

RESEARCH

Open Access



# Vaccine design of coronavirus spike (S) glycoprotein in chicken: immunoinformatics and computational approaches

Eman A. Awadelkareem<sup>1</sup> and Sumaia A. Ali<sup>2\*</sup> 

## Abstract

**Background:** Infectious bronchitis (IB) is a highly contagious respiratory disease in chickens and produces economic loss within the poultry industry. This disease is caused by a single stranded RNA virus belonging to Coronaviridae family. This study aimed to design a potential multi-epitopes vaccine against infectious bronchitis virus spike protein (S). Protein characterization was also performed for IBV spike protein.

**Methods:** The present study used various tools in Immune Epitope Database (IEDB) to predict conserved B and T cell epitopes against IBV spike (S) protein that may perform a significant role in provoking the resistance response to IBV infection.

**Results:** In B cell prediction methods, three epitopes ( $_{1139}KKSSYY_{1144}$ ,  $_{1140}KSSYYT_{1145}$ ,  $_{1141}SSYYT_{1145}$ ) were selected as surface, linear and antigenic epitopes.

Many MHC I and MHC II epitopes were predicted for IBV S protein. Among them  $_{982}YYITARDMY_{990}$  and  $_{983}YITARDMYS_{991}$  epitopes displayed high antigenicity, no allergenicity and no toxicity as well as great linkage with MHC I and MHC II alleles. Moreover, docking analysis of MHC I epitopes produced strong binding affinity with BF<sub>2</sub> alleles.

**Conclusion:** Five conserved epitopes were expected from spike glycoprotein of IBV as the best B and T cell epitopes due to high antigenicity, no allergenicity and no toxicity. In addition, MHC epitopes showed great linkage with MHC alleles as well as strong interaction with BF<sub>2</sub> alleles. These epitopes should be designed and incorporated and then tested as multi-epitope vaccine against IBV.

**Keywords:** IBV, Spike protein, B-and T-cell epitopes, Computational approaches, Vaccine design

## Introduction

Infectious bronchitis virus (IBV) is a single Positive stranded RNA that belonging to coronavirus of the chicken (*Gallus gallus*). It's a highly contagious respiratory disease in chickens that's mainly severe for very young chicks. The signs of illness include tracheal rales, coughing, sneezing, nasal discharge and some strains may cause kidney damage [1, 2]. The disease can be transmitted in

respiratory discharges and feces by infected chickens, and it is spread by aerosol, ingestion of contaminated feed and water, and contact with contaminated equipment or clothing. The virus is not transmitted via eggs [3]. The disease causes economic loss within the poultry industry, affecting the performance of meat-type and egg-laying birds. The disease can affect all ages, but the clinical disease is more severe in young chicks. Chicks become more resistant to IBV-induced mortality with the increasing age [4].

There are four structural proteins related to the envelope, the spike (S), membrane (M), envelope (E), and nucleocapsid (N) protein [5]. The spike 'S' glycoprotein which located

\* Correspondence: somiahadool@yahoo.com

<sup>2</sup>Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, Sudan University of Science and Technology, Khartoum, Sudan  
Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Table 1** Accession numbers, date and area of collection of the retrieved sequences of Spike protein of IBV

No	Accession No	Country	Year	No	Accession No	Country	Year
1	NP_040831.1 <sup>a</sup>	UK	2018	47	AAV98206.1	USA	2002
2	AHX25911.1	China	2016	48	AVX27612.1	India	2004
3	AHX25902.1	China	2016	49	ALE71331.1	India	2018
4	AHX25893.1	China	2016	50	AJP16712.1	China	2015
5	AMK51938.1	China	2016	51	AJP16739.1	China	2015
6	AEP84746.1	China	2016	52	APP50306.1	Korea	2015
7	AEP84736.1	China	2016	53	APP50302.1	Korea	2012
8	ACX71849.1	China	2011	54	APP50294.1	Korea	2012
9	ACX71844.1	China	2011	55	APP50274.1	Korea	2012
10	ACX71842.1	China	2011	56	AEL12221.1	China	2012
11	AAU09490.1	China	2011	57	ADY62552.1	China	2012
12	AAV24433.1	Singapore	2005	58	ADV71785.1	Netherlands	2010
13	AAV24423.1	Singapore	2005	59	ACQ55230.1	Netherlands	2011
14	AAV21248.1	Singapore	2005	60	ARE67884.1	Pakistan	2017
15	AAV21247.1	Singapore	2005	61	ARB66180.1	China	2017
16	AAV21246.1	Singapore	2005	62	AQY55821.1	China	2017
17	AAV21245.1	Singapore	2005	63	AHX26172.1	China	2016
18	AAV21244.1	Singapore	2005	64	AHX26163.1	China	2016
19	AAV21243.1	Singapore	2005	65	AHX26154.1	China	2016
20	AAV21242.1	Singapore	2005	66	AHX26145.1	China	2016
21	AGW24533.1	India	2015	67	AHX26136.1	China	2016
22	AAW33786.1	USA	2006	68	AHX26127.1	China	2016
23	AER08740.1	Sweden	2012	69	AHX26118.1	China	2016
24	AER08739.1	Sweden	2012	70	AHX26109.1	China	2016
25	AER08729.1	Sweden	2012	71	AHX26073.1	China	2016
26	AER08728.1	Sweden	2012	72	AHX26064.1	China	2016
27	AER08727.1	Sweden	2012	73	AHX26055.1	China	2016
28	AER08726.1	Sweden	2012	74	AHX26046.1	China	2016
29	AER08725.1	Sweden	2012	75	AHX26037.1	China	2016
30	AER08724.1	Sweden	2012	76	AHX26028.1	China	2016
31	AER08723.1	Sweden	2012	77	AHX26019.1	China	2016
32	AER08722.1	Sweden	2012	78	AHX26010.1	China	2016
33	AER08721.1	Sweden	2012	79	AHX26001.1	China	2016
34	ADA83557.1	USA	2011	80	AHX25992.1	China	2016
35	ADA83467.1	USA	2011	81	AHX25983.1	China	2016
36	ABH01142.1	USA	2007	82	AHX25974.1	China	2016
37	ABH01141.1	USA	2007	83	AHX25965.1	China	2016
38	ABI26423.1	USA	2006	84	AHX25956.1	China	2016
39	AAK27168.1	China	2005	85	AHX25947.1	China	2016
40	ACH72794.1	China	2009	86	AHX25938.1	China	2016
41	AAW83034.1	China	2006	87	AHX25929.1	China	2016
42	ARS23139.1	Egypt	2014	88	ACJ50199.1	Singapore	2005
43	AHX25920.1	China	2016	89	ACO37566.1	Singapore	2005
44	ADP06504.1	USA	2012	90	AYG86360.1	SouthKorea	2018
45	AAA66578.1	UK	1995	91	AYG86347.1	SouthKorea	2018
46	AAA70235.1	USA	2002	92	AAV28722.1	China	2006

