As for MHC class II five peptides were selected as most potential peptides to actvate T-helper cell (CD4+) according to their population coverage globally, binding energy with the most frequent MHC-II allele HLA DRB1*01:01 and strength of interaction by the IC50 scores, HMIFASFIR, VFVVQKQLI, IAYPQLFNE, IFASFIRTA, and LRWAVSEAV (Global coverage of 95.03%, 91.03%, 93.63%, 92.82%, 79.58%, respectively). With an ideal total coverage of 99.70% of world population.

HMIFASFIR being one of the highly recommended peptides for giving the highest population coverage, binding affinity to HLA DRB1*01:01 of -9.356 (Kcal/mol) and IC50 (which indicated the strength of the interaction between the peptide and selected MHC allele) of 8.2 with HLA DRB5*01:01. LRWAVSEAV peptide had highest binding number to MHC II alleles (six alleles), a relatively high global population coverage (79.58%). Besides that, the peptide had a low binding energy required to bind to HLA-DRB1*01:01 (-18.8068 (Kcal/mol)) and strong interaction with HLA-DRB1*01:01 with and IC50 of 21. Another strong candidate MHC-II peptide is IFASFIRTA which shares few peptides with HMIFASFIR. The peptide had a high global population coverage estimated with 90.79%, binds to six MHC-II alleles, had a strong interaction with HLA-DQA1*01:02/DQB1*06:02 with an IC50 of 30.5, and gave the best docking result to HLA DRB1*01:01 with a binding energy of (-20.997 Kcal/mol)

Peptides binding to MHC class I and II alleles Combined were 61 peptides. The most important five peptides had the highest global population coverage percentage, HMIFASFIR, NFSHGSYEY, TETCAVAAV, VYKGRPLAI, and LRWAVSEAV (97.02%, 79.43%, 80.74%, 79.89%, and 80.56%, respectively). The five peptides altogether provided a total coverage of 99.85%.

Pyruvate kinase PK was selected for this study because it has proven to be a highly immunogenic protein in several organisms. A study by de Klerk N *et* al., 2012 has isolated PK and Frucose-bisphosphate Aldolase FBA as immunogens using flowcytometry from Madurella mycetomatis which causes Eumycetoma and it was expressed on its hyphae.[59] A more recent study regarding M. Mycetomatis by Manofali A et al., 2018 confirmed used PK as an immunogenic for peptide-based vaccine design giving a moderately high global population coverage percentages in conclusion. Another study suggested, using N-terminal sequencing technique that Mycoplasma Synoviae and Mycoplasma Gallisepticum (major poultry pathogens) have 14 immunogenic proteins, among which, pyruvate kinase PK enzyme was by Bercic R. et al., 2007.[60]

In addition, various studies suggested PK to be an immunogenic stimulant and a useful target for vaccine design of candida albicans particularly. However, none of which actually conducted a peptide based study for PK in C. albicans. A study by Pitarch A. et al., 1999, using Gel electrohpresis, followed by western blotting technique for patients with systemic candidiasis detected over 18 immunogenic proteins in C. albicans, one of these was glycolytic enzymes including PK .[61]Another study by Mercedes Pardo et al, 2000. Using nanoelectrospray tandem mass spectrometry tech, suggested that C. albicans PK is an immunogenic protein.[62] Other paper by swoboda et al. 1993, have suggested that pyruvate kinase protein gives strong interactions with IgM and IgG antibodies for candida albicans using pooled sera and concluded it being a proper antigen for candida albicans.[63]

This study was limited by being strictly computational, additionally, during analysis Calculations of some HLA class II alleles /epitopes (HLA DRB3*01:01, HLA DRB 5*01:01, HLA DPA 1*01,

and HLA DRB 4*01:01) for population coverage were missed, that may influence in the accuracy of the population coverage results.

Many recent studies has proved the efficiency of peptide- based vaccines, moreover, few of them even made it to the clinical trials. A study by Matsumoto et al. 2016, reported a phase I study of a personalized peptide vaccine (PPV) for advanced urothelial carcinoma patients who failed treatment with MVAC. Ten patients with MVAC□refractory advanced or metastatic urothelial cancer were treated twelve times using positive peptides chosen from 14 and 16 peptides in patients with HLA A24 and A2, respectively. The peptide vaccination was well tolerated with no major adverse effects and gave a notable rise in the IgG titer in patients.[64] Another study by Schwartzentruber et al. 2011, had shown that testing two groups of patients with advanced melanoma. One group treated with interluken-2 only, and the second group treated with vaccine-interluken-2. The study concluded that patients of the second group response to the treatment was significantly improved than those of Group one.[65]

Peptide-based vaccination is a key role of combining a specific, desirable immune response, easy production, less laboring, with minimal immunological hazards, combining both T- cell and B-cell epitopes and could be used for both prevention or as a therapeutic tool .[66,67] peptide vaccine is poorly immunogenic when used alone, thus, a substance that facilitates and enhance antigen-antibody binding (adjuvant) is needed for better immune response.[68] Future studies are required and highly recommended to subject these suggested peptides for clinical trials and select a suitable delivery system or adjuvant and strategies which could improve in-vivo and in-vitro T-cell and B-cell immune responses.

Conclusion:

Vaccine development for *C. albicans* is becoming a necessity due to, the growing resistance of the fungi to anti-fungal drugs, poor response to treatments and multiple recurrences to hospitalized patients. Peptide vaccines are well known for their efficiency, lack of hazard and less time consumption and that was the scope of this study. Fifteen peptides gave good population coverage percentages, in addition to strong interactions to MHC alleles of both classes I and II with varying degrees and poses of best possible binding energies to the most frequent peptides in the study HLA A*68:01 for MHC-I and HLA DRB1*01:01 or MHC-II. However, YRGVYPFIY, AVAAVSAAY, HMIFASFIR, LRWAVSEAV and IFASFIRTA T-cell peptides gave the best responses to the tests and therefore, are the most recommended peptides of this study for future studies in vitro and in vivo.

