ORIGINAL RESEARCH

Vaccinomics approach for scheming potential epitope-based peptide vaccine by targeting ʟ-protein of Marburg virus

Tahmina Pervin[1] [ORCID] · **Arafat Rahman Oany[2,3]** [ORCID]

Abstract

Marburg virus is one of the world's most threatening diseases, causing extreme hemorrhagic fever, with a death rate of up to 90%. The Food and Drug Administration (FDA) currently not authorized any treatments or vaccinations for the hindrance and post-exposure of the Marburg virus. In the present study, the vaccinomics methodology was adopted to design a potential novel peptide vaccine against the Marburg virus, targeting RNA-directed RNA polymerase (ʟ). A total of 48 ʟ-proteins from diverse variants of the Marburg virus were collected from the NCBI GenBank server and used to classify the best antigenic protein leading to predict equally T and B-cell epitopes. Initially, the top 26 epitopes were evaluated for the attraction with major histocompatibility complex (MHC) class I and II alleles. Finally, four prospective central epitopes NLSDLTFLI, FRYEFTRHF, YRLRNSTAL, and YRVRNVQTL were carefully chosen. Among these, FRYEFTRHF and YRVRNVQTL peptides showed 100% conservancy. Though YRLRNSTAL showed 95.74% conservancy, it demonstrated the highest combined score as T cell epitope (2.5461) and population coverage of 94.42% among the whole world population. The epitope was found non-allergenic, and docking interactions with human leukocyte antigens (HLAs) also verified. Finally, in vivo analysis of the recommended peptides might contribute to the advancement of an efficient and exclusively prevalent vaccine that would be an active route to impede the virus spreading.

Keywords Vaccinomics · HLA · Hemorrhagic fever · Allergenicity · Molecular docking

Introduction

A member of the genera Marburg virus (MABV) belongs to Filoviridae family that causes an acute life-threatening hemorrhagic fever in human being and other primates. The fever caused by MABV is identical to the Ebola virus (EBOV) disease and therefore is characterized by fever, severe inflammatory reaction, pathological coagulation, and systemic hemorrhaging (Mehedi et al. 2011). To date, Marburg virus disease (MVD) is related to 469 total cases and 376 reported deaths (Peterson et al. 2006; Nyakarahuka et al. 2017). Though fewer cases for MABV recorded, forthcoming epidemics and the rapid expansion of MABV to non-endemic areas are of boundless distress. The mortality rate of this virus reported about 81% (Mehedi et al. 2011).

The Egyptian rousette (*Rousettus aegyptiacus*) is publicly classified as the Marburg virus-host and has a wide geographical distribution. The pretended cause for the dispersal of MVD might be the close contact between humans and animals like, other than human primates, bats, and cattle (Towner et al. 2009). It has been listed by World Health Organization (WHO) as a category 4 risk agent because of its lethality (Chosewood and Wilson 2009).

The Marburg virus possesses about 19 kilobase-pair long, negative sense, un-segmented RNA genome, encoding 7 open reading frames (ORF); nucleoprotein NP, virion protein (VP) 35, VP40, glycoprotein (GP), VP24, and viral RNA-dependent RNA polymerase (ʟ) (Feldmann et al. 1992). The whole-genome of filovirus is enveloped into a single threadlike virion, having 790–970 nm length and 80 nm width (Geisbert 1995). Based on the structure and functions of Marburg virus ʟ-protein, the

✉ Arafat Rahman Oany
  arafatr@outlook.com

[1] Biotechnology and Genetic Engineering Discipline, Life Science School, Khulna University, Khulna 9208, Bangladesh

[2] Department of Biotechnology and Genetic Engineering, Faculty of Life Science, Mawlana Bhashani Science and Technology University, Tangail 1902, Bangladesh

[3] Aristopharma Limited, Dhaka, Bangladesh

enzymatically active subunit (L) of the MARV polymerase comprises of 2331 amino acids (Mühlberger et al. 1992). The L-protein conjunction with VP 35, the polymerase cofactor, forms the RNA-dependent polymerase complex that is crucial for transcription and duplication of the virus (Mühlberger et al. 1998).

Because of the emergence of the Marburg virus outbreak, novel therapeutic targets against this pathogen need to be identified immediately. Currently, no FDA approved vaccines or drugs are available to defend the human against the Marburg virus (Brauburger et al. 2012). The only prime treatment of patients during the epidemics was sympathetic care (fluids, antimicrobials, blood transfusion) (Martines et al. 2015; Olival and Hayman 2014).

Identifying precise epitopes from pathogens causing infections has ominously advanced the progress of peptide vaccines grounded on the epitope. Better knowledge on the bio-molecular base of target identification and human leukocyte antigen (HLA) binding motives, culminated in the development of logically engineered vaccines based uniquely on algorithms forecasting the binding of the epitope to HLA (Oany et al. 2014, 2015a, b, 2017). The epitope-based vaccine is chemically constant, more precise, and without any potentially pathogenic or oncogenic exposure (Holland and Domingo 1998; Sette et al. 2002), but the formulation of laboratory-based aspirant epitope is not only costly but also painstaking, requiring diverse laboratory medicine experiments for ultimate epitope choice.

Vaccinomics means the implementation of combined expertise from various fields, including immunogenetics and immunogenomics, to establish and understand the immune response of candidates for the next generation vaccine (Poland et al. 2009). At present, different vaccinomics databases are available for identifying particular B lymphocyte epitopes and highly sensitive and specific HLA ligands (Brusic et al. 1998; Rammensee et al. 1999). The vaccinomics strategy has already proved its promise with optimal findings in the detection of the retained epitope for human coronavirus (Oany 2014), Ebola virus (Oany et al. 2015b), and *Shigella* (Oany et al. 2017).

In the present work, vaccinomics approaches have been executed to design potential conserved epitope candidate that can be used for the vaccine origination against the deadly Marburg virus, by targeting protein L, with an anticipation of further wet lab endorsement.

Methods

The overall schematic diagram for vaccine designing is shown in Fig. 1.

Sequence collection and antigenic protein determination

The L-protein sequences of the Marburg virus were collected from GenBank (Clark et al. 2016) and evaluated by the Immunomedicine Group server based on the Kolaskar method (Kolaskar and Tongaonkar 1990), to ascertain the maximum effective antigenic protein.

T-cell epitope extrapolation and affinity with MHC

The epitope estimation for the corresponding protein and its affinity score with MHC class I and class II allele were determined using the previously used approach (Oany 2014; Oany et al. 2015b). Briefly, the NetCTL server (Larsen et al. 2007) assessed the effective cytotoxic T-lymphocyte epitopes from the most immunogenic protein. The default algorithms were used for this assessment. Epitopes having maximum scores were selected.

The Immune Epitope Database (IEDB) T-cell epitope detection methods were utilized to determine an association with both MHC molecules (Buus et al. 2003; Wang et al. 2008, 2010). To measure the half-maximum inhibitory concentration (IC_{50}) for the binding of a pre-selected 9.0-mer epitope with MHC-I, the stabilized matrix method (SMM) was employed. The IEDB-recommended approach was used to study MHC class II interaction for the particular HLA-DP, HLA-DQ, and HLA-DR loci. Taking into consideration the pre-defined 9-mer epitope and its preserved region in the Marburg virus, 15-mer epitopes were intended for binding interaction analysis of MHC-II. In case of MHC-I, the epitopes consisting of $IC_{50} < 250$ nM and the epitopes containing percentile rank < 50 for MHC-II alleles were chosen for even more assessments.

Epitope management and population distribution study

The aspirant epitope preservation was checked using a web-based epitope conservation method available in the IEDB research database (Bui et al. 2007). The degree of conservation of each possible epitope was evaluated by taking attributes from the database of all L-protein sequences of different strains. Additionally multiple
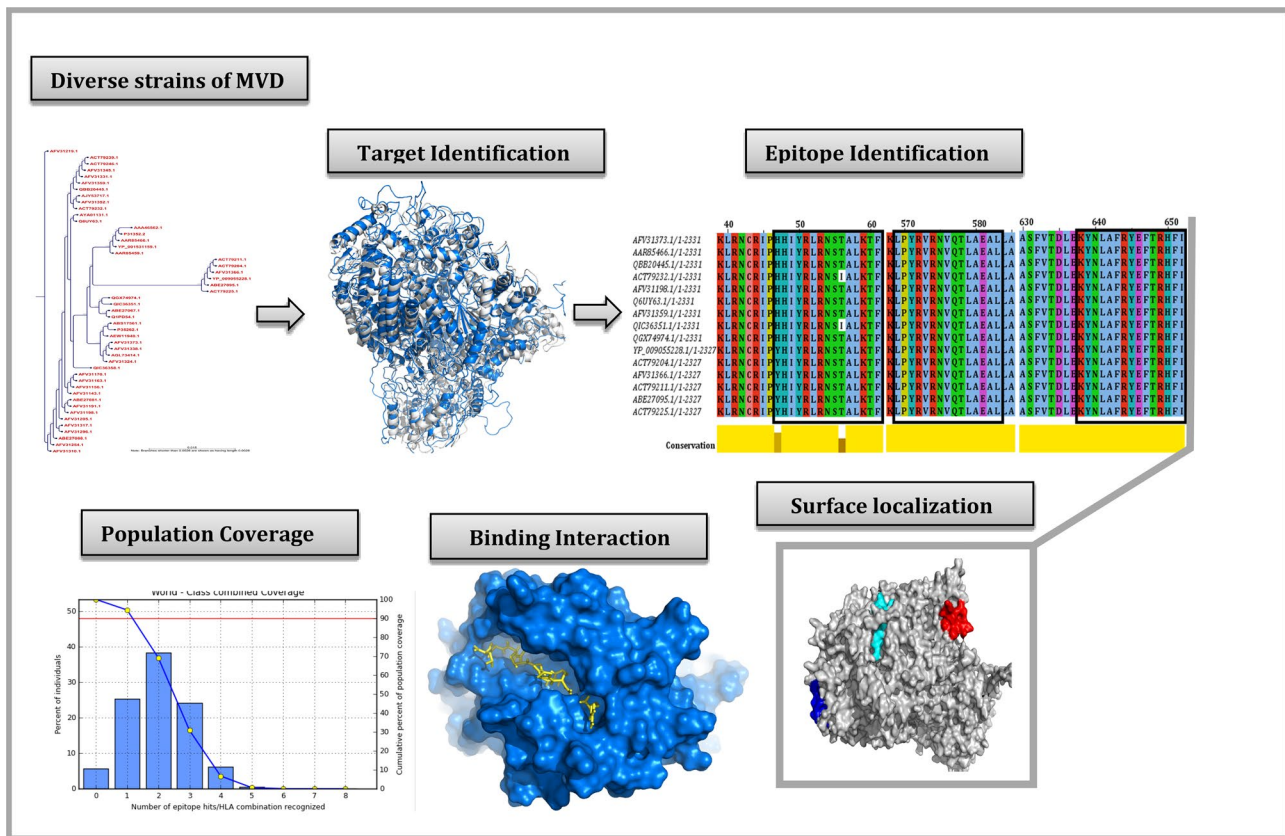
Fig. 1 The overall schematic diagram for vaccine designing

sequence alignment (MSA) was constructed by BioEdit software (Hall and Biosciences 2011) and retrieved through Jalview (http://www.jalview.org/) tool to identify the epitope locations inside the sequences. The phylogram was generated through the CLC Sequence Viewer (Bio-Qiagen 2016). Furthermore, the WebLogo server (Crooks et al. 2004) was used to visualize the conserved region of the MSA. The population distribution for the epitope was analyzed by IEDB estimation tool (Bui et al. 2006), and the collective value for both MHC groups was calculated.

## Homology modeling, energy minimization and structural authentication

The homology model of the targeted protein was acquired by MODELLERv9 (Šali et al. 1995) and the PROCHECK server of the SWISS-MODEL Workspace evaluated the projected structure (Laskowski et al. 1996; Arnold et al. 2006). The best suitable template was used for the construction of the model to cover the whole region of the protein. Before the assessment, the energy minimization of the structure was done by the GROMOS 96 (Scott et al.

1999) force field. The superposition of the prototype and model configuration had been examined to ensure the model's fair consistency. Additionally, the ProSA-web server (Oany et al. 2020; Wiederstein and Sippl 2007) was also utilized for assessing the model quality through Z-score measurement.