<sup>a</sup>Refseq

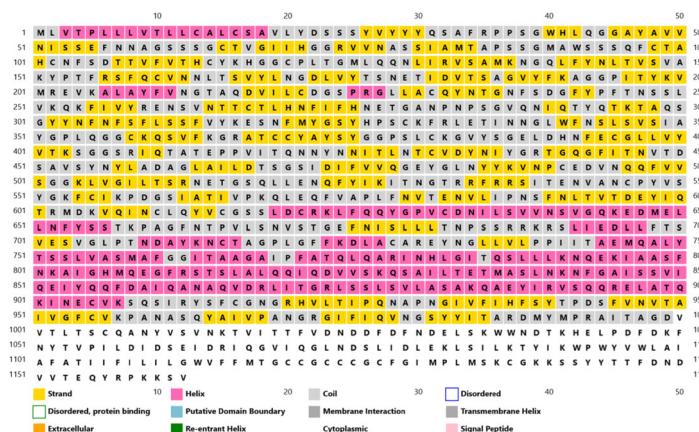
at the surface of the virion. The membrane 'M' glycoprotein is partially exposed at the surface of the virion and also the nucleocapsid 'N' protein that located internally. The spike glycoprotein of IBV induces virus neutralizing (VN) and HI antibodies and has been considered as the most likely inducer of protection [2, 4]. The S protein is either a dimer or trimer. It has two recognized functions; binding the virus to receptor molecules on host cells, and activating fusion of the virion membrane with host cell membranes, releasing the viral genome into the cell [2]. The spike gene in particular the S1 part, is highly variable, due to insertions, deletions, substitutions and recombination events [6]. Application of vaccine is the most effective way to protect against pathogenic diseases, particularly when these pathogens have a high mortality rate such as IBV and viruses in general. On the other hand, the large number of serotypes and strains (genotype) of IBV specifically complicate control method. IBV has shift and drift property [7].

Inactivated and live-attenuated vaccines are employed to control the disease. However, inactivated vaccines often fail to induce strong cellular immunity, while live-attenuated vaccines can contribute to development of antigenic variant viruses [5]. The increasing number of new IBV serotypes, caused by frequent gene mutation and recombination, poses a major challenge for the prevention and control of infectious bronchitis disease [8].

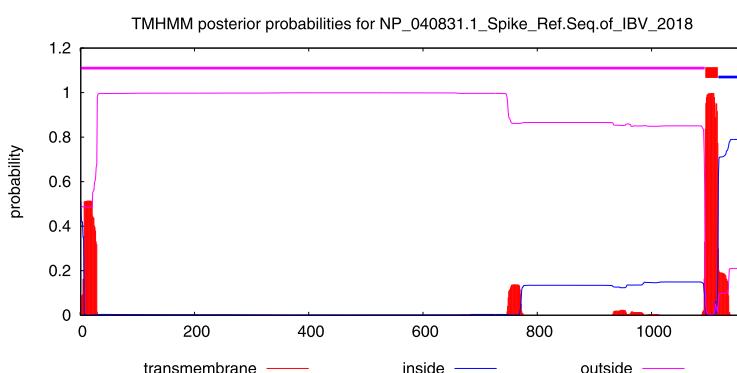
RNA viruses such as IBV have high mutational rates. Thus, the most important step in the design of cross-protective peptide vaccine against IBV is to target the conserved epitopes of different IBV serotypes [5].

Presentation by MHC molecules is important for developing vaccinal immunity. MHC class I and class II molecules are typically highly polymorphic and polygenic [9]. Avian MHC class I and class II genes are localized into two regions (MHC-B and MHC-Y) on the chromosome 16. The MHC-B and MHC-Y haplotypes assort independently as the result of an intervening

(a)



(b)



**Fig. 1 (a):** The secondary structure of IBV spike protein; **(b):** transmembrane topology of spike protein; **(c):** the position of disulphide bond (S-S) in spike protein of IBV

(c)

Predicted bonds	
13 - 358	LVTLLCALCSA - PLQGGCKQSVF
16 - 267	LLCALCSAVLY - SVNTTCTLHNF
65 - 120	GSSSGCTVGII - YKHGGCPLTGM
98 - 610	SSSQPCTAHCN - KVQINCLQYVC
102 - 431	FCTAHCNFSDT - ITLNTCVDYNI
114 - 395	VVFTHCYKHCG - DHNFECGLLVY
160 - 1121	FRSFQCVCNNLT - FMTGCCCCCG
221 - 370	QDVILCDGSPR - GRATCCYAYSY
231 - 632	RGLLACQYNTG - QYGPVCNDILS
329 - 546	SYHPSCKFRL - ENVANCPVPSY
369 - 905	KGRATCCYAYS - QKINECVKSQS
381 - 916	CGPSLCKGVYS - IRYSPFCGNRHF
491 - 1123	YKVNPCEDVNQ - TGCCCCCCCCF
555 - 1137	SYGKPCIKPDG - PLMSKCGKKSS
615 - 728	CLQYVCGSSL - FKDLACAREYN
621 - 1006	GSSLDCRKLPQ - VTLTSCQANYV
715 - 955	DAYKNCTAGPL - AIVGFCVKPAN
1120 - 1125	FFMTGCCCCCCC - CGGCCCCGCFGI
1124 - 1127	GCCCCCGGCFG - GCCCCCGFCIMP

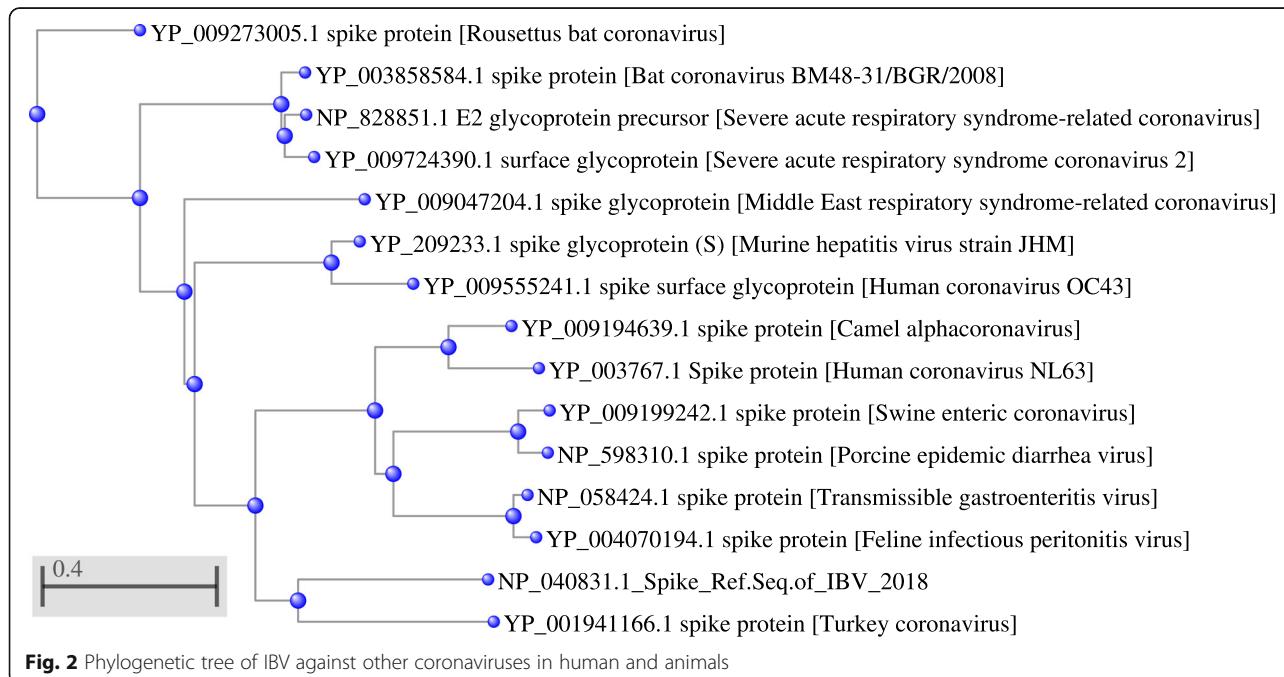
**Table 2** Blastp similarity search of IBV against other coronaviruses in human and animals