5. References:

1. Li W-R, Shi Q-S, Dai H-Q, et al. Antifungal activity, kinetics and molecular mechanism of action of garlic oil against Candida albicans. *Sci Rep*. 2016;6(1):22805. doi:10.1038/srep22805

2. Marc G, Araniciu C, Dafina Oniga S, et al. molecules New N-(oxazolylmethyl)-thiazolidinedione Active against Candida albicans Biofilm: Potential Als Proteins Inhibitors. 2018. doi:10.3390/molecules23102522

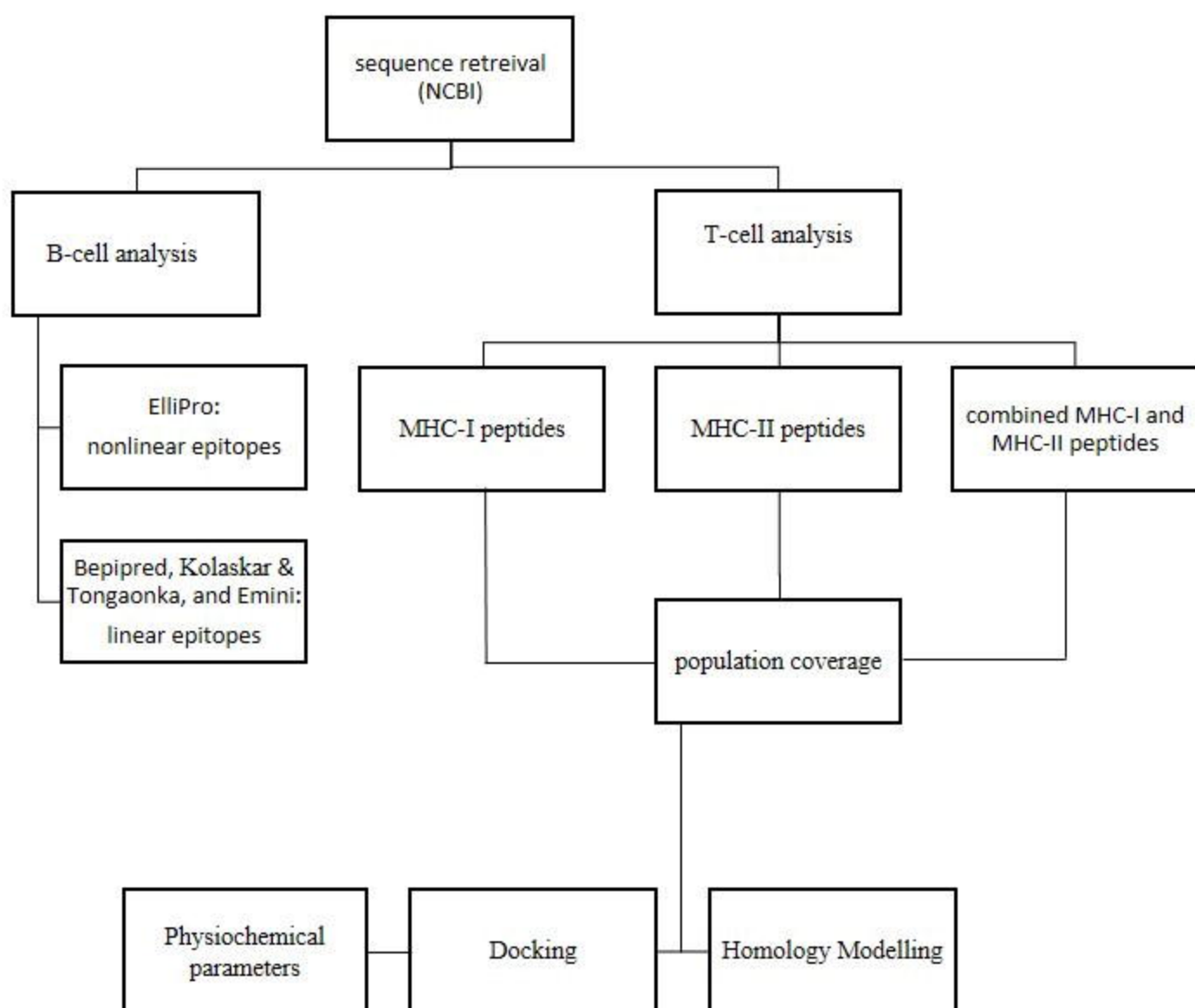3. Ensaf A Azzazy, Nahla E El-Gammal, Wafaa S Metwally MHA. Role of Candida

Albicans Germ Tube Antibody in Diagnosis of Invasive Candidiasis in End Stage Liver Disease Patients | Read by QxMD. *Egypt J Immunol*. 2018. https://read.qxmd.com/read/30243008/role-of-candida-albicans-germ-tube-antibody-in-diagnosis-of-invasive-candidiasis-in-end-stage-liver-disease-patients. Accessed May 8, 2019.

4. Lee J-H, Kim Y-G, Gupta VK, Manoharan RK, Lee J. Suppression of Fluconazole Resistant Candida albicans Biofilm Formation and Filamentation by Methylindole Derivatives. *Front Microbiol*. 2018;9:2641. doi:10.3389/fmicb.2018.02641

5. Li Y, Sun L, Lu C, Gong Y, Li M, Sun S. Promising Antifungal Targets Against Candida albicans Based on Ion Homeostasis. *Front Cell Infect Microbiol*. 2018;8:286. doi:10.3389/fcimb.2018.00286

6. Barelle CJ, Priest CL, MacCallum DM, Gow NAR, Odds FC, Brown AJP. Niche-specific regulation of central metabolic pathways in a fungal pathogen. *Cell Microbiol*. 2006;8(6):961-971. doi:10.1111/j.1462-5822.2005.00676.x

7. Jayaraja S, Dakhama A, Yun B, et al. Cytosolic phospholipase A2 contributes to innate immune defense against Candida albicans lung infection. *BMC Immunol*. 2016;17(1):27. doi:10.1186/s12865-016-0165-9

8. Lu Y, Su C, Ray S, Yuan Y, Liu H. CO2 Signaling through the Ptc2-Ssn3 Axis Governs Sustained Hyphal Development of Candida albicans by Reducing Ume6 Phosphorylation and Degradation. *MBio*. 2019;10(1):e02320-18. doi:10.1128/mBio.02320-18

9. McCarthy CGP, Fitzpatrick DA. Pan-genome analyses of model fungal species. *Microb Genomics*. 2019;5(2). doi:10.1099/mgen.0.000243

10. Masłyk M, Janeczko M, Demchuk OM, et al. A representative of arylcyanomethylenequinone oximes effectively inhibits growth and formation of hyphae in Candida albicans and influences the activity of protein kinases in vitro. *Saudi Pharm J*. 2018;26(2):244-252. doi:10.1016/J.JSPS.2017.12.004

11. Grahl N, Demers EG, Lindsay AK, et al. Mitochondrial Activity and Cyr1 Are Key Regulators of Ras1 Activation of C. albicans Virulence Pathways. Cowen LE, ed. *PLOS Pathog*. 2015;11(8):e1005133. doi:10.1371/journal.ppat.1005133

12. Mancuso R, Chinnici J, Tsou C, Busarajan S, Munnangi R, Maddi A. Functions of Candida albicans cell wall glycosidases Dfg5p and Dcw1p in biofilm formation and HOG MAPK pathway. *PeerJ*. 2018;6:e5685. doi:10.7717/peerj.5685