Molecular docking and association of HLA allele's scrutiny

Studies of docking were also carried out using the best available epitope, adopting the method used in previous studies (Oany et al. 2014, 2015a). Auto-Dock Vina (Trott and Olson 2010) was used for the docking analysis. In our analysis we designated HLA-C*07:02 and DRB1*04:01 as the aspirants for MHC-I and II, respectively, for docking study since they are accessible hits in the Protein Data Bank (PDB) server. The structure 5VGE, HLA-C*07:02 complexes with RYR peptide, and 5JLZ-human HLA-DRB1*04:01 complexes with modified alpha-enolase peptide-were extracted from the Research Collaboratory for Structural Bioinformatics (RCSB) protein databank (Berman et al. 2000). Finally, the

Fig. 2 Phylogenetic tree showing the evolutionary divergence among the different RNA-dependent RNA polymerase-ʟ proteins of the Marburg virus. Here, the phylogramic view is shown with appropriate distance among the different strains. The blue dotted view indicates the node of the tree

Note: Branches shorter than 0.0026 are shown as having length 0.0026
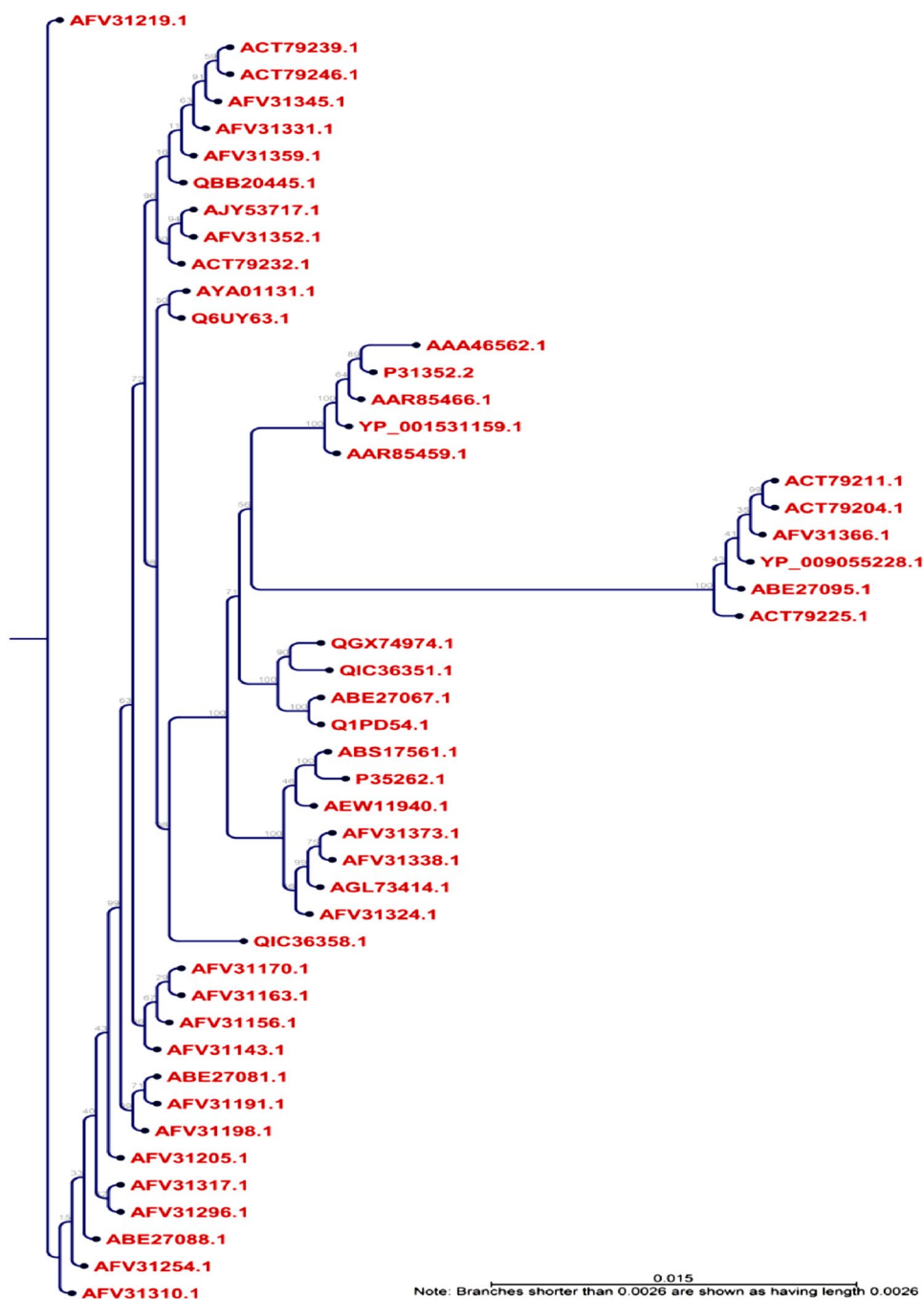
complex-structures were generalized through PyMOL (version 1.5.0.4) for the ultimate docking analysis.

For epitope 3D structure transformation, the preferred PEP-FOLD server was utilized to convert the structure from the sequences (Thévenet et al. 2012). To evaluate the relationship with HLA alleles, the 9-mer epitope of the MHC I and 15-mer epitope of the MHC II molecule were used. Finally, all the proteins and the peptides were minimized using GROMOS 96 (Scott et al. 1999) force field for the

docking analysis through using Swiss-Pdb viewer (Guex and Peitsch 1997).

The grid-line for the docking of MHC-I molecules was of X: 58.9062, Y: 28.6331, and Z: 30.7545 (Angstrom) at the center of X: 16.9834, Y: −61.6639, and Z: 17.8431 and the grid-line for the docking of MHC-II molecules was of X: 26.9773, Y: 33.5139, and Z: 26.0699 (Angstrom) at the center of X: −41.0931, Y: 9.7731, and Z: −26.5406. A control docking was also performed with an experimentally acknowledged peptide-MHC-compound. The human

Table 1 Predicted T-cell epitopes by NetCTL server on the basis of combined score

Epitope	Supertypes	Start position	Combined score
QSSLPVWLY	A1	1595	3.4352
YSGNIVHRY		1288	2.9568
FLEKEELFY		945	2.8701
NLFDWMHFL	A2	663	1.5315
NLSDLTFLI		1512	1.4496
YILFFPFGL		2315	1.3993
RLRNSTALK	A3	51	1.7195
VTFRLMLNK		1367	1.6401
HLHSLMLDY		2210	1.5581
NYPASLHKF	A24	2161	1.9153
SYQNFINNF		1579	1.8991
KWKKTDYLF		1961	1.8145
WTIGNRAPY	A26	1210	2.4674
DVANFLRAY		1111	2.3636
EFIVSVASY		229	1.9937
APSYRNFSF	B7	545	1.8800
RVSRSTLSL		1421	1.7969
HPSVNRLAW		1202	1.6682
KLKLKSSVM	B8	735	2.0038
FIKDRATAV		484	1.9265
KPKVMFETF		403	1.8823
NRWKSWFSY	B27	1083	1.9066
FRYEFTRHF		643	1.8129
YRLRNSTAL		50	1.7871
FHSIWDHIL	B39	72	2.6241
YRLRNSTAL		50	2.5461
YRVRNVQTL		571	2.4425
WEICARAWL	B44	254	1.9816
SESESTINL		1161	1.8007
AEALLADGL		580	1.7186
LTVPFHSIW	B58	68	1.9674
SSLPVWLYF		1596	1.9236
KSSDVHEDF		1803	1.9065
KQHIVSNSF	B62	443	1.4985
WTIGNRAPY		1210	1.4914
FQSQMIKSY		322	1.4696

Here, all the 12 different supertypes' epitopes are showed

HLA-C*07:02 bound with RYR peptide (PDB ID: 5VGE) was selected for this study. The grid-line for the docking focusing at X: 25.6546, Y: − 60.4063, and Z: 11.8845.

Allergenicity assessment and B-cell epitope detection

A web-based AlgPred server was used to estimate the allergenicity of the possible epitopes (Saha and Raghava 2006)

by using the support vector machine (SVM) algorithm at a threshold value of − 0.4. Estimating technique follows guidelines of the Food and Agriculture Organization/World Health Organization, 2003. IEDB-AR screened the projected T-cell epitope (15-mer) using multiple web-based methods for the feasibility of the B-cell epitope (Kolaskar and Tongaonkar 1990; Chou and Fasman 1978; Emini et al. 1985; Karplus and Schulz 1985; Larsen et al. 2006; Parker et al. 1986).

Results

Sequence collection and antigenic protein determination

From the NCBI GenBank database, entirely 48 L-protein molecules from diverse varieties of the Marburg virus were obtained. The L-proteins MSA was retrieved from the BioEdit software via ClustalW using 1000 bootstrap replicates (Figure S1). The CLC Sequence Viewer had been utilized to construct the phylogram based on MSA obtained from BioEdit, to determine the distinction between the sequences received and is depicted in the Fig. 2. Finally, after subsequent evaluation from Immunomedicine Group server, the maximum antigenicity score of 1.0392 was found for the accession number AFV31373.1 (Table S1), and subsequently analyzed for the highly immunogenic epitope.

T-cell epitope extrapolation and affinity with MHC

The NetCTLv1.2 server was utilized for the identification of T-cell epitopes where the calculation was restricted to 12 MHC-I supertypes. The top 26 epitopes (Table 1) were selected from the supertypes for further comprehensive analysis based on the cumulative values. Table 2 lists the restricted MHC-I alleles for which the epitope exhibited increased specificity ($IC_{50} < 250$ nM) and Table 3 displays findings from the MHC-II interaction study (percentile rank < 50).

Epitope management and population distribution study

The IEDB conservation analysis method analyzed the epitope conservation of the anticipated epitopes as shown in Table 4. The positions of the predicted epitopes are shown in the MSA of L-proteins (Fig. 3). Here, we used only 15 diverse sequences from the total retrieved proteins, 48 sequences, for the proper annotation. The entire world population coverage was calculated based on the combined

Table 2 Epitopes for CD8 + T-cell along with their interacting MHC class I alleles with affinity < 250 nM