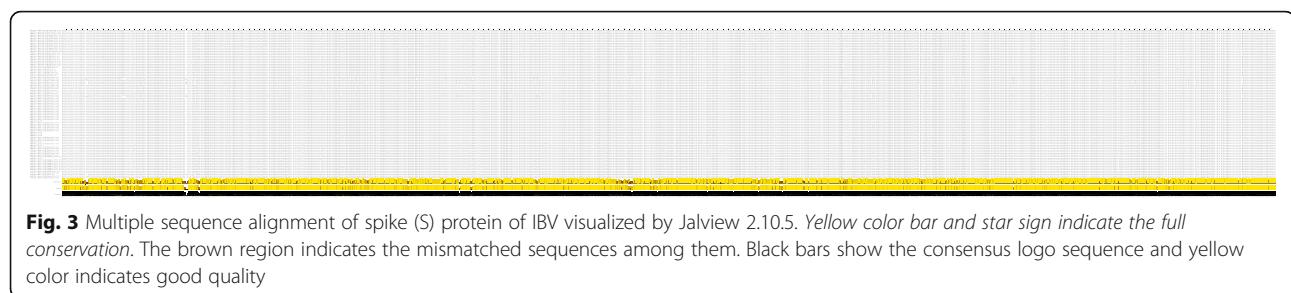
NCBI Protein ID	Protein Name	E- value	Identity
YP_001941166.1	Turkey coronavirus	0.0	38.59%
YP_009194639.1	Camel alphacoronavirus	8e-126	34.05%
YP_009199242.1	Swine enteric coronavirus	4e-124	31.47%
YP_003767.1	Human coronavirus NL63	2e-121	34.01%
NP_598310.1	Porcine epidemic diarrhea virus	8e-120	31.36%
YP_009273005.1	Rousettus bat coronavirus	1e-115	32.86%
NP_058424.1	Transmissible gastroenteritis virus	2e-109	32.03%
YP_209233.1	Murine hepatitis virus strain JHM	9e-109	37.20%
YP_004070194.1	Feline infectious peritonitis virus	1e-108	31.70%
YP_003858584.1	Bat coronavirus BM48-31/BGR/2008	4e-107	35.69%
NP_828851.1	E2 glycoprotein precursor [Severe acute respiratory syndrome-related coronavirus]	5e-107	36.28%
YP_009724390.1	Severe acute respiratory syndrome coronavirus 2	1e-106	36.31%
YP_009555241.1	Human coronavirus OC43	3e-105	31.42%
YP_009047204.1	Middle East respiratory syndrome-related coronavirus	2e-104	34.71%

region that supports highly frequent recombination [9, 10]. Chicken MHC B-F molecules have been structurally and functionally related to mammalian MHC class I molecules and have been involved in the presentation of antigen to CD8 + T lymphocytes, which is important for antiviral immune response [11]. Recently, the design of epitope-based vaccines has been expanded by developments in genomics, proteomics and the understanding of pathogens. Epitope is the negligible immunogenic region of a sequence of proteins that specifically produces

accurate immune responses [12]. The identification of specific B and T cell epitopes produced more desirable manipulation of immune response [13]. It is known that designing of multi-epitope vaccines using bioinformatics tools can significantly reduce the time and cost of production and produce satisfactory results [14, 15].

The production of safer and more reliable vaccines for controlling IBV is important. Therefore, the aim of this study is to analyze strains of spike (S) glycoprotein of infectious bronchitis virus reported in NCBI database





using immunoinformatics and computational approaches to select all possible epitopes that can be used as multi-epitopes vaccine. Protein characterization was also achieved for IBV spike protein.