13. Mullick A, Leon Z, Min-Oo G, et al. Cardiac Failure in C5-Deficient A/J Mice after Candida albicans Infection. *Infect Immun*. 2006;74(8):4439-4451. doi:10.1128/IAI.00159-06

14. Barelle CJ, Priest CL, MacCallum DM, Gow NAR, Odds FC, Brown AJP. Niche-specific regulation of central metabolic pathways in a fungal pathogen. *Cell Microbiol*. 2006;8(6):961-971. doi:10.1111/j.1462-5822.2005.00676.x

15. Guo C, Li G, Hou J, et al. Tumor pyruvate kinase M2: A promising molecular target of gastrointestinal cancer. *Chinese J Cancer Res*. 2019;30(6):669-676. doi:10.21147/j.issn.1000-9604.2018.06.11

16. Israelsen WJ, Vander Heiden MG. Pyruvate kinase: Function, regulation and role in cancer. *Semin Cell Dev Biol*. 2015;43:43-51. doi:10.1016/j.semcdb.2015.08.004
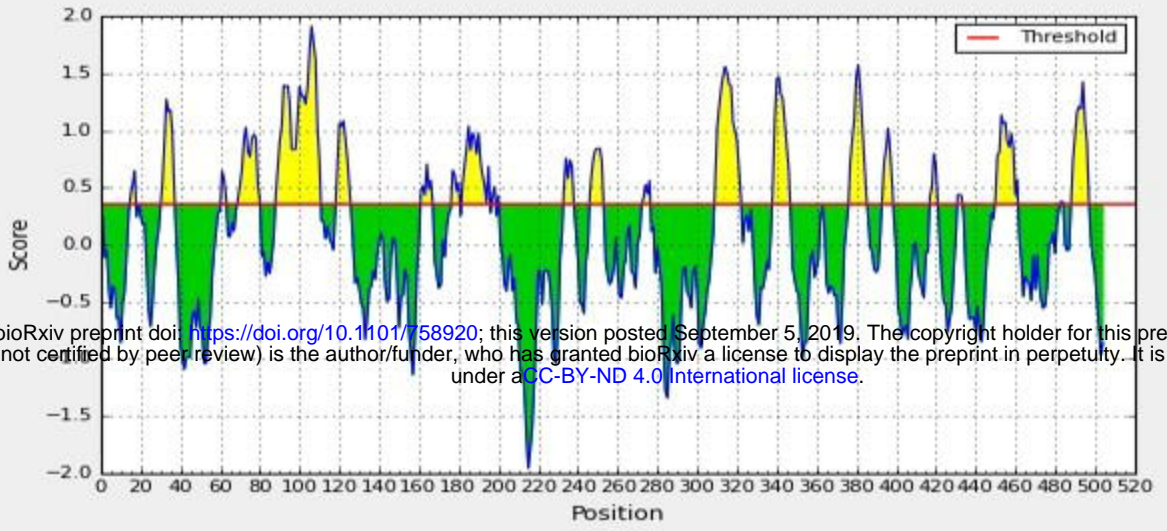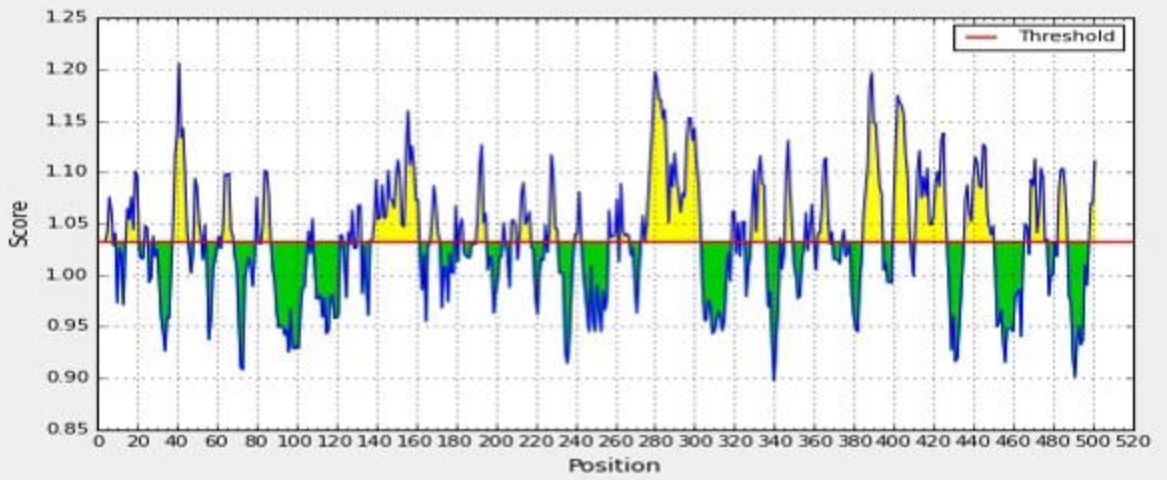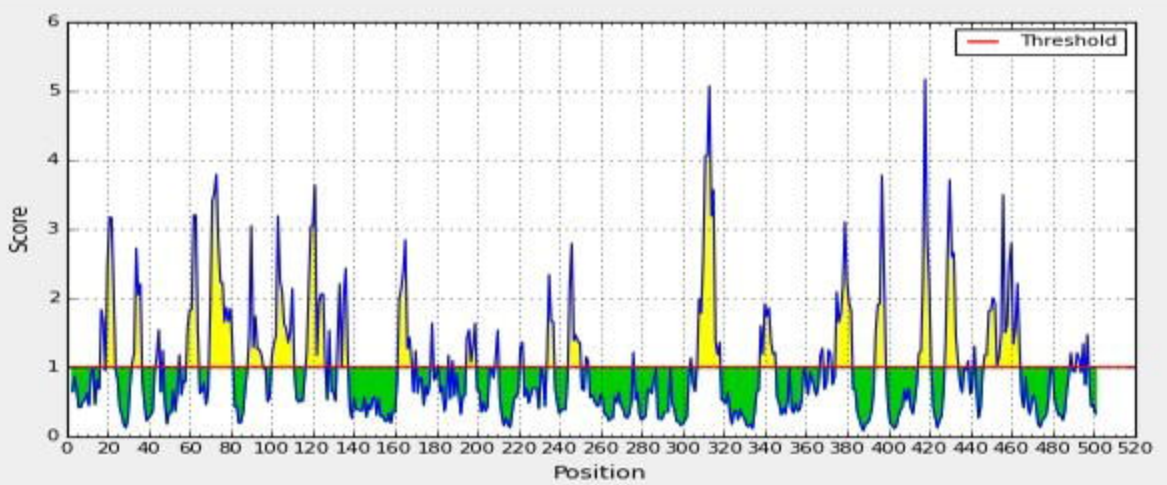
17. Kung C, Hixon J, Kosinski PA, et al. AG-348 enhances pyruvate kinase activity in red blood cells from patients with pyruvate kinase deficiency. *Blood*. 2017;130(11):1347-1356. doi:10.1182/blood-2016-11-753525