Serial no.	Epitope	Interacting MHC-I allele (IC50) on the nM scale
1.	QSSLPVWLY	HLA-B*27:20 (2.03), HLA-A*32:07 (11.35), HLA-B*15:17 (19.71), HLA-C*12:03 (27.49), HLA-A*68:23 (28.10), HLA-C*03:03 (35.86), HLA-B*40:13 (42.75), HLA-C*07:01 (52.26), HLA-B*58:01 (99.15), HLA-A*32:15 (115.41), HLA-A*29:02 (120.44), HLA-B*15:02 (136.53), HLA-C*05:01 (193.05), HLA-A*01:01 (211.62)
2.	YSGNIVHRY	HLA-C*12:03 (22.97), HLA-A*68:23 (26.17), HLA-A*32:07 (31.05), HLA-B*15:17 (40.34), HLA-A*32:15 (44.49), HLA-B*40:13 (72.10), HLA-B*27:20 (73.03), HLA-A*29:02 (85.46), HLA-C*07:01 (90.82), HLA-A*80:01 (122.18), HLA-C*03:03 (165.81), HLA-A*01:01 (182.20), HLA-C*14:02 (212.63), HLA-A*30:02 (235.75)
3.	FLEKEELFY	HLA-B*27:20 (6.99), HLA-A*80:01 (11.46), HLA-C*05:01 (16.13), HLA-A*32:07 (26.98), HLA-C*12:03 (28.07), HLA-C*03:03 (65.56), HLA-A*68:23 (68.19), HLA-B*40:13 (70.30), HLA-C*14:02 (119.29), HLA-A*29:02 (147.49), HLA-A*32:15 (158.95)
4.	NLFDWMHFL	HLA-A*02:50 (1.48), HLA-A*02:11 (3.25), HLA-A*02:16 (4.22), HLA-A*02:02 (4.58), HLA-A*02:01 (5.31), HLA-A*02:12 (5.93), HLA-A*02:19 (8.78), HLA-A*02:03 (10.71), HLA-A*02:06 (18.77), HLA-A*32:07 (22.08), HLA-A*69:01 (23.93), HLA-B*40:13 (25.82), HLA-B*15:02 (28.26), HLA-A*68:23 (37.30), HLA-C*12:03 (39.65), HLA-A*32:15 (44.18), HLA-B*27:20 (78.61), HLA-C*07:02 (105.72), HLA-A*02:17 (108.31), HLA-C*03:03 (128.41), HLA-C*14:02 (148.46), HLA-A*68:02 (170.84)
5.	NLSDLTFLI	HLA-A*02:50 (3.96), HLA-A*02:02 (5.74), HLA-A*02:19 (8.41), HLA-A*32:07 (11.70), HLA-A*02:12 (21.46), HLA-A*02:01 (25.18), HLA-A*02:11 (30.43), HLA-A*02:16 (38.59), HLA-B*27:20 (40.50), HLA-A*02:03 (41.78), HLA-A*02:06 (59.78), HLA-C*12:03 (74.17), HLA-B*40:13 (81.46), HLA-A*68:23 (89.28), HLA-A*32:15 (95.12), HLA-C*03:03 (165.81), HLA-A*68:02 (169.66), HLA-C*07:01 (211.92), HLA-A*02:17 (228.38), HLA-A*69:01 (242.58)
6.	RLRNSTALK	HLA-A*30:01 (6.40), HLA-A*32:07 (11.86), HLA-A*03:01 (13.73), HLA-B*27:20 (22.41), HLA-A*68:23 (35.62), HLA-C*12:03 (60.15), HLA-A*32:15 (110.47), HLA-A*31:01 (111.55), HLA-A*11:01 (228.19)
7.	VTFRLMLNK	HLA-A*11:01 (5.61), HLA-A*68:23 (9.88), HLA-B*27:20 (12.96), HLA-C*12:03 (18.93), HLA-A*03:01 (28.69), HLA-A*32:07 (28.78), HLA-B*40:13 (34.75), HLA-A*30:01 (75.16), HLA-C*14:02 (109.55), HLA-A*68:01 (120.36), HLA-A*32:15 (179.58), HLA-A*31:01 (230.92)
8.	NYPASLHKF	HLA-A*24:03 (2.16), HLA-C*14:02 (5.00), HLA-A*32:07 (16.52), HLA-B*40:13 (20.56), HLA-A*02:50 (20.71), HLA-A*23:01 (30.23), HLA-A*68:23 (34.97), HLA-B*27:20 (37.02), HLA-C*12:03 (44.90), HLA-A*32:15 (63.42), HLA-B*15:02 (70.67), HLA-A*24:02 (75.79), HLA-C*07:02 (84.17)
9.	SYQNFINNF	HLA-A*24:03 (1.37), HLA-A*32:07 (8.05), HLA-C*14:02 (12.35), HLA-A*68:23 (21.07), HLA-C*12:03 (43.07), HLA-A*23:01 (50.28), HLA-B*40:13 (58.87), HLA-A*24:02 (63.48), HLA-B*27:20 (68.94), HLA-C*07:02 (77.83), HLA-B*15:02 (98.45), HLA-B*15:03 (123.05), HLA-A*32:15 (124.52), HLA-A*02:50 (198.25)
10.	WTIGNRAPY	HLA-A*68:23 (1.84), HLA-C*03:03 (3.02), HLA-B*15:17 (4.96), HLA-C*12:03 (22.76), HLA-A*26:02 (25.02), HLA-A*32:15 (32.68), HLA-A*29:02 (39.42), HLA-A*26:03 (48.62), HLA-A*26:01 (54.78), HLA-A*32:07 (67.46), HLA-A*30:02 (68.15), HLA-B*40:13 (72.43), HLA-B*35:01 (93.24), HLA-B*15:01 (101.86), HLA-B*15:02 (106.96), HLA-B*27:20 (112.33), HLA-C*14:02 (113.14), HLA-A*25:01 (211.39)
11.	DVANFLRAY	HLA-A*26:02 (2.78), HLA-A*68:23 (5.11), HLA-A*25:01 (14.03), HLA-A*26:01 (39.87), HLA-C*12:03 (41.13), HLA-A*32:07 (53.22), HLA-C*03:03 (66.31), HLA-B*35:01 (94.97), HLA-B*15:02 (111.49), HLA-B*40:13 (131.20), HLA-B*27:20 (136.30), HLA-A*29:02 (162.09), HLA-A*32:15 (166.82), HLA-A*26:03 (172.10)
12.	APSYRNFSF	HLA-B*27:20 (16.01), HLA-B*42:01 (23.05), HLA-A*32:07 (23.94), HLA-C*03:03 (25.45), HLA-A*32:15 (33.90), HLA-A*68:23 (37.22), HLA-B*07:02 (38.01), HLA-A*02:50 (42.88), HLA-C*12:03 (74.00), HLA-B*40:13 (133.95), HLA-B*35:01 (161.65), HLA-B*83:01 (177.39)
13.	RVSRSTLSL	HLA-A*32:07 (10.26), HLA-B*27:20 (12.26), HLA-A*68:23 (21.97), HLA-B*15:17 (27.33), HLA-A*02:50 (53.98), HLA-C*03:03 (59.93), HLA-B*07:02 (79.23), HLA-C*12:03 (81.32), HLA-A*02:17 (99.46), HLA-B*40:13 (128.21), HLA-A*32:01 (138.02), HLA-C*15:02 (141.07), HLA-B*15:02 (141.65), HLA-A*32:15 (229.22)
14.	KLKLKSSVM	HLA-B*27:20 (9.97), HLA-C*12:03 (12.08), HLA-A*32:07 (19.14), HLA-A*02:17 (19.44), HLA-B*15:03 (22.08), HLA-A*02:50 (54.35), HLA-A*68:23 (90.31), HLA-B*08:01 (98.53), HLA-B*15:01 (106.66), HLA-A*32:15 (125.39), HLA-C*14:02 (165.81), HLA-B*40:13 (202.27), HLA-C*03:03 (211.65)
15.	FIKDRATAV	HLA-C*12:03 (4.00), HLA-A*02:50 (4.23), HLA-A*02:03 (9.01), HLA-A*02:11 (23.35), HLA-A*02:06 (24.47), HLA-A*68:23 (31.03), HLA-A*32:07 (49.10), HLA-B*40:13 (53.45), HLA-B*08:01 (58.15), HLA-C*03:03 (58.70), HLA-B*27:20 (58.82), HLA-A*02:17 (59.52), HLA-A*02:02 (61.52), HLA-A*32:15 (101.22), HLA-A*02:12 (132.95), HLA-C*14:02 (193.03), HLA-C*15:02 (200.18)
16.	NRWKSWFSY	HLA-B*27:20 (2.35), HLA-B*40:13 (30.47), HLA-A*32:07 (39.63), HLA-C*12:03 (41.61), HLA-A*68:23 (68.51), HLA-B*15:02 (78.57), HLA-B*73:01 (103.21), HLA-B*27:05 (146.41), HLA-C*07:02 (148.65), HLA-C*07:01 (162.24), HLA-A*32:15 (168.75), HLA-C*14:02 (181.39)

Table 2 (continued)

Serial no.	Epitope	Interacting MHC-I allele (IC50) on the nM scale
17.	FRYEFTRHF	HLA-B*27:20 (1.08), HLA-C*06:02 (5.17), HLA-C*07:02 (16.08), HLA-C*07:01 (16.34), HLA-B*15:03 (20.37), HLA-B*40:13 (21.18), HLA-B*15:02 (36.16), HLA-A*68:23 (38.17), HLA-A*02:50 (42.98), HLA-A*32:07 (43.96), HLA-C*03:03 (46.41), HLA-C*12:03 (49.68), HLA-C*14:02 (77.02), HLA-B*27:05 (89.86)
18.	FHSIWDHIL	HLA-A*02:50 (3.54), HLA-B*39:01 (5.36), HLA-B*38:01 (6.87), HLA-B*27:20 (8.96), HLA-C*03:03 (16.62), HLA-B*40:13 (17.46), HLA-A*32:07 (19.55), HLA-B*15:02 (19.78), HLA-A*68:23 (23.54), HLA-A*32:15 (24.11), HLA-B*15:09 (42.99), HLA-C*12:03 (79.11), HLA-C*07:02 (187.14), HLA-A*02:02 (207.50), HLA-A*02:17 (239.14)
19.	YRLRNSTAL	HLA-B*27:20 (2.10), HLA-C*03:03 (4.35), HLA-B*39:01 (9.87), HLA-C*14:02 (23.91), HLA-B*15:02 (24.79), HLA-A*02:50 (26.50), HLA-A*68:23 (44.85), HLA-A*32:07 (60.54), HLA-A*02:17 (74.93), HLA-B*14:02 (93.55), HLA-C*07:02 (94.44), HLA-C*07:01 (101.90), HLA-B*27:05 (105.82), HLA-C*12:03 (151.78), HLA-C*06:02 (199.44)
20.	YRVRNVQTL	HLA-B*27:20 (1.64), HLA-C*03:03 (7.83), HLA-A*02:50 (8.26), HLA-C*07:02 (17.47), HLA-B*15:02 (19.83), HLA-B*39:01 (28.14), HLA-A*32:07 (43.16), HLA-C*07:01 (45.62), HLA-C*12:03 (58.10), HLA-A*68:23 (65.73), HLA-C*06:02 (68.84), HLA-C*14:02 (71.39), HLA-A*02:17 (120.13), HLA-B*40:13 (144.52), HLA-B*27:05 (160.17)
21.	WEICARAWL	HLA-A*02:50 (4.28), HLA-B*27:20 (18.56), HLA-C*03:03 (19.26), HLA-B*40:01 (21.62), HLA-A*68:23 (30.53), HLA-B*15:02 (34.77), HLA-B*40:13 (38.45), HLA-A*02:17 (68.18), HLA-C*12:03 (70.02), HLA-A*32:07 (99.55), HLA-B*18:01 (131.84), HLA-A*32:15 (170.71)
22.	SESESTINL	HLA-A*32:07 (17.18), HLA-A*68:23 (25.51), HLA-B*15:02 (32.60), HLA-B*40:01 (34.51), HLA-A*02:50 (39.83), HLA-B*40:02 (41.77), HLA-B*27:20 (54.89), HLA-B*40:13 (72.77), HLA-C*12:03 (88.15), HLA-C*03:03 (109.30), HLA-A*32:15 (141.66), HLA-A*02:17 (146.78), HLA-C*07:02 (241.64)
23.	LTVPFHSIW	HLA-B*15:17 (3.81), HLA-B*58:01 (13.47), HLA-A*68:23 (14.55), HLA-B*57:01 (21.08), HLA-C*12:03 (23.29), HLA-C*03:03 (35.12), HLA-B*27:20 (75.25), HLA-A*32:07 (79.63), HLA-A*32:15 (233.48)
24.	SSLPVWLYF	HLA-B*15:17 (4.53), HLA-A*32:07 (7.35), HLA-B*27:20 (9.21), HLA-B*15:03 (18.71), HLA-B*58:01 (37.26), HLA-A*32:15 (71.82), HLA-B*40:13 (72.93), HLA-A*68:23 (86.85), HLA-B*15:02 (136.85), HLA-C*12:03 (163.01), HLA-C*03:03 (234.21), HLA-B*57:01 (239.23), HLA-C*14:02 (240.23)
25.	KQHIVSNSF	HLA-B*27:20 (1.24), HLA-B*15:03 (1.53), HLA-B*40:13 (5.21), HLA-A*32:07 (7.13), HLA-B*15:01 (18.71), HLA-A*68:23 (35.87), HLA-A*32:01 (37.93), HLA-A*02:50 (56.91), HLA-A*32:15 (57.44), HLA-C*12:03 (99.13), HLA-C*14:02 (162.04), HLA-A*30:01 (238.22), HLA-A*02:06 (244.09)
26.	WTIGNRAPY	HLA-A*68:23 (1.84), HLA-C*03:03 (3.02), HLA-B*15:17 (4.96), HLA-C*12:03 (22.76), HLA-A*26:02 (25.02), HLA-A*32:15 (32.68), HLA-A*29:02 (39.42), HLA-A*26:03 (48.62), HLA-A*26:01 (54.78), HLA-A*32:07 (67.46), HLA-A*30:02 (68.15), HLA-B*40:13 (72.43), HLA-B*35:01 (93.24), HLA-B*15:01 (101.86), HLA-B*15:02 (106.96), HLA-B*27:20 (112.33), HLA-C*14:02 (113.14), HLA-A*25:01 (211.39)

MHC-I and MHC-II class with the top chosen interrelated alleles (Fig. 4 and Table S2).

Homology modeling, energy minimization and structural authentication

The three-dimensional configuration of the selected protein was constructed by MODELLER using the paramount template based modeling method and the template 6V85_A was utilized for this purpose. The energy minimization of the model was done from energy level 43,319.410 to − 60,879.637 through GROMOS 96 force field. The superposition view between the model and template is shown in Fig. 5a with an RMSD (root-mean-square deviation) of 2.08. The validation of the model was assessed by the PROCHECK server through the Ramachandran plot and is shown in Fig. 5c, where 90.51% of amino acid residues were originated within the most-favored area. Furthermore, the ProSA Z score was also implemented for validation and shown in

Fig. 5b. Additionally, the top four epitopes were shown on the protein surface in Fig. 5c.