## Material and method

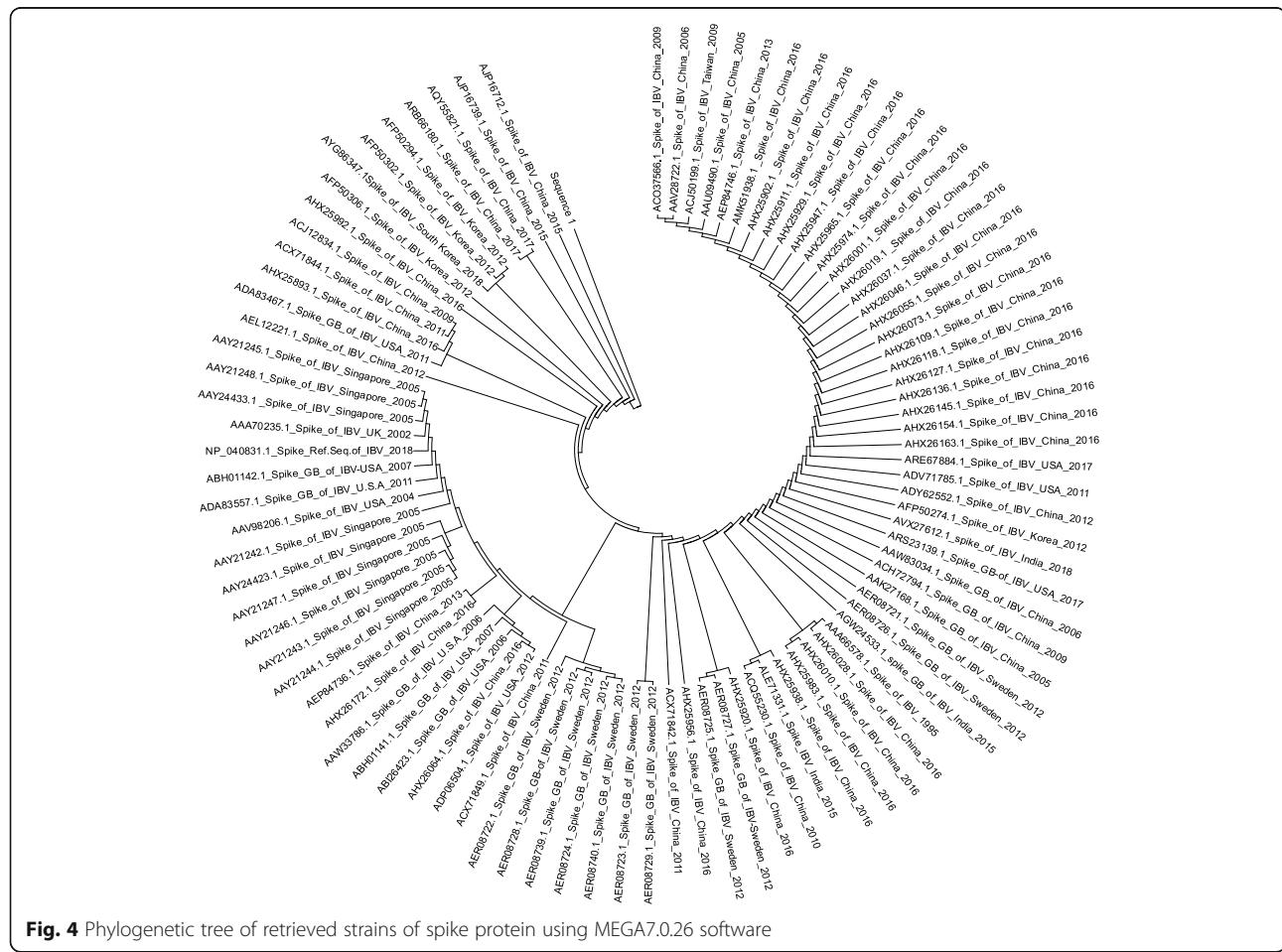
### Protein sequence retrieval

Spike (S) protein sequences of different infectious bronchitis virus (IBV) strains were retrieved from the Gene-Bank of National Central Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/protein/>) database in March 2019. The sequences were saved in FASTA format (Table 1).

### Structural analysis

Reference sequence of spike S protein (NP\_040831.1) was analyzed to identify chemicals and physical properties including GRAVY (grand average of hydropathicity), half-life, molecular weight, stability index and amino acids atomic composition using an online tool ProtParam [16]

Secondary structure of spike S protein of IBV was analyzed through PSIPRED [17]. The secondary structure of protein including helix, sheet, turn, and coil parameters was predicted using GOR IV server at [https://npsa-prabi.ibcp.fr/cgi-bin/secpred\\_sopma.pl](https://npsa-prabi.ibcp.fr/cgi-bin/secpred_sopma.pl). TMHMM an online tool (<http://www.cbs.dtu.dk/services/TMHMM/>),



**Table 3** Conservancy assessment of B cell linear epitopes

Epitope no	Epitope sequence	Start	End	Epitope length	Percent of protein sequence matches at identity <= 100%
1	MTAPSSGMAW	83	92	10	<b>89.13% (82/92)</b>
2	GGPI	193	196	4	<b>90.22% (83/92)</b>
3	TGNFSD	235	240	6	<b>97.83% (90/92)</b>
4	GPLQGGCK	352	359	8	<b>94.57% (87/92)</b>
5	DSAV	450	453	4	<b>91.30% (84/92)</b>
6	VNPCEDV	488	494	7	<b>96.74% (89/92)</b>
7	RNETGSQ	512	518	7	<b>94.57% (87/92)</b>
8	VGQKE	642	646	5	<b>81.52% (75/92)</b>
9	STKPAGFNTP	656	665	10	<b>81.52% (75/92)</b>
10	PQNAPN	926	931	6	<b>98.91% (91/92)</b>
11	ANASQY	959	964	6	<b>98.91% (91/92)</b>
12	IVPA	966	969	4	<b>86.96% (80/92)</b>
13	DFDFN	1026	1030	5	<b>84.78% (78/92)</b>
14	SKWWNDTKHELP	1034	1045	12	<b>94.57% (87/92)</b>
15	GKKSSYYTT	1138	1146	9	<b>97.83% (90/92)</b>

used to examine the trans-membrane topology of S protein. Presence of disulphide-bonds were predicted through an online tool DIANNA v1.1. It makes prediction based on trained neural system [18]. CDD-BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) [19–21] and PFAM (<https://pfam.xfam.org/>) [22] were used to search the defined conserved domains in the targeted protein sequences. Blastp in NCBI database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) using reference sequence (refseq-protein) database was used to compare spike reference sequences of different coronaviruses in human and animals against IBV spike protein sequence. Phylogenetic tree was also constructed based on COBALT multiple alignment (<https://www.ncbi.nlm.nih.gov/blast/treeview/treeView.cgi>) [19, 20].

#### Multiple sequence alignment and epitope conservancy assessment

The retrieved sequences of IBV S protein were aligned using Clustal program and consensus sequence was generated using the multiple sequence alignment (MSA) tool, Jalview version 2.10.5. (<http://www.jalview.org/about/jalview-scientific-advisory-committee>) [23]. Epitope conservancy analysis in Immune Epitope Database (IEDB) was used to detect potential epitope conservancy (<http://tools.iedb.org/conservancy/>) [24]. For calculating the conservancy score, the sequence identity threshold was kept at 80%.

#### Phylogeny analysis

Phylogenetic tree of the retrieved sequences of spike (S) protein was performed using MEGA7.0.26 (7170509) software using maximum likelihood parameter [25].

**Table 4** List of shortened B cell epitopes predicted by different B cell scale

No.	Peptide	Start	End	Length	Emini	koleskar
1	MTAP	83	86	4	0.949	0.966
2	GSRIQT	406	411	6	1.273	0.973
3	SRIQT	407	411	5	1.583	0.992
4	SRIQ	407	410	4	1.355	1.013
5	STKP	656	659	4	2.543	0.979
6	VGLP	704	707	4	0.398	1.143
7	VGLPT	704	708	5	0.465	1.096
8	NASQY	960	964	5	2.034	1.006
9	SKWW	1034	1037	4	1.26	0.932
10	KKSSYYTT	1139	1146	8	6.723	1.003
11	KSSYYTT	1140	1146	7	4.166	1.013
12	SSYYTT <sup>a</sup>	1141	1146	6	2.568	1.027
13	SYYTT <sup>a</sup>	1142	1146	5	2.359	1.03
14	YYTT <sup>a</sup>	1143	1146	4	1.26	1.035
15	KKSSYYT	1139	1145	7	5.773	1.016
16	KKSSYY <sup>a</sup>	1139	1144	6	4.931	1.034
17	KSSYYT <sup>a</sup>	1140	1145	6	3.559	1.031
18	KKSSY	1139	1143	5	3.875	1.009
19	KKSS	1139	1142	4	3.054	0.971
20	SSYYT <sup>a</sup>	1141	1145	5	2.191	1.051
21	SYYT <sup>a</sup>	1142	1145	4	2.019	1.061

<sup>a</sup>Shortened peptide that has high score in both Emini and kolaskar

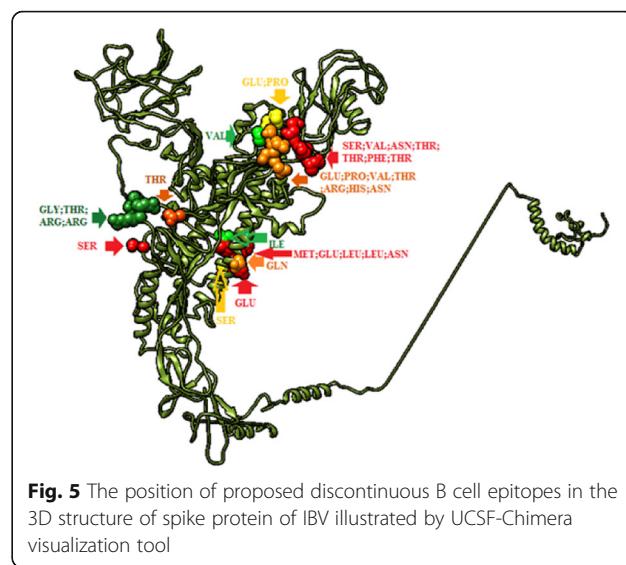
**Table 5** Discontinuous epitopes predicted through DISCOTOPe 2.0 Server