18. Gonçalves S, Silva PM, Felício MR, de Medeiros LN, Kurtenbach E, Santos NC. Psd1 Effects on Candida albicans Planktonic Cells and Biofilms. *Front Cell Infect Microbiol*. 2017;7. doi:10.3389/fcimb.2017.00249

19. Pandolfi F, D'Acierno F, Bortolami M, et al. Searching for new agents active against Candida albicans biofilm: A series of indole derivatives, design, synthesis and biological evaluation. *Eur J Med Chem*. 2019;165:93-106. doi:10.1016/J.EJMECH.2019.01.012

20. Weiland-Bräuer N, Malek I, Schmitz RA. Metagenomic quorum quenching enzymes affect biofilm formation of Candida albicans and Staphylococcus epidermidis. *PLoS One*. 2019;14(1):e0211366. doi:10.1371/JOURNAL.PONE.0211366

21. Alim D, Sircaik S, Panwar S, Alim D, Sircaik S, Panwar SL. The Significance of Lipids to Biofilm Formation in Candida albicans: An Emerging Perspective. *J Fungi*. 2018;4(4):140. doi:10.3390/jof4040140

22. Lagree K, Mon HH, Mitchell AP, Ducker WA. Impact of surface topography on biofilm formation by Candida albicans. Coste AT, ed. *PLoS One*. 2018;13(6):e0197925. doi:10.1371/journal.pone.0197925

23. Lopes JP, Stylianou M, Backman E, et al. Evasion of Immune Surveillance in Low Oxygen Environments Enhances Candida albicans Virulence. *MBio*. 2018;9(6):e02120-18. doi:10.1128/mBio.02120-18

24. Templeton S, Nett JE, Drummond R, Hoi Wan Tso G, Antonio Reales-Calderon J, Pavelka N. The elusive Anti-Candida vaccine: Lessons From the Past and Opportunities for the Future. 2018;9:1. doi:10.3389/fimmu.2018.00897

25. Templeton S, Nett JE, Drummond R, Hoi Wan Tso G, Antonio Reales-Calderon J, Pavelka N. The elusive Anti-Candida vaccine: Lessons From the Past and Opportunities for the Future. 2018;9:1. doi:10.3389/fimmu.2018.00897

26. Qin Y, Zhang L, Xu Z, et al. Innate immune cell response upon *Candida albicans* infection. *Virulence*. 2016;7(5):512-526. doi:10.1080/21505594.2016.1138201

27. Cárdenas-Freytag L, Steele C, Jr FLW, Cheng E, Clements JD, Jr PLF. Partial protection against experimental vaginal candidiasis after mucosal vaccination with heat-killed *Candida albicans* and the mucosal adjuvant LT(R192G). *Med Mycol*. 2002;40(3):291-299. doi:10.1080/mmy.40.3.291.299

28. Cárdenas-Freytag L, Steele C, Wormley FL, Cheng E, Clements JD, Fidel PL. Partial protection against experimental vaginal candidiasis after mucosal vaccination with heat-killed Candida albicans and the mucosal adjuvant LT(R192G). *Med Mycol*. 2002;40(3):291-299. http://www.ncbi.nlm.nih.gov/pubmed/12146759. Accessed May 8, 2019.

29. Qin Y, Zhang L, Xu Z, et al. Innate immune cell response upon *Candida albicans* infection. *Virulence*. 2016;7(5):512-526. doi:10.1080/21505594.2016.1138201

30. Segal E, Frenkel M. Experimental in Vivo Models of Candidiasis. *J fungi (Basel, Switzerland)*. 2018;4(1). doi:10.3390/jof4010021

31. Wang X, Sui X, Yan L, Wang Y, Cao Y, Jiang Y. Vaccines in the treatment of invasive candidiasis. *Virulence*. January 2015:1-7. doi:10.4161/21505594.2014.983015

32. Edwards JE, Schwartz MM, Schmidt CS, et al. A Fungal Immunotherapeutic Vaccine (NDV-3A) for Treatment of Recurrent Vulvovaginal Candidiasis—A Phase 2 Randomized, Double-Blind, Placebo-Controlled Trial. *Clin Infect Dis*. 2018;66(12):1928-1936. doi:10.1093/cid/ciy185

33. Khalili S, Jahangiri A, Borna H, Ahmadi Zanoos K, Amani J. Computational vaccinology and epitope vaccine design by immunoinformatics. *Acta Microbiol Immunol Hung*. 2014;61(3):285-307. doi:10.1556/AMicr.61.2014.3.4

34. Mohammed AA, ALnaby AM, Sabeel SM, et al. Epitope-Based Peptide Vaccine Against Fructose-Bisphosphate Aldolase of *Madurella mycetomatis* Using Immunoinformatics Approaches. *Bioinform Biol Insights*. 2018;12:117793221880970. doi:10.1177/1177932218809703

35. Chaudhuri R, Ramachandran S. Immunoinformatics as a Tool for New Antifungal Vaccines. In: *Methods in Molecular Biology (Clifton, N.J.)*. Vol 1625. ; 2017:31-43. doi:10.1007/978-1-4939-7104-6_3

36. N Altayb H, M El Amin N, M. Mukhtar M, Ahmed Salih M, A M Siddig M. Molecular Characterization and in Silico<i> </i>Analysis of a Novel Mutation in TEM-1 Beta-Lactamase Gene among Pathogenic <i>E. coli </i>infecting a Sudanese Patient. *Am J Microbiol Res*. 2014;2(6):217-223. doi:10.12691/ajmr-2-6-8

37. Emini EA, Hughes J V, Perlow DS, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol*. 1985;55(3):836-839. http://www.ncbi.nlm.nih.gov/pubmed/2991600. Accessed May 8, 2019.

38. Larsen J, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res*. 2006;2(1):2. doi:10.1186/1745-7580-2-2

39. Ponomarenko J V, Bourne PE. Antibody-protein interactions: benchmark datasets and prediction tools evaluation. *BMC Struct Biol*. 2007;7(1):64. doi:10.1186/1472-6807-7-64

40. Haste Andersen P, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci*. 2006;15(11):2558-2567. doi:10.1110/ps.062405906

41. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*. 1990;276(1-2):172-174. http://www.ncbi.nlm.nih.gov/pubmed/1702393. Accessed May 8, 2019.