Molecular docking and association of HLA allele's scrutiny

The structural minimization of the peptide was done before the docking analysis and the minimization obtained from energy level 62.3 to − 523.53 for 15-mer and − 287 to − 539.43 for 9-mer, respectively. The 9.0-mer (YRLRN-STAL) and its 15-mer (HHIYRLRNSTALKTF) form of the epitope were docked in the furrow of the HLA-C*07:02 and DRB1*04:01 with docking score of − 8.6 and − 6.3 kcal/mol, correspondingly. Auto Dock Vina developed various configurations of the docked peptide and picked the better one at an RMSD score of 0.0 for the ultimate measurement. The PyMOL (version 1.5.0.4) was used for the visualization of the docking poses. The amino acid residues-Glu-63, Lys-66, Gln-70, Ser-77, Asp-114, Thr-143, and Gln-155 form hydrogen bond (Fig. 6) with the 9.0-mer epitope and the

Table 3 The potential CD4 + T-cell epitopes along with their interacting MHC class II alleles based on percentile rank < 50

Serial no.	Epitope	Interacting MHC-II allele based on percentile rank < 50
1.	QSSLPVWLYFPSEGQ	HLA-DRB1*04:01 (7.8), HLA-DRB1*04:05 (9.9), HLA-DQA1*01:01/DQB1*05:01 (12), HLA-DPA1*01:03/DPB1*02:01 (17), HLA-DPA1*01:03/DPB1*04:01 (18), HLA-DPA1*02:01/DPB1*01:01 (18), HLA-DRB1*15:01 (18), HLA-DPA1*02:01/DPB1*05:01 (25), HLA-DRB1*11:01 (26), HLA-DQA1*05:01/DQB1*02:01 (32), HLA-DQA1*03:01/DQB1*03:02 (36), HLA-DQA1*04:01/DQB1*04:02 (37), HLA-DRB1*12:01 (37.5), HLA-DPA1*03:01/DPB1*04:02 (39), HLA-DRB4*01:01 (42), HLA-DRB1*08:02 (43), HLA-DRB1*09:01 (44), HLA-DRB5*01:01 (48)
2.	FLPTHYSGNIVHRYN	HLA-DRB3*02:02 (4.7), HLA-DRB3*01:01 (8), HLA-DQA1*01:02/DQB1*06:02 (17), HLA-DRB1*13:02 (23), HLA-DQA1*05:01/DQB1*03:01 (24), HLA-DRB1*15:01 (28), HLA-DPA1*03:01/DPB1*04:02 (29), HLA-DRB1*04:01 (32), HLA-DRB1*01:01 (35), HLA-DRB1*07:01 (40), HLA-DRB1*09:01 (40), HLA-DRB1*11:01 (41), HLA-DPA1*01:03/DPB1*04:01 (47), HLA-DPA1*02:01/DPB1*14:01 (48), HLA-DRB1*04:05 (48), HLA-DRB1*12:01 (49)
3.	LEFLEKEELFYILIA	HLA-DPA1*01:03/DPB1*04:01 (0.54), HLA-DPA1*01:03/DPB1*02:01 (1.2), HLA-DPA1*02:01/DPB1*05:01 (1.3), HLA-DPA1*02:01/DPB1*01:01 (1.4), HLA-DQA1*01:01/DQB1*05:01 (3.8), HLA-DQA1*05:01/DQB1*02:01 (5.4), HLA-DPA1*03:01/DPB1*04:02 (6.3), HLA-DRB3*01:01 (7.2), HLA-DRB1*12:01 (15.55), HLA-DRB1*03:01 (17), HLA-DQA1*03:01/DQB1*03:02 (26), HLA-DPA1*02:01/DPB1*14:01 (30), HLA-DRB4*01:01 (37), HLA-DRB5*01:01 (39), HLA-DRB1*15:01 (40), HLA-DQA1*04:01/DQB1*04:02 (41), HLA-DQA1*01:02/DQB1*06:02 (49)
4.	VKNLFDWMHFLIPLC	HLA-DQA1*01:01/DQB1*05:01 (0.39), HLA-DPA1*01:03/DPB1*02:01 (1.3), HLA-DPA1*01:03/DPB1*04:01 (1.3), HLA-DPA1*02:01/DPB1*01:01 (1.5), HLA-DPA1*03:01/DPB1*04:02 (2.2), HLA-DRB3*01:01 (6.5), HLA-DPA1*02:01/DPB1*05:01 (7.5), HLA-DRB1*04:05 (11), HLA-DRB1*12:01 (16.5), HLA-DRB1*15:01 (22), HLA-DQA1*05:01/DQB1*02:01 (26), HLA-DRB1*08:02 (29), HLA-DRB4*01:01 (31), HLA-DRB1*04:01 (32), HLA-DPA1*02:01/DPB1*14:01 (41), HLA-DQA1*03:01/DQB1*03:02 (44), HLA-DRB1*09:01 (44), HLA-DQA1*04:01/DQB1*04:02 (47), HLA-DRB1*07:01 (47), HLA-DRB1*11:01 (47)
5.	SRIKSIKNLSDLTFL	HLA-DRB1*04:05 (0.43), HLA-DRB1*04:01 (2.1), HLA-DRB3*02:02 (8.8), HLA-DRB4*01:01 (13), HLA-DRB1*12:01 (14.7), HLA-DPA1*02:01/DPB1*14:01 (18), HLA-DRB1*11:01 (18), HLA-DRB1*15:01 (19), HLA-DPA1*01:03/DPB1*04:01 (20), HLA-DRB1*08:02 (26), HLA-DRB1*13:02 (26), HLA-DPA1*02:01/DPB1*05:01 (27), HLA-DRB5*01:01 (32), HLA-DRB1*01:01 (33), HLA-DRB1*07:01 (35), HLA-DQA1*01:01/DQB1*05:01 (37), HLA-DRB1*03:01 (37), HLA-DPA1*03:01/DPB1*04:02 (41), HLA-DPA1*02:01/DPB1*01:01 (43), HLA-DRB3*01:01 (43), HLA-DRB1*09:01 (45), HLA-DQA1*05:01/DQB1*02:01 (46)
6.	HHIYRLRNSTALKTF	HLA-DRB1*01:01 (0.27), HLA-DRB1*04:01 (0.35), HLA-DRB3*02:02 (0.52), HLA-DPA1*02:01/DPB1*14:01 (0.91), HLA-DRB1*07:01 (1.2), HLA-DRB1*08:02 (1.7), HLA-DRB1*13:02 (2.7), HLA-DRB1*04:05 (4.6), HLA-DRB1*09:01 (5), HLA-DRB1*15:01 (6.8), HLA-DRB5*01:01 (7.5), HLA-DRB1*12:01 (12), HLA-DPA1*02:01/DPB1*05:01 (13), HLA-DRB1*11:01 (13), HLA-DRB3*01:01 (15), HLA-DRB4*01:01 (15), HLA-DRB1*03:01 (17), HLA-DQA1*01:02/DQB1*06:02 (18), HLA-DPA1*01:03/DPB1*04:01 (28), HLA-DPA1*02:01/DPB1*01:01 (28), HLA-DPA1*03:01/DPB1*04:02 (30), HLA-DPA1*01:03/DPB1*02:01 (34), HLA-DQA1*05:01/DQB1*03:01 (42)
7.	NVTFRLMLNKCCTRH	HLA-DRB1*11:01 (2.5), HLA-DRB5*01:01 (3.8), HLA-DRB1*04:05 (5.6), HLA-DRB1*04:01 (7.4), HLA-DRB1*01:01 (9.7), HLA-DRB3*02:02 (11), HLA-DRB1*15:01 (13), HLA-DRB1*08:02 (20), HLA-DRB1*12:01 (23), HLA-DRB1*03:01 (24), HLA-DPA1*03:01/DPB1*04:02 (32), HLA-DPA1*01:03/DPB1*04:01 (38), HLA-DPA1*02:01/DPB1*14:01 (40), HLA-DRB1*13:02 (43), HLA-DQA1*01:02/DQB1*06:02 (45)
8.	LKHIEKNYPASLHKF	HLA-DRB1*13:02 (0.75), HLA-DRB3*02:02 (2), HLA-DRB1*09:01 (5.3), HLA-DRB1*01:01 (8.5), HLA-DRB1*07:01 (21), HLA-DRB1*03:01 (22), HLA-DRB5*01:01 (23), HLA-DPA1*02:01/DPB1*14:01 (25), HLA-DRB4*01:01 (29), HLA-DRB3*01:01 (30), HLA-DRB1*15:01 (31), HLA-DRB1*11:01 (36), HLA-DPA1*02:01/DPB1*05:01 (37), HLA-DRB1*12:01 (45.5), HLA-DRB1*08:02 (48), HLA-DQA1*05:01/DQB1*03:01 (49)
9.	SYQNFINNFSCLIKK	HLA-DRB3*02:02 (0.19), HLA-DRB1*13:02 (0.5), HLA-DRB1*04:05 (0.81), HLA-DRB1*04:01 (2.5), HLA-DRB3*01:01 (7.8), HLA-DPA1*01:03/DPB1*04:01 (9.3), HLA-DRB1*15:01 (9.4), HLA-DPA1*01:03/DPB1*02:01 (11), HLA-DPA1*03:01/DPB1*04:02 (12), HLA-DRB1*11:01 (12), HLA-DRB1*07:01 (13), HLA-DRB1*01:01 (14), HLA-DRB1*12:01 (17.75), HLA-DQA1*01:01/DQB1*05:01 (18), HLA-DRB1*08:02 (18), HLA-DPA1*02:01/DPB1*05:01 (20), HLA-DRB5*01:01 (20), HLA-DPA1*02:01/DPB1*01:01 (24), HLA-DPA1*02:01/DPB1*14:01 (28), HLA-DQA1*05:01/DQB1*02:01 (28), HLA-DQA1*01:02/DQB1*06:02 (34), HLA-DRB1*09:01 (39), HLA-DRB1*03:01 (41)

Table 3 (continued)

Serial no.	Epitope	Interacting MHC-II allele based on percentile rank < 50
10.	RLAWTIGNRAPYIGS	HLA-DRB1*13:02 (1.9), HLA-DRB3*02:02 (2), HLA-DRB1*09:01 (10), HLA-DPA1*02:01/DPB1*14:01 (11), HLA-DQA1*05:01/DQB1*03:01 (23), HLA-DRB3*01:01 (23), HLA-DRB1*08:02 (25), HLA-DRB1*07:01 (26), HLA-DRB1*12:01 (26.5), HLA-DRB1*11:01 (27), HLA-DQA1*01:01/DQB1*05:01 (28), HLA-DRB1*03:01 (28), HLA-DPA1*01:03/DPB1*04:01 (32), HLA-DRB1*15:01 (32), HLA-DRB1*01:01 (44), HLA-DPA1*02:01/DPB1*01:01 (45), HLA-DQA1*01:02/DQB1*06:02 (48), HLA-DQA1*04:01/DQB1*04:02 (49)
11.	DVANFLRAYSWSDVL	HLA-DRB1*04:05 (0.73),HLA-DRB1*15:01 (3.1), HLA-DPA1*01:03/DPB1*04:01 (7.1), HLA-DQA1*01:01/DQB1*05:01 (8), HLA-DRB1*08:02 (8.7), HLA-DRB3*02:02 (9.2),HLA-DRB1*07:01 (12), HLA-DRB5*01:01 (13), HLA-DRB1*04:01 (14), HLA-DRB1*11:01 (16), HLA-DQA1*05:01/DQB1*02:01 (17), HLA-DPA1*01:03/DPB1*02:01 (18), HLA-DRB1*09:01 (18), HLA-DPA1*02:01/DPB1*14:01 (20), HLA-DRB1*01:01 (26), HLA-DQA1*04:01/DQB1*04:02 (28), HLA-DRB1*13:02 (28), HLA-DRB1*12:01 (30.5), HLA-DQA1*03:01/DQB1*03:02 (31), HLA-DQA1*01:02/DQB1*06:02 (34), HLA-DRB3*01:01 (34), HLA-DRB4*01:01 (37), HLA-DPA1*03:01/DPB1*04:02 (38), HLA-DPA1*02:01/DPB1*01:01 (41), HLA-DPA1*02:01/DPB1*05:01 (43), HLA-DQA1*05:01/DQB1*03:01 (45)
12.	APSYRNFSFSLKEKE	HLA-DPA1*02:01/DPB1*05:01 (1.2), HLA-DPA1*02:01/DPB1*01:01 (3.8), HLA-DPA1*01:03/DPB1*02:01 (6.7), HLA-DPA1*01:03/DPB1*04:01 (7.9), HLA-DRB1*11:01 (10), HLA-DRB1*09:01 (12), HLA-DPA1*02:01/DPB1*14:01 (16), HLA-DRB1*04:05 (16), HLA-DRB3*02:02 (16), HLA-DRB5*01:01 (18), HLA-DPA1*03:01/DPB1*04:02 (26), HLA-DQA1*03:01/DQB1*03:02 (27), HLA-DQA1*01:01/DQB1*05:01 (29), HLA-DRB1*15:01 (29), HLA-DRB3*01:01 (30), HLA-DRB1*01:01 (31), HLA-DRB1*07:01 (35), HLA-DQA1*04:01/DQB1*04:02 (37), HLA-DRB1*13:02 (37), HLA-DRB1*04:01 (40), HLA-DRB1*03:01 (42)
13.	LGRVSRSTLSLSLNV	HLA-DPA1*02:01/DPB1*14:01 (0.8), HLA-DRB1*07:01 (3.8), HLA-DPA1*02:01/DPB1*01:01 (9.7), HLA-DQA1*01:02/DQB1*06:02 (9.8), HLA-DRB3*02:02 (15), HLA-DPA1*02:01/DPB1*05:01 (18), HLA-DPA1*01:03/DPB1*04:01 (20), HLA-DRB1*09:01 (20), HLA-DPA1*03:01/DPB1*04:02 (21), HLA-DRB1*08:02 (21), HLA-DRB1*13:02 (25),HLA-DRB4*01:01 (25), HLA-DRB1*03:01 (28), HLA-DRB1*04:01 (28), HLA-DRB1*15:01 (32), HLA-DPA1*01:03/DPB1*02:01 (33), HLA-DRB1*01:01 (33), HLA-DQA1*05:01/DQB1*03:01 (34), HLA-DRB1*11:01 (36), HLA-DRB1*04:05 (46)
14.	VELKTKLKLKSSVMG	HLA-DRB1*08:02 (1.1), HLA-DRB1*11:01 (2.4), HLA-DRB1*12:01 (4.85), HLA-DPA1*02:01/DPB1*05:01 (8.7), HLA-DRB1*15:01 (9.9), HLA-DRB1*01:01 (11), HLA-DPA1*02:01/DPB1*14:01 (18), HLA-DRB4*01:01 (18), HLA-DRB1*09:01 (22), HLA-DPA1*03:01/DPB1*04:02 (25), HLA-DRB1*03:01 (25), HLA-DRB1*07:01 (27), HLA-DRB5*01:01 (30), HLA-DPA1*02:01/DPB1*01:01 (34), HLA-DRB1*04:05 (35), HLA-DRB1*13:02 (35), HLA-DRB1*04:01 (37), HLA-DRB3*02:02 (37), HLA-DPA1*01:03/DPB1*04:01 (48), HLA-DQA1*01:02/DQB1*06:02 (48)
15.	LSIFIKDRATAVNQE	HLA-DRB3*01:01 (0.07), HLA-DRB3*02:02 (0.32), HLA-DRB1*13:02 (1.3), HLA-DRB1*08:02 (2.8), HLA-DRB1*04:01 (3.9), HLA-DRB1*03:01 (5.2), HLA-DQA1*04:01/DQB1*04:02 (7.6), HLA-DQA1*03:01/DQB1*03:02 (8.5), HLA-DRB1*11:01 (12), HLA-DRB4*01:01 (13), HLA-DRB1*07:01 (14), HLA-DPA1*02:01/DPB1*14:01 (16), HLA-DRB1*01:01 (17), HLA-DRB1*12:01 (18.75), HLA-DRB5*01:01 (20), HLA-DPA1*03:01/DPB1*04:02 (24), HLA-DRB1*15:01 (25), HLA-DRB1*09:01 (28), HLA-DRB1*04:05 (32), HLA-DQA1*01:01/DQB1*05:01 (37), HLA-DPA1*01:03/DPB1*02:01 (39), HLA-DPA1*02:01/DPB1*01:01 (39), HLA-DQA1*01:02/DQB1*06:02 (41), HLA-DPA1*01:03/DPB1*04:01 (42), HLA-DQA1*05:01/DQB1*02:01 (45), HLA-DPA1*02:01/DPB1*05:01 (46), HLA-DQA1*05:01/DQB1*03:01 (48)
16.	NRWKSWFSYIDALDD	HLA-DRB1*04:05 (0.84), HLA-DQA1*01:01/DQB1*05:01 (1.2), HLA-DQA1*05:01/DQB1*02:01 (1.7), HLA-DPA1*01:03/DPB1*04:01 (3.4), HLA-DQA1*03:01/DQB1*03:02 (5.1), HLA-DPA1*01:03/DPB1*02:01 (6.9), HLA-DPA1*02:01/DPB1*01:01 (8), HLA-DQA1*04:01/DQB1*04:02 (11), HLA-DPA1*02:01/DPB1*05:01 (12), HLA-DPA1*03:01/DPB1*04:02 (13), HLA-DRB1*04:01 (16), HLA-DRB1*07:01 (18), HLA-DPA1*02:01/DPB1*14:01 (23), HLA-DRB1*11:01 (24), HLA-DRB3*01:01 (25), HLA-DRB3*02:02 (36), HLA-DRB5*01:01 (41), HLA-DRB1*01:01 (48)
17.	KYNLAFRYEFTRHFI	HLA-DRB3*01:01 (0.63), HLA-DPA1*01:03/DPB1*04:01 (2.3), HLA-DRB5*01:01 (4), HLA-DPA1*01:03/DPB1*02:01 (4.6), HLA-DRB1*04:01 (7.1), HLA-DPA1*02:01/DPB1*05:01 (8), HLA-DRB1*11:01 (8.7), HLA-DRB1*07:01 (8.8), HLA-DRB1*04:05 (9), HLA-DPA1*02:01/DPB1*14:01 (11), HLA-DQA1*01:01/DQB1*05:01 (17), HLA-DRB3*02:02 (17), HLA-DRB1*08:02 (18), HLA-DRB1*15:01 (20), HLA-DPA1*02:01/DPB1*01:01 (21), HLA-DRB1*03:01 (21), HLA-DPA1*03:01/DPB1*04:02 (27), HLA-DQA1*05:01/DQB1*02:01 (32), HLA-DRB1*01:01 (32), HLA-DRB1*09:01 (33), HLA-DRB1*12:01 (38)