Residue ID	Residue Name	Contact Number	Propensity Score	Discotope Score
262	SER	2	-3.91	-3.69
263	VAL	4	-2.626	-2.784
264	ASN	0	-0.238	-0.211
265	THR	19	-1.418	-3.44
266	THR	5	0.483	-0.148
267	PHE	25	-0.627	-3.43
268	THR	7	-0.463	-1.215
387	GLY	1	-3.678	-3.37
414	GLU	7	-0.476	-1.226
415	PRO	8	0.187	-0.754
417	VAL	5	-0.324	-0.862
419	THR	6	1.351	0.506
420	ARG	0	1.529	1.353
421	HIS	11	0.482	-0.838
422	ASN	12	-2.504	-3.596
515	THR	4	-3.071	-3.178
531	GLY	5	-1.472	-1.877
532	THR	5	1.933	1.136
533	ARG	0	1.404	1.243
534	ARG	0	-0.425	-0.376
648	MET	5	-1.103	-1.551
649	GLU	16	-1.752	-3.39
650	LEU	26	-4.013	-6.541
651	LEU	10	-2.379	-3.256
652	ASN	12	-1.72	-2.902
655	SER	7	-2.994	-3.454
685	SER	0	-3.842	-3.4
741	ILE	15	-1.303	-2.878
893	GLN	7	-2.411	-2.939
896	GLU	9	-2.901	-3.602

### B cell prediction

The Immune Epitope Database (IEDB) (<http://tools.iedb.org/mhci/>) was used to predict B and T cell epitopes of IBV reference sequence of S protein (NP\_040831.1) [26]. Linear B-cell epitopes were predicted using BepiPred from IEDB [27]. Emini surface accessibility prediction tool was used to predict surface located epitopes [28]. Whereas, the antigenic epitopes were investigated using kolaskar and Tongaonkar antigenicity method [29].

Discontinuous epitopes were predicted using Discotope server [30]. The parameter was set at  $\geq 0.5$  which indicated 90% specificity and 23% sensitivity. This method is based on surface accessibility and amino acid



**Fig. 5** The position of proposed discontinuous B cell epitopes in the 3D structure of spike protein of IBV illustrated by UCSF-Chimera visualization tool

statistics in a collected form dataset of discontinuous epitopes found out by X-ray crystallography of antigen/antibody protein buildings. Chimera software was used to display the position of predicted epitopes clusters on 3D structure of S protein [31].

### T-cell epitope prediction

The T cell epitopes were predicted in human among different alleles of major histocompatibility complex class I (MHC I) and class II (MHC II).

MHC-I binding epitopes were predicted by the IEDB MHC I prediction tool at <http://tools.iedb.org/mhci>. The binding affinity of peptides to MHC I molecules was

**Table 6** Antigenic, non-allergic and non-toxic MHC I epitopes

Peptide	Start	End	Antigenicity	Allele	IC50
YYITARDMY	982	990	<b>0.8845</b>	HLA-A*29:02	14.52
				HLA-A*30:02	160.94
				HLA-C*14:02	27.32
YITARDMYM	983	991	<b>0.7901</b>	HLA-A*02:01	233.08
				HLA-A*02:06	212.86
				HLA-C*03:03	29
				HLA-C*06:02	200.39
				HLA-C*07:01	267.22
				HLA-C*14:02	49.52
				HLA-C*15:02	77.63
TARDMYMPR	985	993	<b>0.6914</b>	HLA-A*30:01	56.23
				HLA-A*31:01	14.3
				HLA-A*68:01	28.24
IIFILILGW	1105	1113	<b>0.6749</b>	HLA-B*57:01	78.45
				HLA-B*58:01	64.27
KKSSYYTTF	1139	1147	<b>1.1865</b>	HLA-A*32:01	182.52

**Table 7** Antigenic, non-allergic and non-toxic MHCII epitopes

Core Sequence	Antigenicity	Peptide Sequence	Start	End	Allele	IC50
IIFILILGW	<b>0.6914</b>	IAFATIIFILILGWV	1100	1114	HLA-DRB1*15:01	454.6
KKSSYYTF	<b>0.6749</b>	MSKCGKKSSYYTFD	1134	1148	HLA-DPA1*01:03/DPB1*02:01	872.7
		SKCGKKSSYYTFD	1135	1149	HLA-DPA1*01/DPB1*04:01	408.1
					HLA-DPA1*01:03/DPB1*02:01	301.5
					HLA-DPA1*02:01/DPB1*05:01	953.4
		KCGKKSSYYTFDND	1136	1150	HLA-DPA1*01/DPB1*04:01	276.8
					HLA-DPA1*02:01/DPB1*05:01	853.9
		CGKKSSYYTFDNDV	1137	1151	HLA-DPA1*02:01/DPB1*05:01	958.9
		MSKCGKKSSYYTFD	1134	1148	HLA-DPA1*01:03/DPB1*02:01	872.7
KSSYYTFD	<b>0.6466</b>	MSKCGKKSSYYTFD	1134	1148	HLA-DRB1*04:05	155
		SKCGKKSSYYTFDN	1135	1149	HLA-DRB1*04:05	125.6
		KCGKKSSYYTFDND	1136	1150	HLA-DRB1*04:05	92.2
		CGKKSSYYTFDNDV	1137	1151	HLA-DRB1*04:05	51.9
		GKKSSYYTFDNDVW	1138	1152	HLA-DRB1*04:05	46.9
		KKSSYYTFDNDWVT	1139	1153	HLA-DRB1*04:05	45.3
TARDMYMPR	<b>0.7901</b>	SYYITARDMYMPRAI	981	995	HLA-DRB1*03:01	269.3
		YYITARDMYMPRAIT	982	996	HLA-DRB1*03:01	281.9
		YITARDMYMPRAITA	983	997	HLA-DRB1*03:01	618.8
YITARDMYM	<b>1.1865</b>	QVNGSYYITARDMYM	977	991	HLA-DRB1*01:01	22
					HLA-DRB1*04:01	145
					HLA-DRB1*04:04	331.2
					HLA-DRB1*07:01	20.3
					HLA-DRB3*01:01	550.7
					HLA-DRB5*01:01	227.8
		VNGSYYITARDMYMP	978	992	HLA-DQA1*01:02/DQB1*06:02	338.6
					HLA-DRB1*01:01	25.8
					HLA-DRB1*03:01	447.6
					HLA-DRB1*04:01	105.8
					HLA-DRB1*04:04	248.3
					HLA-DRB1*07:01	27.8
					HLA-DRB1*15:01	380.6
					HLA-DRB3*01:01	577.8
					HLA-DRB5*01:01	198.6
		NGSYYITARDMYMPR	979	993	HLA-DQA1*01:02/DQB1*06:02	393.3
					HLA-DQA1*05:01/DQB1*03:01	817.3
					HLA-DRB1*01:01	19.8
					HLA-DRB1*03:01	176.5
					HLA-DRB1*04:01	65.2
					HLA-DRB1*04:04	225
					HLA-DRB1*07:01	40.2
					HLA-DRB1*15:01	291.2
					HLA-DRB3*01:01	635
					HLA-DRB5*01:01	93.5
		GSYYITARDMYMPRA	980	994	HLA-DQA1*01:02/DQB1*06:02	218