42. Ponomarenko J, Bui H-H, Li W, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*. 2008;9(1):514. doi:10.1186/1471-2105-9-514

43. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*. 2016;32(4):511-517. doi:10.1093/bioinformatics/btv639

44. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11. *Nucleic Acids Res*. 2008;36(suppl_2):W509-W512. doi:10.1093/nar/gkn202
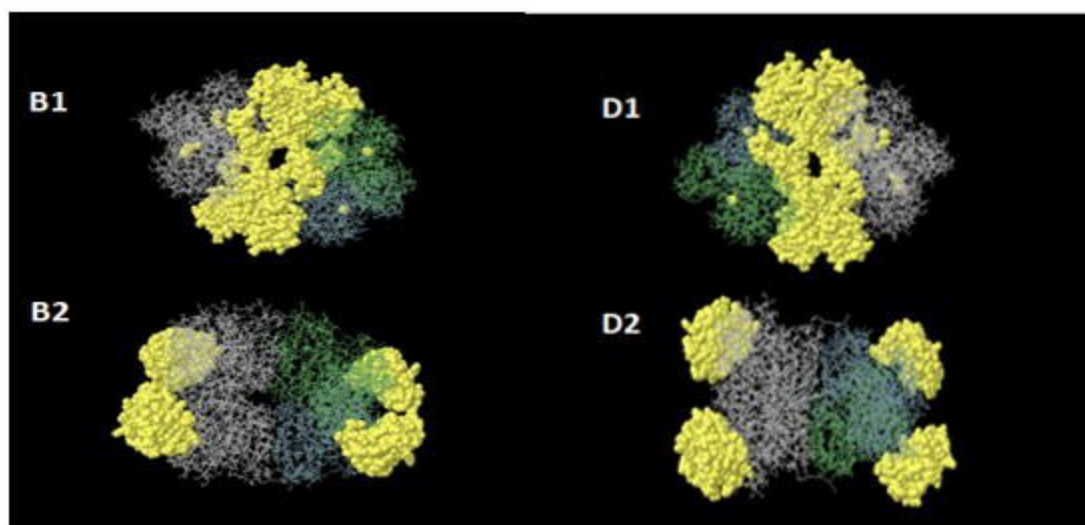
45.   Lundegaard C, Nielsen M, Lund O. The validity of predicted T-cell epitopes. *Trends Biotechnol*. 2006;24(12):537-538. doi:10.1016/j.tibtech.2006.10.001

46.   Lundegaard C, Lund O, Nielsen M. Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics*. 2008;24(11):1397-1398. doi:10.1093/bioinformatics/btn128

47.   Nielsen M, Lundegaard C, Worning P, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci*. 2003;12(5):1007-1017. doi:10.1110/ps.0239403

48.   Buus S, Lauemøller SL, Worning P, et al. Sensitive quantitative predictions of peptide-MHC binding by a "Query by Committee" artificial neural network approach. *Tissue Antigens*. 2003;62(5):378-384. http://www.ncbi.nlm.nih.gov/pubmed/14617044. Accessed May 8, 2019.

49.   Nielsen M, Lund O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics*. 2009;10(1):296. doi:10.1186/1471-2105-10-296

50.   Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 2006;7(1):153. doi:10.1186/1471-2105-7-153

51.   Shen Y, Maupetit J, Derreumaux P, Tufféry P. Improved PEP-FOLD Approach for Peptide and Miniprotein Structure Prediction. *J Chem Theory Comput*. 2014;10(10):4745-4758. doi:10.1021/ct500592m

52.   Thevenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tuffery P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Res*. 2012;40(W1):W288-W293. doi:10.1093/nar/gks419

53.   Spassov VZ, Yan L. *p* H-selective mutagenesis of protein-protein interfaces: *In silico* design of therapeutic antibodies with prolonged half-life. *Proteins Struct Funct Bioinforma*. 2013;81(4):704-714. doi:10.1002/prot.24230

54.   Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera?A visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-1612. doi:10.1002/jcc.20084

55.   Gasteiger E, Hoogland C, Gattiker A, et al. *Protein Analysis Tools on the ExPASy Server 571 571 From: The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server*. http://www.expasy.org/tools/. Accessed May 17, 2019.

56.   Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera?A visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-1612. doi:10.1002/jcc.20084

57.   Chen JE, Huang CC, Ferrin TE. RRDistMaps: a UCSF Chimera tool for viewing and comparing protein distance maps. *Bioinformatics*. 2015;31(9):1484-1486. doi:10.1093/bioinformatics/btu841