Table 3 (continued)

Serial no.	Epitope	Interacting MHC-II allele based on percentile rank < 50
18.	VPFHSIWDHILTSIQ	HLA-DRB1*04:05 (2.7), HLA-DQA1*01:01/DQB1*05:01 (7.6), HLA-DPA1*01:03/DPB1*04:01 (7.9), HLA-DPA1*01:03/DPB1*02:01 (15), HLA-DRB1*04:01 (15), HLA-DRB5*01:01 (15), HLA-DRB3*01:01 (16), HLA-DPA1*03:01/DPB1*04:02 (17), HLA-DRB1*07:01 (21), HLA-DPA1*02:01/DPB1*01:01 (24), HLA-DPA1*02:01/DPB1*14:01 (26), HLA-DPA1*02:01/DPB1*05:01 (27), HLA-DRB3*02:02 (27), HLA-DQA1*05:01/DQB1*02:01 (30), HLA-DRB1*01:01 (30), HLA-DRB1*09:01 (31), HLA-DRB1*13:02 (35), HLA-DRB1*11:01 (36), HLA-DRB1*12:01 (36), HLA-DRB1*15:01 (37), HLA-DRB4*01:01 (40), HLA-DRB1*08:02 (48)
19.	HHIYRLRNSTALKTF	HLA-DRB1*01:01 (0.27), HLA-DRB1*04:01 (0.35), HLA-DRB3*02:02 (0.52), HLA-DPA1*02:01/DPB1*14:01 (0.91), HLA-DRB1*07:01 (1.2), HLA-DRB1*08:02 (1.7), HLA-DRB1*13:02 (2.7), HLA-DRB1*04:05 (4.6), HLA-DRB1*09:01 (5), HLA-DRB1*15:01 (6.8), HLA-DRB5*01:01 (7.5), HLA-DRB1*12:01 (12), HLA-DPA1*02:01/DPB1*05:01 (13), HLA-DRB1*11:01 (13), HLA-DRB3*01:01 (15), HLA-DRB4*01:01 (15), HLA-DRB1*03:01 (17), HLA-DQA1*01:02/DQB1*06:02 (18), HLA-DPA1*01:03/DPB1*04:01 (28), HLA-DPA1*02:01/DPB1*01:01 (28), HLA-DPA1*03:01/DPB1*04:02 (30), HLA-DPA1*01:03/DPB1*02:01 (34), HLA-DQA1*05:01/DQB1*03:01 (42)
20.	LPYRVRNVQTLAEAL	HLA-DRB1*04:05 (1.3), HLA-DRB1*08:02 (5.5), HLA-DRB3*02:02 (5.9), HLA-DQA1*03:01/DQB1*03:02 (7.6), HLA-DPA1*02:01/DPB1*14:01 (7.7), HLA-DRB1*07:01 (13), HLA-DQA1*01:02/DQB1*06:02 (15), HLA-DRB1*13:02 (17), HLA-DQA1*05:01/DQB1*02:01 (19), HLA-DQA1*04:01/DQB1*04:02 (20), HLA-DRB1*09:01 (21), HLA-DRB3*01:01 (24), HLA-DRB4*01:01 (27), HLA-DRB1*11:01 (28), HLA-DRB5*01:01 (28), HLA-DPA1*03:01/DPB1*04:02 (30), HLA-DRB1*01:01 (31), HLA-DPA1*02:01/DPB1*01:01 (32), HLA-DRB1*04:01 (37), HLA-DPA1*02:01/DPB1*05:01 (39), HLA-DPA1*01:03/DPB1*04:01 (41), HLA-DPA1*01:03/DPB1*02:01 (42), HLA-DRB1*03:01 (44), HLA-DQA1*01:01/DQB1*05:01 (47), HLA-DRB1*12:01 (48.5)
21.	WEICARAWLEDSDGA	HLA-DQA1*01:01/DQB1*05:01 (6.2), HLA-DQA1*03:01/DQB1*03:02 (11), HLA-DQA1*04:01/DQB1*04:02 (25), HLA-DPA1*01:03/DPB1*02:01 (34), HLA-DQA1*05:01/DQB1*02:01 (34), HLA-DRB3*01:01 (43)
22.	LSSESESTINLLPYD	HLA-DQA1*03:01/DQB1*03:02 (5.8), HLA-DQA1*01:02/DQB1*06:02 (18), HLA-DQA1*01:01/DQB1*05:01 (19), HLA-DPA1*03:01/DPB1*04:02 (26), HLA-DQA1*04:01/DQB1*04:02 (26), HLA-DPA1*02:01/DPB1*01:01 (31), HLA-DQA1*05:01/DQB1*02:01 (32), HLA-DPA1*01:03/DPB1*02:01 (41), HLA-DRB1*12:01 (47), HLA-DPA1*02:01/DPB1*05:01 (48)
23.	LQNCSILTVPFHSIW	HLA-DPA1*01:03/DPB1*04:01 (1.4), HLA-DPA1*03:01/DPB1*04:02 (4.9), HLA-DPA1*01:03/DPB1*02:01 (5.7), HLA-DPA1*02:01/DPB1*14:01 (11), HLA-DPA1*02:01/DPB1*01:01 (15), HLA-DRB1*07:01 (15), HLA-DRB1*12:01 (20), HLA-DQA1*01:02/DQB1*06:02 (21), HLA-DRB4*01:01 (24), HLA-DRB1*13:02 (26), HLA-DRB1*11:01 (27), HLA-DRB5*01:01 (27), HLA-DPA1*02:01/DPB1*05:01 (29), HLA-DRB1*08:02 (29), HLA-DRB1*15:01 (29), HLA-DQA1*01:01/DQB1*05:01 (34), HLA-DRB1*01:01 (36), HLA-DRB1*04:01 (36), HLA-DRB3*02:02 (38), HLA-DRB1*04:05 (40), HLA-DQA1*04:01/DQB1*04:02 (41), HLA-DRB3*01:01 (44), HLA-DQA1*05:01/DQB1*03:01 (47), HLA-DRB1*09:01 (47)
24.	SSLPVWLYFPSEGQQ	HLA-DRB1*04:05 (6), HLA-DRB1*04:01 (7.8), HLA-DQA1*01:01/DQB1*05:01 (12), HLA-DPA1*01:03/DPB1*02:01 (17), HLA-DRB1*15:01 (17), HLA-DPA1*02:01/DPB1*01:01 (18), HLA-DPA1*01:03/DPB1*04:01 (19), HLA-DPA1*02:01/DPB1*05:01 (26), HLA-DRB1*11:01 (26), HLA-DRB1*08:02 (31), HLA-DQA1*05:01/DQB1*02:01 (33), HLA-DQA1*03:01/DQB1*03:02 (36), HLA-DRB4*01:01 (37),, HLA-DPA1*03:01/DPB1*04:02 (39), HLA-DRB1*12:01 (44.5), HLA-DQA1*04:01/DQB1*04:02 (46), HLA-DRB5*01:01 (48)
25.	LKQHIVSNSFPSQAE	HLA-DRB3*02:02 (1.7), HLA-DRB1*13:02 (7.5), HLA-DRB1*04:01 (11), HLA-DRB1*08:02 (11), HLA-DRB1*09:01 (14), HLA-DPA1*01:03/DPB1*02:01 (18), HLA-DQA1*03:01/DQB1*03:02 (20), HLA-DPA1*02:01/DPB1*14:01 (21), HLA-DRB1*04:05 (21), HLA-DPA1*01:03/DPB1*04:01 (22), HLA-DRB4*01:01 (22), HLA-DRB1*07:01 (23), HLA-DQA1*04:01/DQB1*04:02 (25), HLA-DRB1*12:01 (30.5), HLA-DQA1*05:01/DQB1*03:01 (31), HLA-DRB1*03:01 (31), HLA-DRB1*11:01 (33), HLA-DQA1*01:02/DQB1*06:02 (35), HLA-DPA1*02:01/DPB1*01:01 (37), HLA-DRB1*15:01 (37), HLA-DRB1*01:01 (47), HLA-DQA1*01:01/DQB1*05:01 (49), HLA-DRB5*01:01 (49)
26.	RLAWTIGNRAPYIGS	HLA-DRB1*13:02 (1.9), HLA-DRB3*02:02 (2), HLA-DRB1*09:01 (10), HLA-DPA1*02:01/DPB1*14:01 (11), HLA-DQA1*05:01/DQB1*03:01 (23), HLA-DRB3*01:01 (23), HLA-DRB1*08:02 (25), HLA-DRB1*07:01 (26), HLA-DRB1*12:01 (26.5), HLA-DRB1*11:01 (27), HLA-DQA1*01:01/DQB1*05:01 (28), HLA-DRB1*03:01 (28), HLA-DPA1*01:03/DPB1*04:01 (32), HLA-DRB1*15:01 (32), HLA-DRB1*01:01 (44), HLA-DPA1*02:01/DPB1*01:01 (45), HLA-DQA1*01:02/DQB1*06:02 (48), HLA-DQA1*04:01/DQB1*04:02 (49)