**Table 7** Antigenic, non-allergic and non-toxic MHCII epitopes (*Continued*)

Core Sequence	Antigenicity	Peptide Sequence	Start	End	Allele	IC50
					HLA-DRB1*01:01	14
					HLA-DRB1*03:01	197.3
					HLA-DRB1*04:01	47.8
					HLA-DRB1*04:04	242.4
					HLA-DRB1*07:01	57.3
					HLA-DRB1*15:01	288.6
					HLA-DRB3*01:01	780.4
					HLA-DRB5*01:01	61.4
		SYYITARDMYMPRAI	981	995	HLA-DRB1*01:01	23.1
					HLA-DRB1*04:01	65.3
					HLA-DRB1*04:04	249.2
					HLA-DRB1*04:05	356.4
					HLA-DRB1*07:01	72.2
					HLA-DRB1*15:01	284.7
					HLA-DRB5*01:01	87.4
		YYITARDMYMPRAIT	982	996	HLA-DRB1*01:01	40.8
					HLA-DRB1*04:01	108.8
					HLA-DRB1*04:04	269.1
					HLA-DRB1*04:05	706.3
					HLA-DRB1*07:01	160.6
					HLA-DRB5*01:01	121.3
		YITARDMYMPRAITA	983	997	HLA-DRB1*04:01	145.4
					HLA-DRB1*04:04	652.4
					HLA-DRB1*07:01	355.4
					HLA-DRB1*08:02	955
					HLA-DRB5*01:01	206.9
YYITARDMY	<b>0.8845</b>	IQVNGSYYITARDMY	976	990	HLA-DQA1*05:01/DQB1*02:01	491.6
					HLA-DRB1*04:01	723.4
					HLA-DRB1*04:04	819.7
					HLA-DRB1*11:01	72
					HLA-DRB1*11:01	72
		QVNGSYYITARDMYM	977	991	HLA-DPA1*01/DPB1*04:01	710.8
					HLA-DPA1*01:03/DPB1*02:01	875.8
					HLA-DQA1*05:01/DQB1*02:01	292.7
					HLA-DRB1*03:01	588
					HLA-DRB1*11:01	32.4
					HLA-DRB1*11:01	32.4
					HLA-DPA1*01/DPB1*04:01	557.6
					HLA-DPA1*01:03/DPB1*02:01	860.8
					HLA-DQA1*05:01/DQB1*02:01	311.9
					HLA-DRB1*11:01	17.9
					HLA-DRB1*11:01	17.9
		NGSYYYITARDMYMPR	979	993	HLA-DPA1*01/DPB1*04:01	503
					HLA-DPA1*01:03/DPB1*02:01	763.2

**Table 7** Antigenic, non-allergic and non-toxic MHCII epitopes (Continued)

Core Sequence	Antigenicity	Peptide Sequence	Start	End	Allele	IC50
					HLA-DQA1*05:01/DQB1*02:01	387.6
					HLA-DRB1*09:01	858.7
					HLA-DRB1*09:01	858.7
					HLA-DRB1*11:01	11
					HLA-DRB1*11:01	11
GSYYITARDMYMPRA	980	994			HLA-DPA1*01/DPB1*04:01	504.5
					HLA-DPA1*01:03/DPB1*02:01	790.4
					HLA-DQA1*05:01/DQB1*02:01	482.9
					HLA-DRB1*11:01	15.2
					HLA-DRB1*11:01	15.2
SYYITARDMYMPRAI	981	995			HLA-DPA1*01/DPB1*04:01	480
					HLA-DPA1*01:03/DPB1*02:01	733.8
					HLA-DQA1*05:01/DQB1*02:01	526.3
					HLA-DRB1*11:01	26.8
					HLA-DRB1*11:01	26.8
YYITARDMYMPRAIT	982	996			HLA-DPA1*01/DPB1*04:01	705.6
					HLA-DPA1*01:03/DPB1*02:01	931.5
					HLA-DQA1*05:01/DQB1*02:01	678.7
					HLA-DRB1*11:01	51.8
					HLA-DRB1*11:01	51.8

measured using artificial neural networks (ANN) method [32, 33]. Prior to prediction, peptide lengths were set as 9 mers. The half maximum inhibitory concentration (IC50) values needed for the binding of peptide to MHC-I molecules were set less than or equal to 300 nM.

The IEDB MHCII prediction tool was used for MHC class II molecules at (<http://tools.iedb.org/mhcii/>) [26]. Human MHC class II alleles (HLA DR, HLADP and HLADQ) were used for MHCII binding predication. The NN-align method was used with IC50 less or equal to 1000 nM [34].

#### Antigenicity, allergenicity and toxicity of epitopes

VaxiJen v2.0 server was used to predict the antigenicity of the conserved regions (<http://www.jenner.ac.uk/VaxiJen>) [35]. The default prediction parameters and a threshold value of 0.4 were used. The in silico allergenicity prediction of epitopes was investigated using AllerTop v .2.0 (<http://www.ddg-pharmfac.net/AllerTOP>) [36]. While ToxiPred server was used to evaluate the toxicity of predicted epitopes. (<http://crdd.osdd.net/raghava/toxinpred/>) [37].

#### Homology modeling

IBV reference sequence and the protein sequences of BF alleles (BF2 \*2101 and BF2\*0401) were submitted to

Raptor X server (<http://raptorgx.uchicago.edu/>) to design their three D structures [38–40]. PEPFOLD3 server was used for the homology modelling of MHCII epitopes (<http://bioserv.rpbs.univ-paris-diderot.fr/services/PEPFOLD3/>) from amino acid sequences [41–43]. Chimera software 1.8 was used to visualize 3D structures of IBV spike S protein reference sequence and BF alleles [31].

#### Molecular docking

To perform molecular docking, 3D structures of MHCII epitopes and BF alleles were submitted simultaneously to the PatchDock online autodock tools; an automatic server for molecular docking (<https://bioinfo3d.cs.tau.ac.il/PatchDock/>) [44]. The five top models were selected using firedock [44]. The results were visualized using the UCSF-Chimera software 1.8 [31].

## Results

#### Structural analysis

The physiochemical properties of the spike S protein, measured through ProtParam, showed that it contained 1162 amino acids (aa) with a molecular weight of 128,046.70 kDa. The spike protein showed an antigenic nature when subjected to VaxiJen v2.0.

Theoretical isoelectric point (PI) of spike protein was 7.71, indicating its positive in nature. An isoelectric point above 7 indicates the protein is charged positively.

Near to 81 aa charges were found negative, whereas 84 aa found positive.