58.   Hertig S, Goddard TD, Johnson GT, Ferrin TE. Multidomain Assembler (MDA) Generates Models of Large Multidomain Proteins. *Biophys J*. 2015;108(9):2097-2102. doi:10.1016/j.bpj.2015.03.051
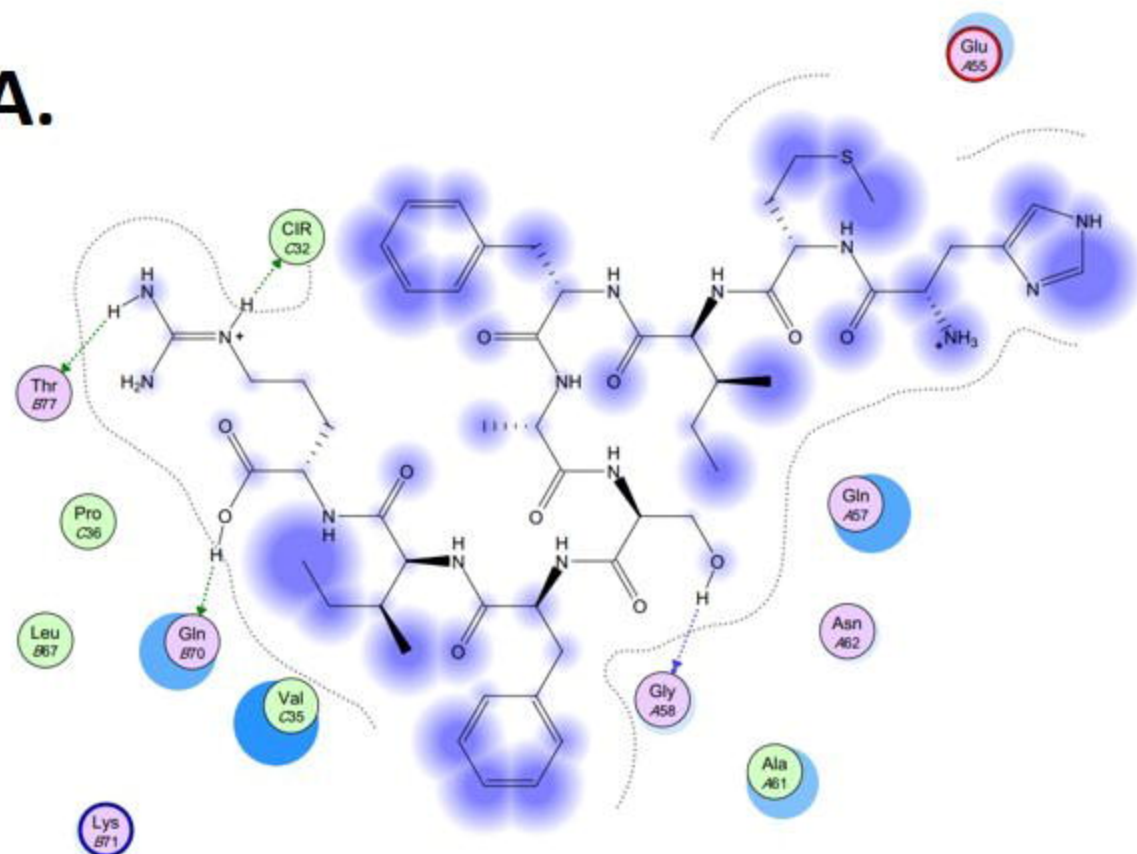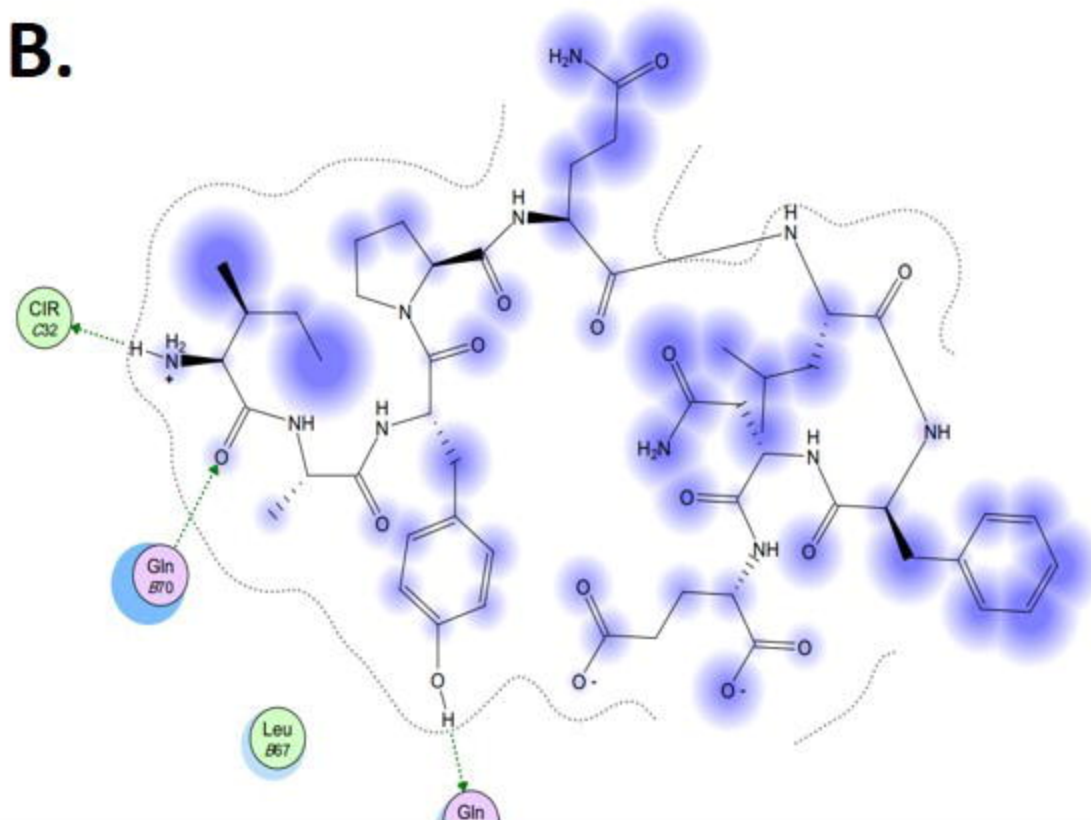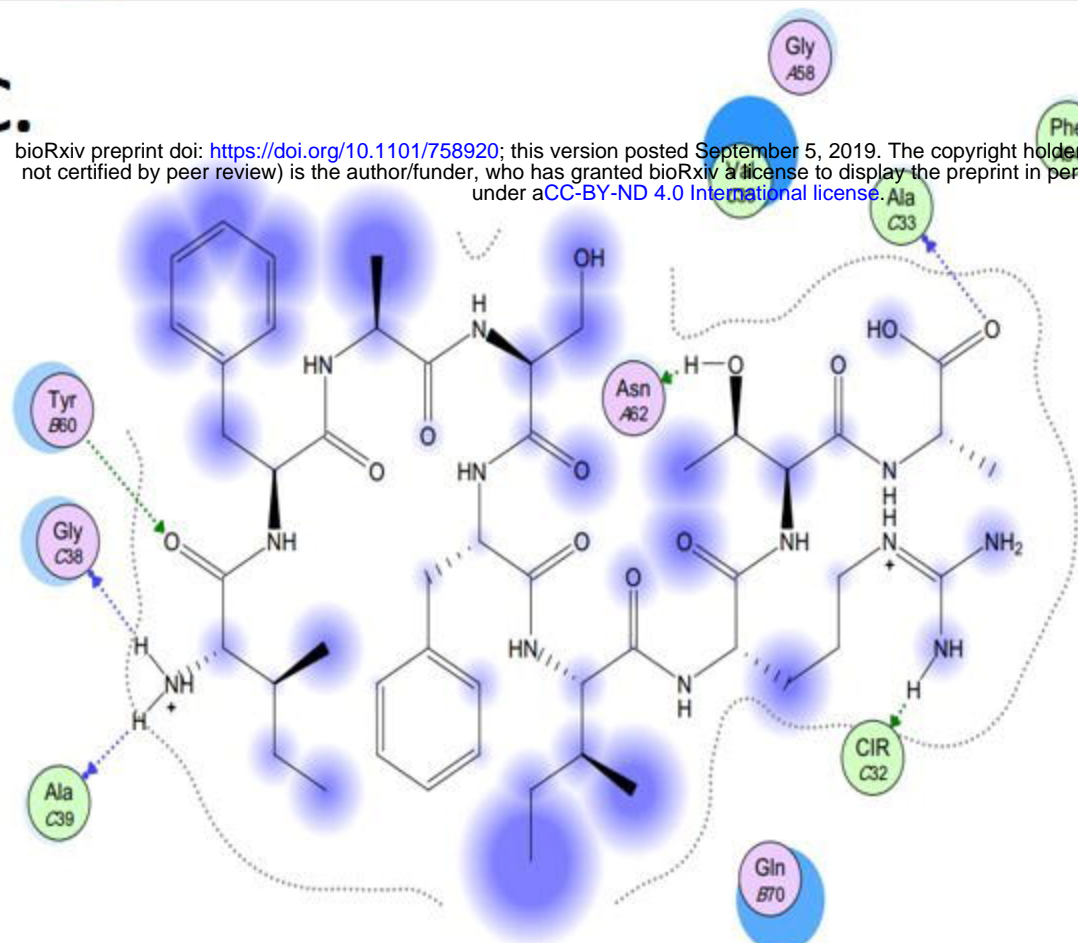