Table 4 Conservancy analysis of all the epitopes with appropriate length

No.	Epitope	Conservancy (%)	Length	Epitope	Conservancy (%)	Length
1.	QSSLPVWLY	87.23	9	QSSLPVWLYFPSEGQ	87.23	15
2.	YSGNIVHRY	100.00	9	FLPTHYSGNIVHRYN	100.00	15
3.	FLEKEELFY	31.91	9	LEFLEKEELFYILIA	21.28	15
4.	NLFDWMHFL	100.00	9	VKNLFDWMHFLIPLC	100.00	15
5.	NLSDLTFLI	87.23	9	SRIKSIKNLSDLTFL	78.72	15
6.	RLRNSTALK	95.74	9	HHIYRLRNSTALKTF	82.98	15
7.	VTFRLMLNK	82.98	9	NVTFRLMLNKCCTRH	82.98	15
8.	NYPASLHKF	14.89	9	LKHIEKNYPASLHKF	14.89	15
9.	SYQNFINNF	48.94	9	SYQNFINNFSCLIKK	10.64	15
10.	WTIGNRAPY	97.87	9	RLAWTIGNRAPYIGS	97.87	15
11.	DVANFLRAY	93.62	9	DVANFLRAYSWSDVL	78.72	15
12.	APSYRNFSF	100.00	9	APSYRNFSFSLKEKE	100.00	15
13.	RVSRSTLSL	95.74	9	LGRVSRSTLSLSLNV	95.74	15
14.	KLKLKSSVM	97.87	9	VELKTKLKLKSSVMG	97.87	15
15.	FIKDRATAV	95.74	9	LSIFIKDRATAVNQE	80.85	15
16.	NRWKSWFSY	87.23	9	NRWKSWFSYIDALDD	87.23	15
17.	FRYEFTRHF	100.00	9	KYNLAFRYEFTRHFI	100.00	15
18.	FHSIWDHIL	87.23	9	VPFHSIWDHILTSIQ	82.98	15
19.	YRLRNSTAL	95.74	9	HHIYRLRNSTALKTF	82.98	15
20.	YRVRNVQTL	100.00	9	LPYRVRNVQTLAEAL	100.00	15
21.	WEICARAWL	87.23	9	WEICARAWLEDSDGA	85.11	15
22.	SESESTINL	14.89	9	LSSESESTINLLPYD	10.64	15
23.	LTVPFHSIW	97.87	9	LQNCSILTVPFHSIW	97.87	15
24.	SSLPVWLYF	87.23	9	SSLPVWLYFPSEGQQ	87.23	15
25.	KQHIVSNSF	8.51	9	LKQHIVSNSFPSQAE	8.51	15
26.	WTIGNRAPY	97.87	9	VNRLAWTIGNRAPYI	34.04	15

residues-Tyr-30 (B), Asn-62 (A), Asp-66 (B), Gln-70 (B), Lys-71 (B) and Thr-77 (B) form hydrogen bond (Fig. 7) with the 15.0-mer. In addition, the control docking energy for MHC-I interaction was found to be − 9.6 kcal/mol and is shown in Figure S2.

Allergenicity assessment and B-cell epitope detection

AlgPred predicted allergenicity of the epitope depending on amino acid structure. The predictive score of AlgPred for the two combined epitopes was − 0.51324781, with a cutoff value − 0.4.

The linear peptide with 15-mer form (HHIYRLRN-STALKTF) through the sequence-based methodology, the B-cell epitope calculation was accomplished and the different prediction parameters were considered and the values ranging from − 0.929 to 3.214. The cutoff values for these predictions were ranging from 0.500 to 1.390 (Fig. 8). The maximum antigenicity score of the peptide calculating through Kolaskar and Tongaonkar antigenicity scale was 1.074 (Fig. 8a). Accessibility of peptide surfaces is another

crucial benchmark for satisfying the criteria of a prospective B-cell epitope and here the Emini surface accessibility calculation supports the prediction with a maximum value of 1.871 (Fig. 8d). The Parker hydrophilicity prediction was also employed with a maximum score of 3.214 and is shown in Fig. 8f. Additionally, the maximum score of the Chou and Fasman beta turn estimation score was 1.083 (Fig. 8c), the flexibility estimation score of Karplus and Schulz was 1.066 (Fig. 8e), and the Bepipred linear epitope prediction analysis was 0.601 (Fig. 8b).

Discussion

Although the epitope-based design of vaccines has become a conjoint method, no substantial research has yet been done in the case of the Marburg virus L-protein. Marburg virus genome is composed of ribonucleic acid rather than deoxyribonucleic acid. It is especially challenging to develop vaccines for RNA viruses due to rapid mutations of various surface proteins (Twiddy et al. 2003). Therefore, targeting transcription or replication machinery is the possible
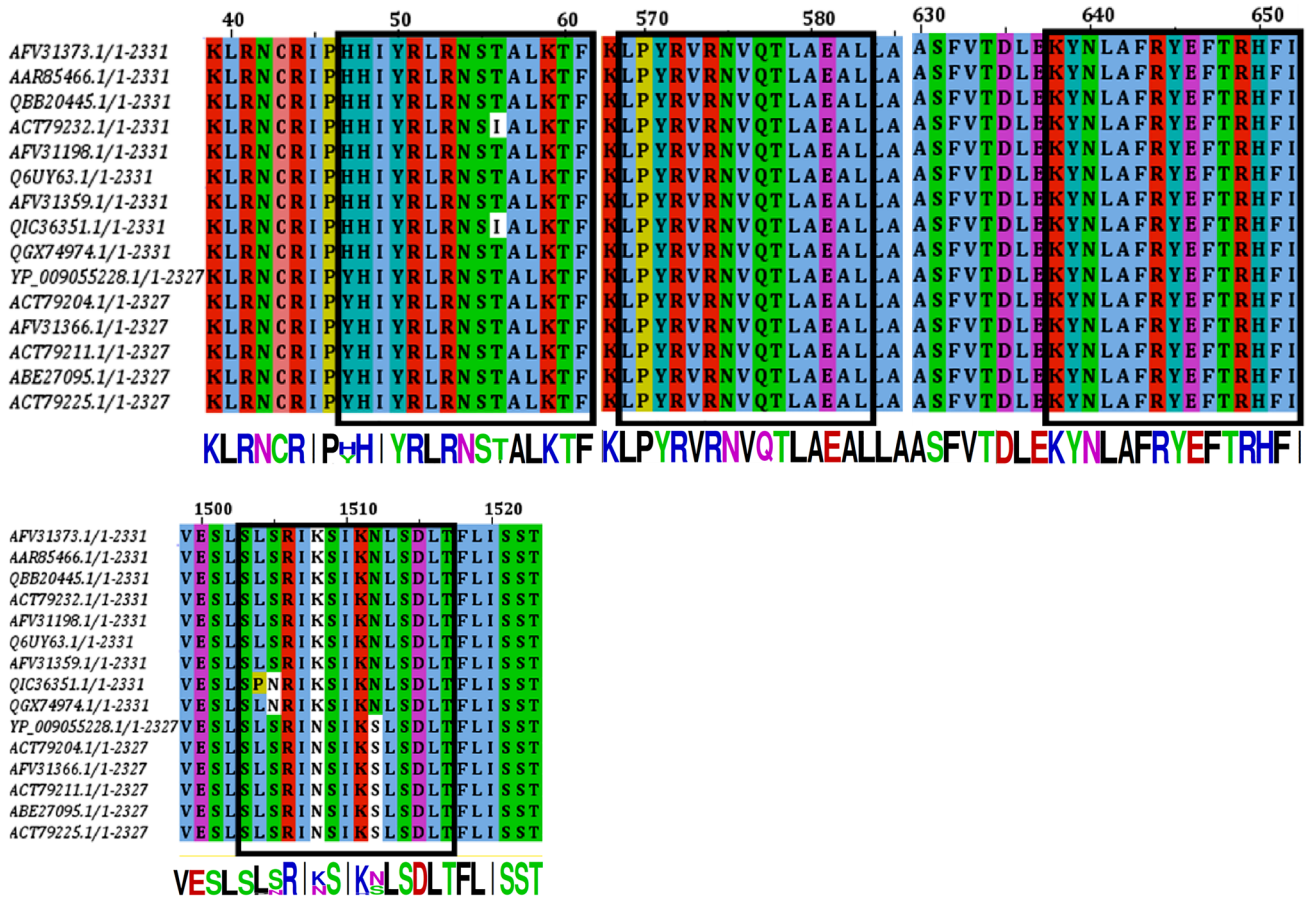
Fig. 3 MSA-based location identification of the different epitopes within the L-proteins of Marburg virus. Here the epitopes are shown by boxed (black) region within the MSA. The conservancies of the sequences are displayed by sequence to logo at the bottom of the total MSA. Here, only the best four epitopes, 15 mer including 9 mer, are shown

Fig. 4 Population coverage analysis for the top predicted epitope (YRLRNSTAL) based on the HLA interaction. Here, the whole world populations are assessed for the proposed epitope. The combined prediction for both of the MHC has been shown. Here, the number 1 bar for all the analyses represents out-predicted epitope. 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
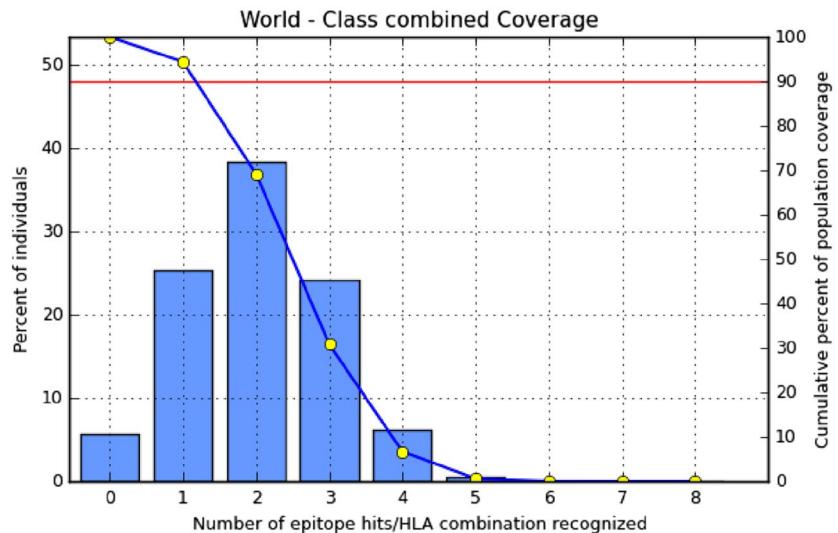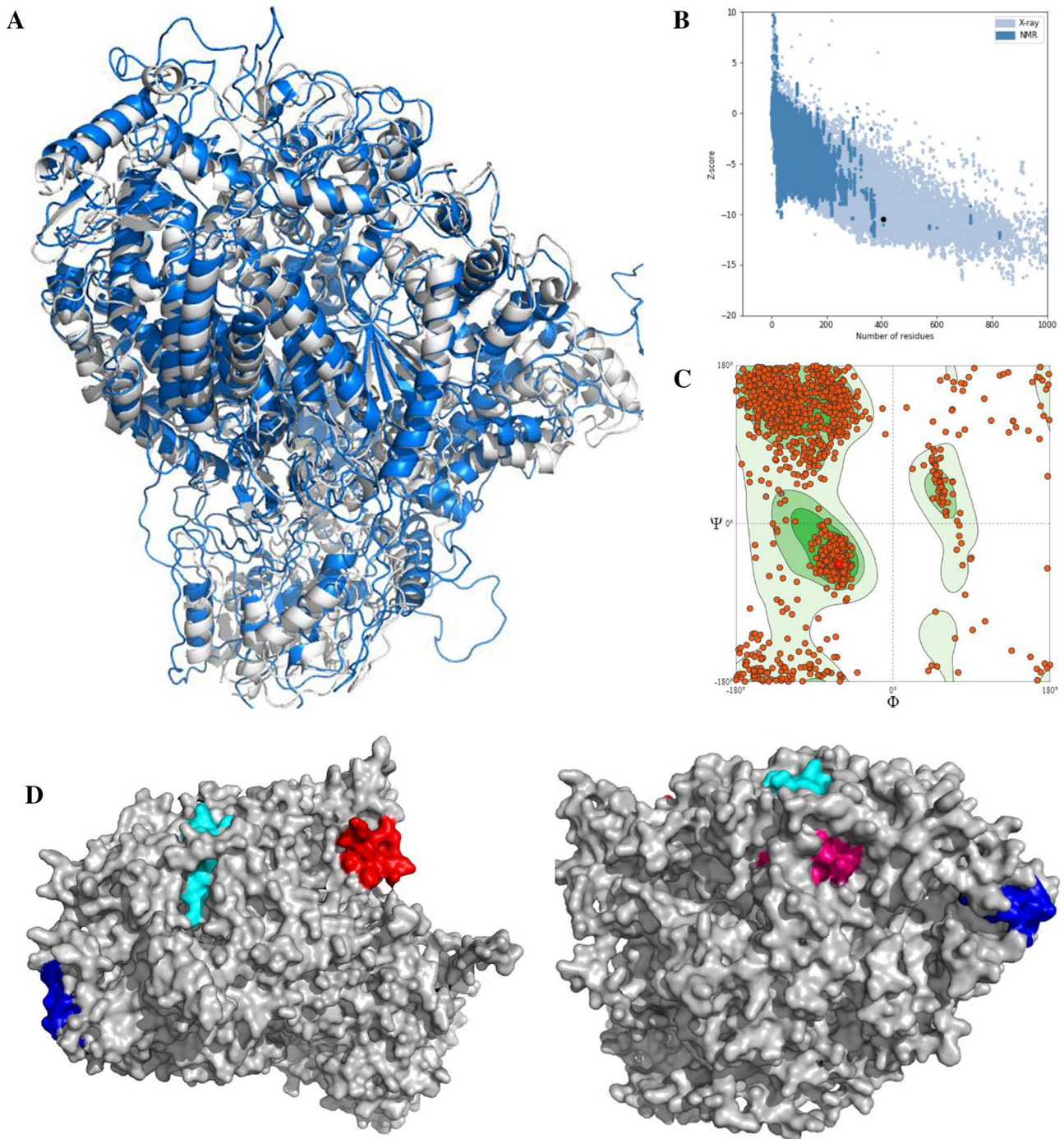
Fig. 5 The three-dimensional model of L protein with structural validation and the superficial localities of the predicted epitopes. **a** The predicted model superposed view (grey color represents the model and the blue one is the template). **b** The ProSA Z score for model quality estimation and **c** the Ramachandran plot of the pre-dicted model. In **d** the proposed epitopes SRIKSIKNLSDLTFL (red), KYNLAFRYEFTRHFI (cyan), HHIYRLRNSTALKTF (blue), and LPYRVRNVQTLAEAL (pink), having 9-mer core epitope NLSDLT-FLI, FRYEFTRHF, YRLRNSTAL, and YRVRNVQTL, respectively are shown