Protparam computed instability-index (II) 35.53, this categorize the protein as stable. Aliphatic-index 86.05, which devotes a thought to the proportional volume holding by aliphatic side chain and GRAVY value of the protein sequence is 0.012. Half-life of S protein shown as the total time taken for its vanishing after it has been synthesized in cell, computed as 30 h for mammalian-reticulocytes, > 20 h for yeast, > 10 h for *Escherichia coli*. The total numbers of Carbon (C), Oxygen (O), Nitrogen (N), Hydrogen (H) and Sulfur (S) were entitled by the formula: C<sub>5737</sub>H<sub>8847</sub>N<sub>1495</sub>O<sub>1718</sub>S<sub>56</sub>.

The secondary structure of IBV spike S protein was analyzed through PSIPRED and GOR IV server. The components of secondary structure prediction by GOR IV server are alpha helix (29.43%), extended strand (27.37%), beta turn (5.25%), and random coil (37.95%) (Fig. 1).

DiANNA1.1 tool calculated 19 disulphides bond (S–S) positions and assign them a score and makes prediction based on trained neural system. The trans-membrane protein topology was investigated via online tool TMHMM. Residues from 1 to 1093 were found to be exposed to the surface, residue from 1094 to 1116 were found inside trans-membrane-region and residues from 1117 to 1162 were buried within the core-region of the S protein (Fig. 1).

In refseq of IBV spike protein two conserved domains (Corona-S2, Corona-S2) were identified. The conserved domains were sequenced by Conserved Domain (CDD) BLAST search. The results revealed that corona-S1 (pfam01600) is the only member of the superfamily cl03276 and corona-S2 domain (pfam01601) is the only member of the superfamily cl20218. The top associated sequences in both domains were Feline infectious peritonitis virus (strain 79–1146), Avian infectious bronchitis virus (strain Beaudette), and Human coronavirus 229E while Severe acute respiratory syndrome-related coronavirus sequences were associated only with corona-S2 domain. The closest homologue obtained from BLASTP (refseq-protein) results was the Turkey coronavirus S protein with E value 0.00 followed by Murine hepatitis virus strain JHM with E value 9e-109 when comparing various coronaviruses in human and animals with IBV spike protein sequence (Table 2). Phylogenetic tree of IBV against other coronaviruses in human and animals was created based on COBALT multiple alignment see Fig. 2.

#### Multiple sequence alignment

Jalview was used to visualize the multiple sequence alignment of the retrieved sequences. Several areas in alignment were shown to have mutation see Fig. 3.

#### Phylogeny

Phylogenetic tree for IBV spike S protein sequences was constructed using MEGA7.0.26 (7170509) software using maximum likelihood parameter see Fig. 4.

#### B-cell epitopes

Several epitopes were predicted in B cell prediction methods using the Bepipred Linear Epitope Prediction tool. The conservancy percentages of these epitopes are presented in Table 3. After shortening of predicted epitopes, 21 linear conserved epitopes were recognized. Of these, seven epitopes with different lengths were identified as linear, surface and antigenic epitopes between the positions 1139–1146 (see Table 4). These epitopes were 1139KKSSYY<sub>1144</sub>, 1140KSSYYT<sub>1145</sub>, 1141SSYYTT<sub>1146</sub>, 1141SSYYT<sub>1145</sub>, 1142SYYTT<sub>1146</sub>, 1142SYYT<sub>1145</sub>, and 1143YYTT<sub>1146</sub>. Based on the length and antigenicity score, three epitopes (1139KKSSYY<sub>1144</sub>, 1140KSSYYT<sub>1145</sub>, 1141SSYYT<sub>1145</sub>) were selected as top B cell epitopes.

Discotope 2.0 server was used to predict the discontinuous epitopes from the 3D structure of S protein (PDB ID: 6CV0), 90% specificity, – 3.700 threshold and 22.000 Angstroms propensity score radius [45]. Total 30 discontinuous epitopes were recognized at different exposed surface areas (Table 5). The position of each predicted epitope on the surface of 3D structure of S protein is shown in Fig. 5 using Chimera visualization tool [31].

#### Prediction of MHC class I epitopes

In this study, the Human MHC class-I HLA alleles were used to explore the interaction of epitopes with MHCI alleles as chicken MHC alleles don't exists in IEDB database. MHC-1 binding prediction tool using IEDB database expected 13 conserved epitopes of spike protein (S) which were interacted with many cytotoxic T cell alleles. These epitopes were 1115FFMTGCCGC<sub>1123</sub>, 590FNLTVT DEY<sub>598</sub>, 734GLLVLPPII<sub>742</sub>, 1105IIFILILGW<sub>1113</sub>, 1139KKSS YYTT<sub>1147</sub>, 1087KTYIKWPWY<sub>1095</sub>, 166SVYLNNGDLV<sub>174</sub>, 985TARDMYMPR<sub>993</sub>, 1145TTFDNDVV<sub>1153</sub>, 983YITARD MYM<sub>991</sub>, 1144YTTFDNDVV<sub>1152</sub>, 982YYITARDMY<sub>990</sub>, 1143YYTTFDNDV<sub>1151</sub>.

#### Prediction of MHC class II epitopes

MHC-II binding prediction tool based on NN-align with half-maximal inhibitory concentration (IC50) ≤ 1000 was used. Thirty one conserved core sequences were predicted to interact with MHCII alleles. These cores were 694EDLLFTSVE<sub>702</sub>, 1147FDNDVVTEQ<sub>1155</sub>, 1115FFMTGC CGC<sub>1123</sub>, 1116FMTGCCGCC<sub>1124</sub>, 590FNLTVTDEY<sub>598</sub>, 734GLLVLPPII<sub>742</sub>, 1105IIFILILGW<sub>1113</sub>, 902INECVKSQS<sub>910</sub>, 984ITARDMYMP<sub>992</sub>, 901KINECVKSQ<sub>909</sub>, 1139KKSSYYTT F<sub>1147</sub>, 1140KSSYYTTFD<sub>1148</sub>, 1187KTYIKWPWY<sub>1195</sub>, 735LLVLPPII<sub>743</sub>, 592LTVTDEYIQ<sub>600</sub>, 1014NKTVIT

*TFV*<sub>1022</sub>    893*QQRELATQK*<sub>901</sub>    894*QRELATQKI*<sub>902</sub>  
 895*RELATQKIN*<sub>903</sub>    589*SFNLTVTDE*<sub>597</sub>    1141*SYYTTFD*  
*N*<sub>1149</sub>    166*SVYLNGLDV*<sub>174</sub>    1142*SYYTTFDND*<sub>1150</sub>  
 985*TARDMYMPR*<sub>993</sub>    1146*TFDNDVVTE*<sub>1154</sub>    593*TVTDEY*  
*IQT*<sub>601</sub>    1013*VNKTVITTF*<sub>1021</sub>    983*YITARDMYM*<sub>991</sub>,  
 1144*YTTFDNDVV*<sub>1152</sub>    982*YYITARDMY*<sub>990</sub> and 1143*YYTT*  
*FDNDV*<sub>1151</sub>.