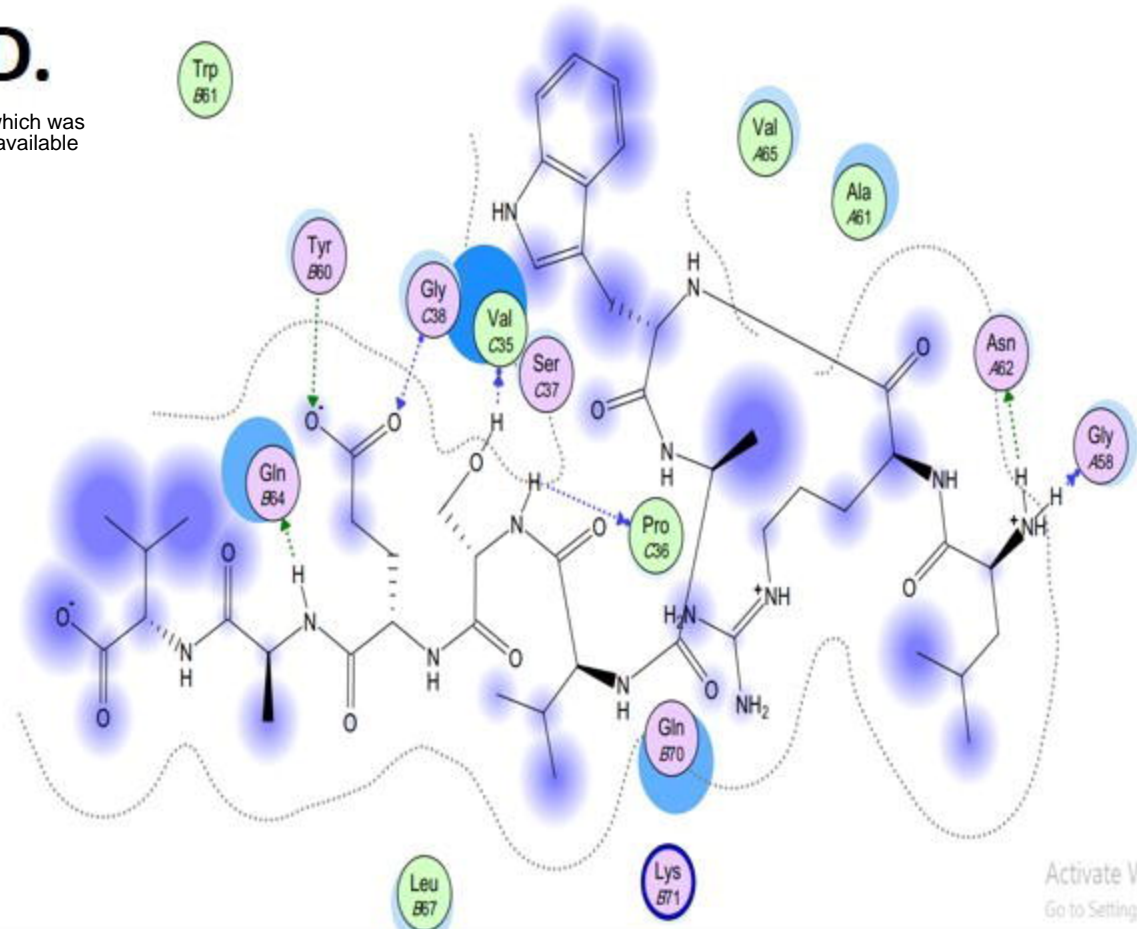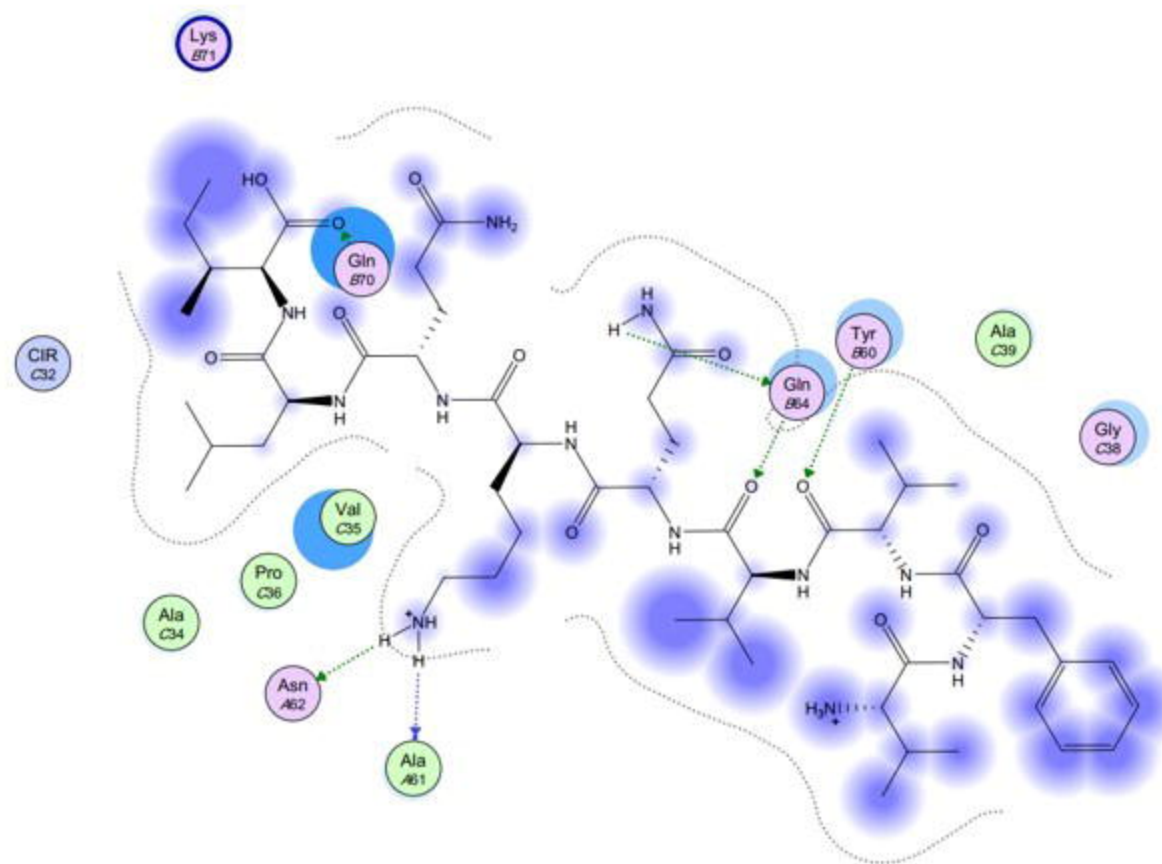
59. de Klerk N, de Vogel C, Fahal A, van Belkum A, van de Sande WWJ. Fructose-bisphosphate aldolase and pyruvate kinase, two novel immunogens in *Madurella mycetomatis*. *Med Mycol*. 2012;50(2):143-151. doi:10.3109/13693786.2011.593005