way to establish effective antiviral therapies against RNA viruses like the Marburg virus. Scientists have discovered that L-protein is a significant cellular element for Marburg

virus genome transcription and duplication. Once a cell is infected by the Marburg virus, its genetic RNA code goes into the cell accompanied by L-protein. The L-protein usually
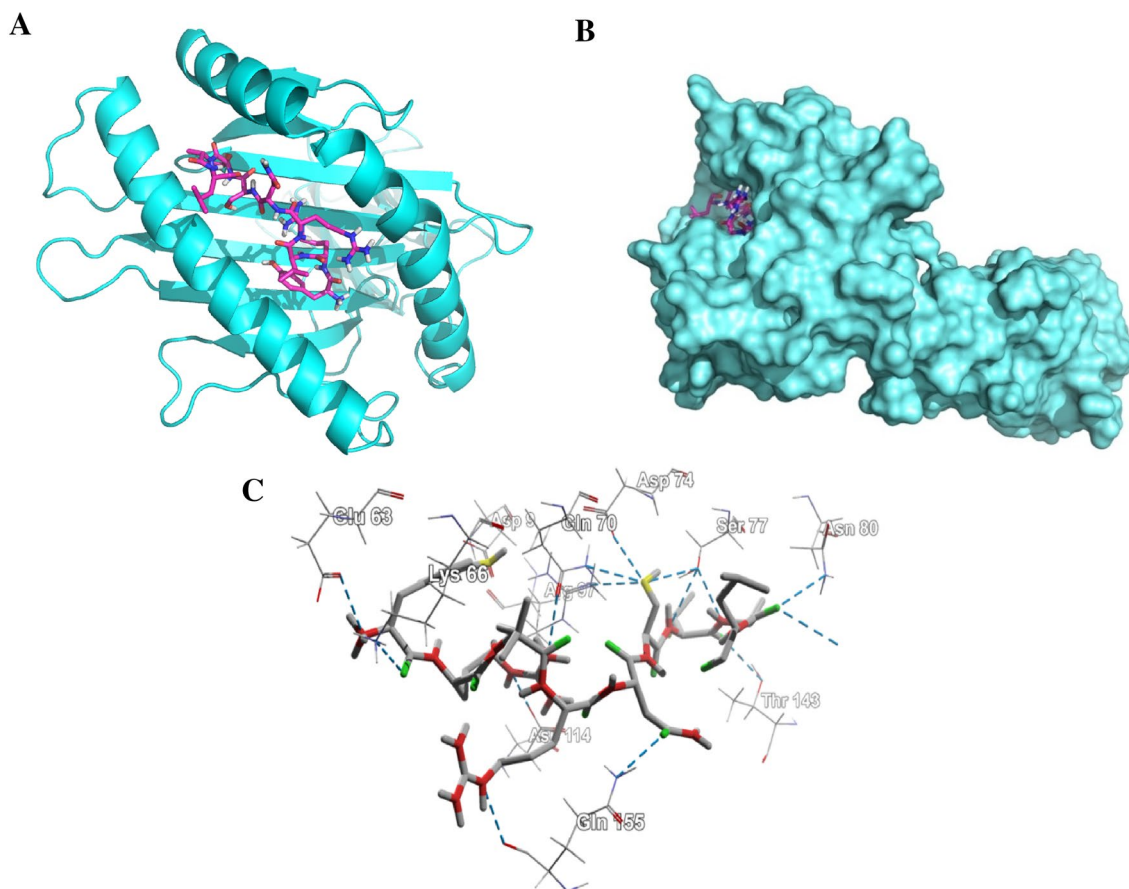
Fig. 6 Docking analysis of the predicted epitope YRLRNSTAL and HLA-C*07:02. **a** Representing the cartoon view. **b** Representing the surface view of the interaction and convincing the perfect binding. **c** Representing the interacted amino-acid residues with the peptide

"reads" the RNA code and produces messenger RNA that creates viral proteins along with duplication and the development of additional viral constituents. For the above mentioned significant contributions, the L-protein was selected to scheme the utmost plausible epitopes by computational methodologies.

Sequence-based bioinformatics methods have been used to model epitopes of both B cells and T cells to impart immunity in distinct manners. Many vaccines are currently constructed on immunity from B cells; but recently vaccinations depending on the T-cell are widely adopted. Antigenic drift can quickly circumvent antibody mediated reactions over durations, whereas cellular immunity also delivers long-term protection (Bacchetta et al. 2005; Igietseme et al. 2004). The cytotoxic CD8+T lymphocytes (CTL) prevent the dissemination of contagious agents through detecting and killing affected cells or releasing exclusive cytokines specific for virus (Oany et al. 2015b; Shrestha and Diamond 2004). Thus, vaccination based on T-cell is a distinctive method for obtaining a robust protective response counter to viruses (Oany et al. 2014; Klein et al. 2005).

From the phylogenetic analysis, it was quite clear that all the different L-proteins have different origins and having diverse groups in the phylogenetic tree (Fig. 2). We have also identified the divergence among the sequences (Figure S1). From that point of view, we selected the most antigenic protein, AFV31373.1, which suggested its ability to elicit a potent immune response.

NetCTL server was used to find 48 epitopes from 12 MHC supertypes (Table 1), and initially, we selected 26 to activate the T-cell response. The epitopes NLSDLTFLI, FRYEFTRHF, YRLRNSTAL, and YRVRNVQTL are principally chosen to scheme a vaccine based on the primary analysis, including the attraction with MHC class I and combined score processed by NetCTL server (Tables 1, 2). The conservancy is an epitope's one of the most critical criteria for selecting it for the production of vaccines (Oany et al. 2014). For our proposed epitopes, study showed conservancy of 87.23, 100, 94.74, and 100 percent, respectively, among all available sequences (Table 4). The locations of the four epitopes anticipated are shown in MSA of L-proteins in Fig. 3.
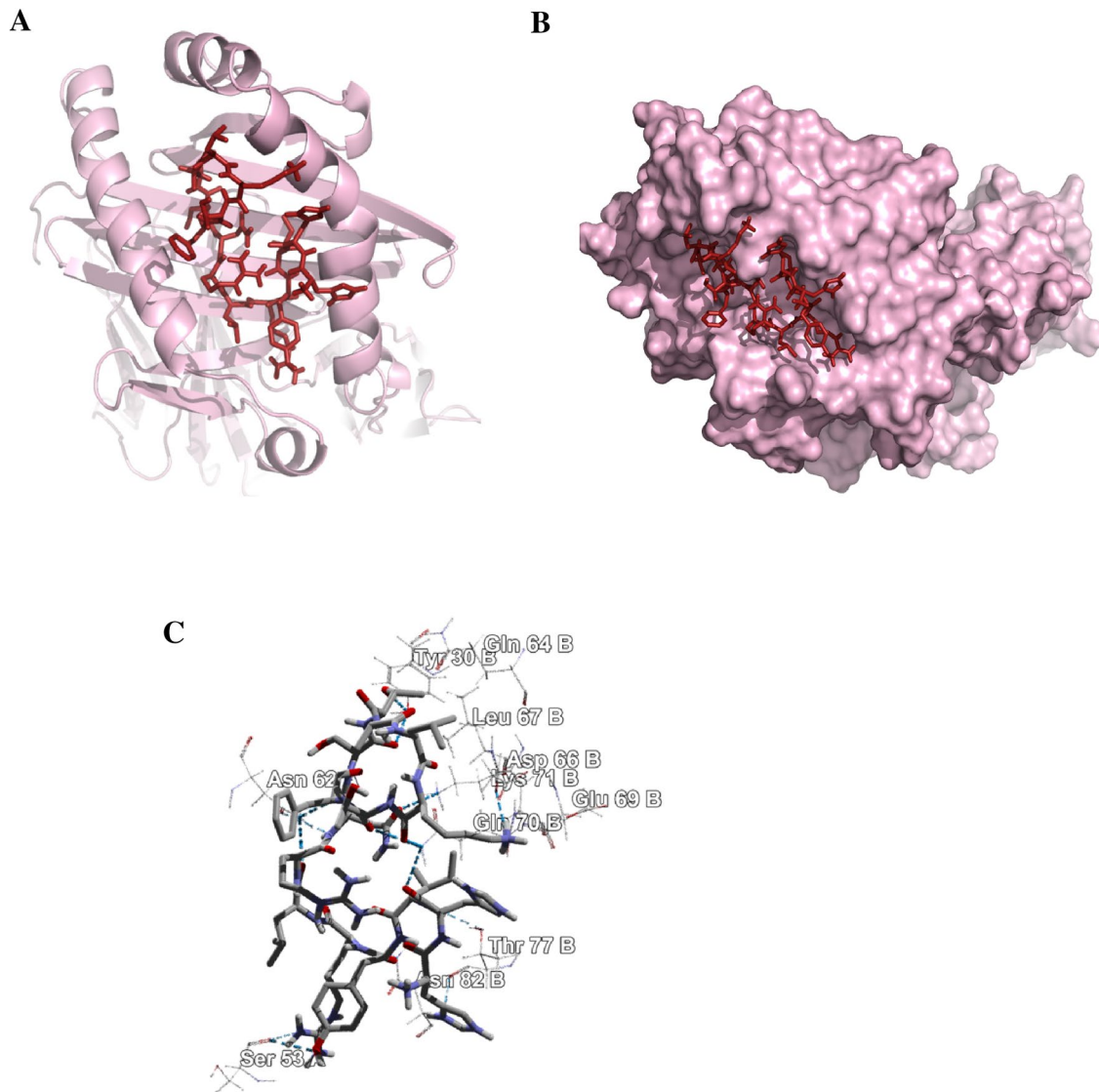
Fig. 7 Docking analysis of the predicted epitope HHIYRLRN-STALKTF and HLA-DRB1*04:01 allele. **a** Representing the cartoon view. **b** Representing the surface view of the interaction and convinc-ing the perfect binding. **c** Representing the interacted amino-acid residues with the peptide

Vaccine candidates should have superior population coverage to achieve acceptability (Oany et al. 2014). Before designing, that is very important. In our analysis, we found that the combined population coverage of our proposed epitopes was 93.48, 93.18, 94.42, and 91.33%, respectively. This outcome revealed that the proposed epitopes have broader in vitro coverage; therefore, they would be supreme candidates for consideration of vaccines.

For visualizing the precise position of the anticipated epitopes, the tertiary structure of the targeted protein was generated and authenticated by Ramachandran Plot (Fig. 5), whereby 90.51% of amino acid molecules were found within the preferred zone. The Z scores from the ProSA server for the protein were − 8.96, which also support the validity of the predicted models, as the values were within the plot and close to zero. That supports reasonably good model, as there is currently no available crystal structure for the L-protein of MABV. As the epitopes were present on the model's surface (Fig. 5), the chance of interaction with the immune system will be improved as early as possible (Oany et al. 2014, 2017).