60. Berčič RL, Slavec B, Lavrič M, et al. Identification of major immunogenic proteins of isolates. doi:10.1016/j.vetmic.2007.07.020ï

61. Pitarch A, Pardo M, Jiménez A, et al. Two-dimensional gel electrophoresis as analytical tool for identifyingCandida albicans immunogenic proteins. *Electrophoresis*. 1999;20(4-5):1001-1010. doi:10.1002/(SICI)1522-2683(19990101)20:4/5<1001::AID-ELPS1001>3.0.CO;2-L

62. Pardo M, Ward M, Pitarch A, et al. Cross-species identification of novelCandida albicans immunogenic proteins by combination of two-dimensional polyacrylamide gel electrophoresis and mass spectrometry. *Electrophoresis*. 2000;21(13):2651-2659. doi:10.1002/1522-2683(20000701)21:13<2651::AID-ELPS2651>3.0.CO;2-3

63. Swoboda RK, Bertram G, Hollander H, et al. Glycolytic enzymes of Candida albicans are nonubiquitous immunogens during candidiasis. *Infect Immun*. 1993;61(10):4263-4271. http://www.ncbi.nlm.nih.gov/pubmed/8406815. Accessed May 8, 2019.

64. Matsumoto K, Noguchi M, Satoh T, et al. A phase I study of personalized peptide vaccination for advanced urothelial carcinoma patients who failed treatment with methotrexate, vinblastine, adriamycin and cisplatin. *BJU Int*. January 2011:no-no. doi:10.1111/j.1464-410X.2010.09933.x

65. Schwartzentruber DJ, Lawson DH, Richards JM, et al. gp100 Peptide Vaccine and Interleukin-2 in Patients with Advanced Melanoma. *N Engl J Med*. 2011;364(22):2119-2127. doi:10.1056/NEJMoa1012863

66. Skariyachan S, Taskeen N, Ganta M, Venkata Krishna B. Recent perspectives on the virulent factors and treatment options for multidrug-resistant *Acinetobacter baumannii*. *Crit Rev Microbiol*. April 2019:1-19. doi:10.1080/1040841X.2019.1600472

67. Obara W, Kanehira M, Katagiri T, Kato R, Kato Y, Takata R. Present status and future perspective of peptide-based vaccine therapy for urological cancer. *Cancer Sci*. 2018;109(3):550-559. doi:10.1111/cas.13506

68. Ahmad Fuaad A, Azmi F, Skwarczynski M, et al. Peptide Conjugation via CuAAC 'Click' Chemistry. *Molecules*. 2013;18(11):13148-13174. doi:10.3390/molecules181113148

A.

B.

C.

PRESSURE over JET of USER_DISCRATCH

World - Class I Coverage

| BAC class | Coverage | Average kb | FCR |

Number of contigs into HLA contributions sequences

MHC class	Coverage	Average hit	PC90
II	99.7%	7.77	4.65

World - Class II Coverage
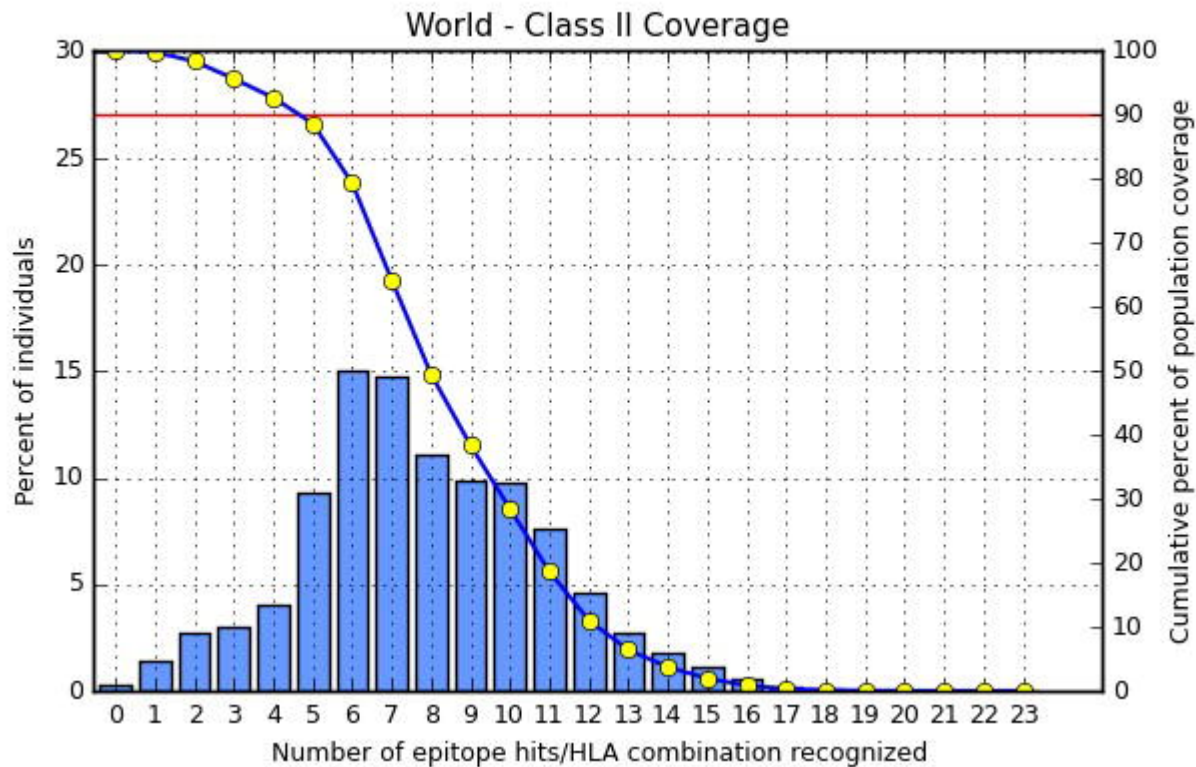
MHC class	Coverage	Average hit	PC90
combined	99.75%	4.88	3.0

World - Class combined Coverage

B

C

A

E

D

360°