Lastly, epitopes YRLRNSTAL and YRVRNVQTL (15.0-mer length, HHIYRLRNSTALKTF and LPYRVRNVQT-LAEAL, respectively) were identified as being the utmost
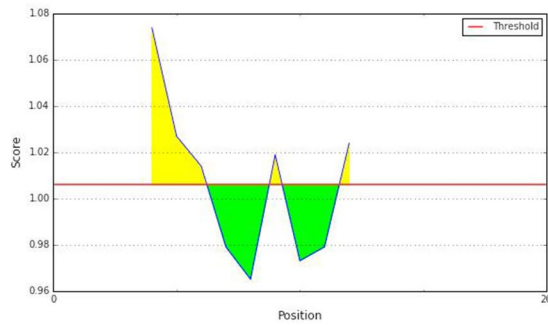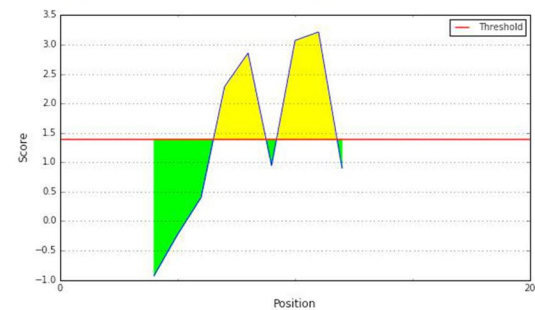
Fig. 8 B-cell epitope prediction. **a** Kolaskar and Tongaonkar antigenicity prediction of the proposed epitope with a threshold value of 1.06. **b** Bepipred linear epitope prediction of the proposed epitope, with a threshold value of 0.50. **c** Chou and Fasman beta-turn prediction of the epitope, with a threshold of 0.98. **d** Emini surface accessibility prediction of the epitope, with a threshold of 1.1. **e** Karplus and Schulz flexibility prediction of the epitope, with a threshold of 1.1. **f** Parker hydrophilicity prediction of the epitope, with a threshold of 3.2. The x-axis and y-axis represent the sequence position and antigenic propensity; respectively. The regions above the threshold are antigenic (desired), shown in yellow
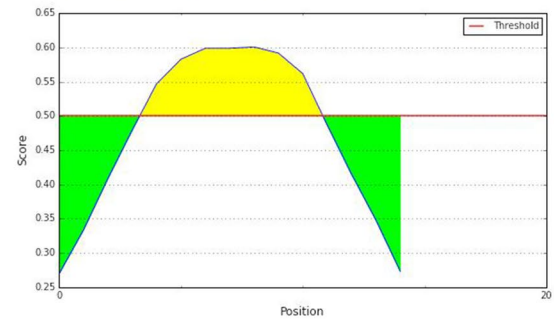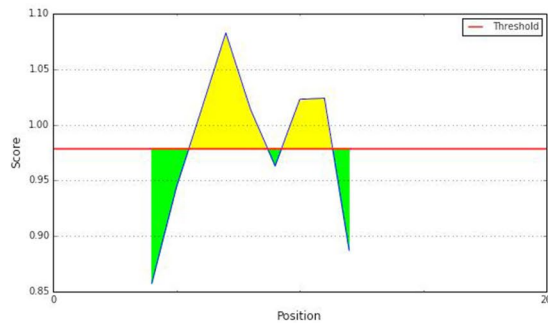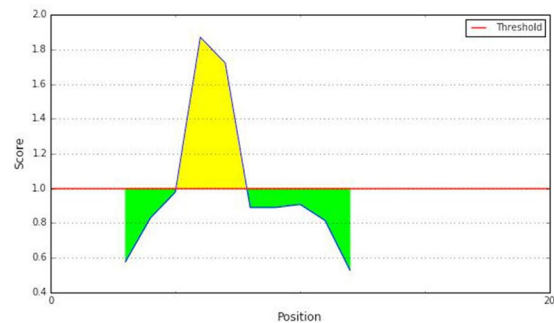
potent and greatly interrelated HLA aspirants for class II MHC molecules (Table 3). Our suggested epitopes are also of a non-allergenic type as per the FAO/WHO Allergenicity Evaluation System, which is another vital peptide vaccine criterion (Oany et al. 2014, 2015b; McKeever et al. 2004).

On the basis of the combined score (2.546), the core epitope YRLRNSTAL would be the paramount epitope applicant, thus further exposed to binding aptitude study. The docking study ensures accuracy with a fairly high binding score and the correctly directed interfaces between the MHC and the predicted epitopes. In addition, comparative research with the empirically observed peptide-MHC complex has showed the specificity of our estimation with comparable binding energy and interacted residues (Figs. 6, 7, S2).

Throughout our research, the B-cell epitope and the T-cell epitope were both given preference, which can activate both primary and secondary antibody mediated immune response (Oany et al. 2015a; Foy et al. 1993). Multiple methods of prediction were used to decide the B-cell epitope, taking into account numerous benchmarks of antigenicity, beta-turn, hydrophilicity, surface accessibility, and two other methods. Our proposed epitope complied with all the criteria of the B-cell prediction techniques mentioned earlier (Fig. 8).

In conclusion, we are optimistic from all of the above in silico studies that our recommended epitope will cause an immune response in vitro and in vivo.

Conclusion

Prevention and monitoring during outbreaks of recently-evolving Marburg virus contagions are both necessary and challenging. Through in silico analyses, the laboratory experimental work can be guided by finding the desired solutions with less tests and recurrences of errors, hence saving the researchers' time as well as costs. We used different computational methods in this analysis to identify a possible epitope against the Marburg virus. Our vaccinomics approaches speculate that the selected part of the ʟ-protein is a promising applicant for a peptide vaccine. However, to validate the effectiveness of the deduced amino acid sequences as an epitope vaccine in case of the deadly Marburg virus, further wet laboratory justification is required.

Compliance with ethical standards

References

Arnold K, Bordoli L, Kopp J, Schwede TJB (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22(2):195–201

Bacchetta R, Gregori S, Roncarolo M-G (2005) CD4+ regulatory T cells: mechanisms of induction and effector function. Autoimmun Rev 4(8):491–496

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28(1):235–242

Brauburger K, Hume AJ, Mühlberger E, Olejnik J (2012) Forty-five years of Marburg virus research. Viruses 4(10):1878–1927

Brusic V, Rudy G, Harrison LC (1998) MHCPEP, a database of MHC-binding peptides: update 1997. Nucleic Acids Res 26(1):368–371

Bui H-H, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A (2006) Predicting population coverage of T-cell epitope-based diagnostics and vaccine. BMC Bioinform 7(1):153

Bui H-H, Sidney J, Li W, Fusseder N, Sette A (2007) Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinform 8(1):361

Buus S, Lauemøller S, Worning P, Kesmir C, Frimurer T, Corbet S, Fomsgaard A, Hilden J, Holm A, Brunak SJ (2003) Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. Tissue Antigens 62(5):378–384

Bio-Qiagen CLC (2016) CLC sequence viewer. Aarhus, Denmark

Chosewood LC, Wilson DE (2009) Biosafety in microbiological and biomedical laboratories: US Department of Health and Human Services. Public Health Service, Washington, DC

Chou PY, Fasman GD (1978) Empirical predictions of protein conformation. Annu Rev Biochem 47(1):251–276

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2016) GenBank. Nucleic Acids Rev 44(D1):D67–D72

Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14(6):1188–1190

Emini EA, Hughes JV, Perlow D, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol 55(3):836–839

Feldmann H, Mühlberger E, Randolf A, Will C, Kiley MP, Sanchez A, Klenk H-D (1992) Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. Virus Res 24(1):1–19

Foy TM, Shepherd D, Durie F, Aruffo A, Ledbetter J, Noelle RJ (1993) In vivo CD40-gp39 interactions are essential for thymus-dependent humoral immunity II Prolonged suppression of the humoral immune response by an antibody to the ligand for CD40, gp39. J Exp Med 178(5):1567–1575

Geisbert T, Jahrling PJ (1995) Differentiation of filoviruses by electron microscopy. Virus Res 39(2–3):129–150

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. Electrophoresis 18(15):2714–2723

Hall T, Biosciences I, Carlsbad C (2011) BioEdit: an important software for molecular biology. GERF Bull Biosci 2(1):60–61

Holland J, Domingo EJ (1998) Origin and evolution of viruses. Virus Genes 16(1):13–21

Igietseme JU, Eko FO, He Q, Black CM (2004) Antibody regulation of T-cell immunity: implications for vaccine strategies against intracellular pathogens. Expert Rev Vaccines 3(1):23–34

Karplus P, Schulz GJN (1985) Prediction of chain flexibility in proteins. Naturwissenschaften 72(4):212–213

Klein RS, Lin E, Zhang B, Luster AD, Tollett J, Samuel MA, Engle M, Diamond MS (2005) Neuronal CXCL10 directs CD8+ T-cell recruitment and control of West Nile virus encephalitis. J Virol 79(17):11457–11466

Kolaskar A, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 276(1–2):172–174

Larsen JEP, Lund O, Nielsen MJ (2006) Improved method for predicting linear B-cell epitopes. Immunome Res 2(1):2

Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen MJ (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinform 8(1):424

Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8(4):477–486

Martines RB, Ng DL, Greer PW, Rollin PE, Zaki SR (2015) Tissue and cellular tropism, pathology and pathogenesis of Ebola and Marburg viruses. J Pathol 235(2):153–174

McKeever TM, Lewis SA, Smith C, Hubbard RJ (2004) Vaccination and allergic disease: a birth cohort study. Am J Public Health 94(6):985–989

Mehedi M, Groseth A, Feldmann H, Ebihara HJ (2011) Clinical aspects of Marburg hemorrhagic fever. Future Virol 6(9):1091–1106

Mühlberger E, Sanchez A, Randolf A, Will C, Kiley MP, Klenk H-D, Feldmann HJ (1992) The nucleotide sequence of the L gene of Marburg virus, a filovirus: homologies with paramyxoviruses and rhabdoviruses. Virology 187(2):534–547

Mühlberger E, Lötfering B, Klenk H-D, Becker SJ (1998) Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. J Virol 72(11):8756–8764

Nyakarahuka L, Shoemaker TR, Balinandi S, Chemos G, Kwesiga B, Mulei S, Kyondo J, Tumusiime A, Kofman A, Masiira BJ (2019) Marburg virus disease outbreak in Kween District Uganda, 2017: epidemiological and laboratory findings. PLoS Negl Trop Dis 13(3):e0007257

Oany AR, Emran A-A, Jyoti TP (2014) Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. Drug Des Dev Ther 8:1139

Oany AR, Ahmad SAI, Hossain MU, Jyoti TP (2015a) Identification of highly conserved regions in L-segment of Crimean–Congo hemorrhagic fever virus and immunoinformatic prediction about potential novel vaccine. Adv Appl Bioinform Chem AABC 8:1

Oany AR, Sharmin T, Chowdhury AS, Jyoti TP, Hasan MA (2015b) Highly conserved regions in Ebola virus RNA dependent RNA polymerase may be act as a universal novel peptide vaccine target: a computational approach. In Silico Pharmacol 3(1):7

Oany AR, Pervin T, Mia M, Hossain M, Shahnaij M, Mahmud S, Kibria KMK (2017) Vaccinomics approach for designing potential peptide vaccine by targeting Shigella spp. serine protease autotransporter subfamily protein SigA. J Immunol Res 2017:6412353–6412353

Oany AR, Mia M, Pervin T, Junaid M, Hosen SZ, Moni MA (2020) Design of novel viral attachment inhibitors of the spike glycoprotein (S) of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) through virtual screening and dynamics. Int J Antimicrob Agents 56(6):106177

Olival KJ, Hayman DT (2014) Filoviruses in bats: current knowledge and future directions. Viruses 6(4):1759–1788

Parker J, Guo D, Hodges RJB (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry 25(19):5425–5432

Peterson AT, Lash RR, Carroll DS, Johnson KM (2006) Geographic potential for outbreaks of Marburg hemorrhagic fever. Am J Trop Med Hyg 75(1):9–15

Poland GA, Ovsyannikova IG, Jacobson RM (2009) Application of pharmacogenomics to vaccines. Pharmacogenomics 10(5):837

Rammensee H-G, Bachmann J, Emmerich NPN, Bachor OA, Stevanović SJ (1999) SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50(3–4):213–219

Saha S, Raghava GJ (2006) AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res 34(suppl_2):W202–W209

Šali A, Potterton L, Yuan F, van Vlijmen H, Karplus MJ (1995) Evaluation of comparative protein modeling by MODELLER. Proteins Struct Function Bioinform 23(3):318–326

Scott WR, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Krüger P, van Gunsteren WF (1999) The GROMOS biomolecular simulation program package. J Phys Chem A 103(19):3596–3607

Sette A, Newman M, Livingston B, McKinney D, Sidney J, Ishioka G, Tangri S, Alexander J, Fikes J, Chesnut R (2002) Optimizing vaccine design for cellular processing, MHC binding and TCR recognition. Tissue Antigens 59(6):443–451

Shrestha B, Diamond MS (2004) Role of CD8+ T cells in control of West Nile virus infection. J Virol 78(15):8312–8321

Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tuffery P (2012) PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res 40(W1):W288–W293

Towner JS, Amman BR, Sealy TK, Carroll SAR, Comer JA, Kemp A, Swanepoel R, Paddock CD, Balinandi S, Khristova ML (2009) Isolation of genetically diverse Marburg viruses from Egyptian fruit bats. PLoS Pathog 5(7):e1000536

Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31(2):455–461

Twiddy SS, Holmes EC, Rambaut A (2003) Inferring the rate and timescale of dengue virus evolution. Mol Biol Evol 20(1):122–129

Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters B (2008) A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput Biol 4(4):e1000048

Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, Peter B (2010) Peptide binding predictions for HLA DR, DP and DQ molecules. BMC Bioinform 11(1):568

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35(suppl_2):W407–W410