#### Antigenicity, allergenicity and toxicity of MHC I and MHC II epitopes

The predicted epitopes of MHC I and MHC II were subjected to VaxiJen v2.0 server, AllerJen v2.0, and ToxiPred to estimate the potential antigenicity, allergenicity and toxicity of epitopes. Five MHC I epitopes were identified as antigenic, non-allergic and non-toxic, but only three epitopes (<sub>985</sub>*TARDMYMPR*<sub>993</sub>, <sub>983</sub>*YITARDMYM*<sub>991</sub> and <sub>982</sub>*YYITARDMY*<sub>990</sub>) showed a high linkage with MHC I alleles (Table 6). Furthermore, six MHC II epitopes were predicted to be antigenic, non-allergic and non-toxic epitopes (Table 7). However, <sub>983</sub>*YITARDMYM*<sub>991</sub> and <sub>982</sub>*YYITARDMY*<sub>990</sub> epitopes which were also presented in MHC II prediction methods, showed high antigenicity, no allergenicity and no toxicity. These epitopes were interacted with 52 and 38 alleles in MHC II see Fig. 6.

#### Molecular docking

The molecular docking was achieved by docking MHC I epitopes with chicken BF alleles (BF2 \* 2101 & BF2 \* 0401) using peptide-binding groove affinity. The chicken alleles were used as receptors, and the top MHC I epitopes <sub>982</sub>*YYITARDMY*<sub>990</sub>, <sub>983</sub>*YITARDMYM*<sub>991</sub> and <sub>985</sub>*TARDMYMPR*<sub>993</sub> were used as ligands. Docking of

<sub>983</sub>*YITARDMYM*<sub>991</sub> epitope with BF2\*2101 and BF2\*0401 alleles showed – 72.11 and – 37.39 global energy respectively, indicating a strong binding affinity between the ligands and the receptors compared to other epitopes (Fig. 7, 8 and 9). In general, the global binding affinity of ligands with the receptor BF2\*2101 alleles was found to be lower compared to BF2\*0401, suggesting strong receptor-ligand interaction.

#### Discussion

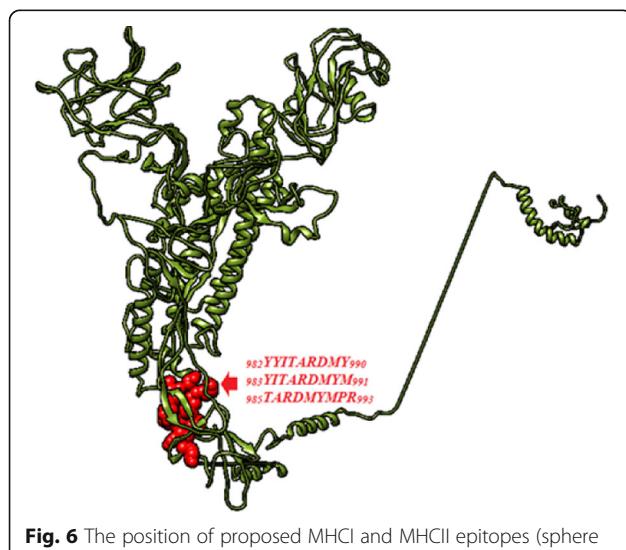
Epitopes capable of inducing immunity in both types (B-cell and T-cell) are considered to be strong candidates for the vaccine [46]. There are several potential benefits offered by peptide vaccine over traditional vaccines against organisms. Most importantly, it allows the immune response to focus only on relevant epitopes and avoid those leading to non-protective responses, immune evasion, or unwanted side effects, such as autoimmunity [47].

IBV vaccination studies have always focused on humoral immune responses regarding protection. Acquired immunity results in the activation of antigen-specific effector mechanisms including B-cells (humoral), T-cells (cellular) and macrophages, and memory cells production [4]. Chickens develop a good humoral response to IBV infections, which measured by ELISA, virus neutralizing (VN) and haemagglutination-inhibition HI antibodies tests [48].

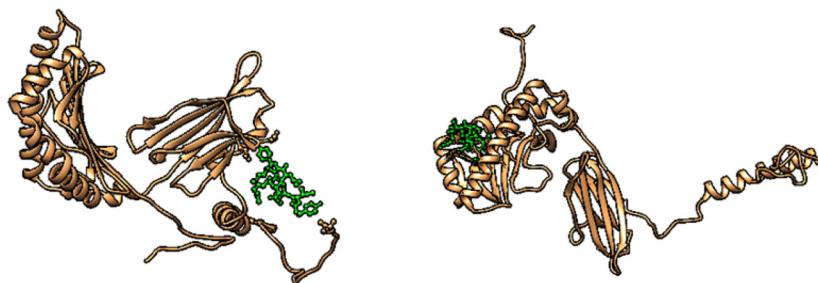
IBV glycoprotein S1 is known to be responsible for virus neutralization (VN) and haemagglutination-inhibition HI antibodies and has been considered the most likely protective inducer [4]. Multi-peptide vaccines using immunoinformatics tools have recently been conducted in Sudan for several viral diseases in chicken such as ILTV, fowlpox, Newcastle and marek's disease virus [15, 49–51].

In the present study, IBV spike protein was analyzed using various prediction servers. Protein characterization of IBV spike S protein using ProtParam confirmed its positive in nature and stable. The protein also exhibited good antigenic properties using VaxiJen 2.0v server.

Corona-S1 and Corona S2 have been identified as major conserved domains in the IBV spike glycoprotein refseq. Conserved Domain (CDD) BLAST search revealed that corona-S1 (pfam01600) is the only member of the superfamily cl03276 and corona-S2 domain (pfam01601) is the only member of the superfamily cl20218. The main related sequences in both domains were Feline infectious peritonitis virus (strain 79–1146), Avian infectious bronchitis virus (strain Beaudette), and Human coronavirus 229E. However, Severe acute respiratory syndrome-related coronavirus sequences was only associated with corona-S2 domain [52].



**Fig. 6** The position of proposed MHC I and MHC II epitopes (sphere red colour) in the 3D structure of spike protein of IBV illustrated by UCSF-Chimera visualization tool



Peptide	Receptor	Energy	Attractive vdw
<b>YITARDMYM</b>	BF <sub>2</sub> 2101	- 72.11	- 37.79
	BF <sub>2</sub> 0401	- 30.33	- 38.52

**Fig. 7** Docking of YITARDMYM with BF<sub>2</sub> alleles

Prediction of B-cell epitopes is essential for the design vaccine components and immuno-diagnostic reagents. B-cell antigenic epitopes are either continuous or discontinuous in nature.

Most epitope prediction methods are based on continuous epitopes [53]. It has been reported that linear B cell epitopes play a role in virus neutralization [11]. IEDB prediction tool was used to predict linear, surface and antigenic epitopes based on the properties of amino acids such as hydrophilicity, surface accessibility, flexibility, and antigenicity [15].

In this study, seven shortened conserved epitopes ( $_{1139}KKSSYY_{1144}$ ,  $_{1140}KSSYYT_{1145}$ ,  $_{1141}SSYYTT_{1146}$ ,  $_{1141}SSYYT_{1145}$ ,  $_{1142}SYYTT_{1146}$ ,  $_{1142}SYYT_{1145}$ , and  $_{1143}YYTT_{1146}$ ) were predicted from B cell prediction methods as surface, linear and antigenic epitopes. The

predicted epitopes were adjacent to each other from the position 1139–1146. In a similar study, only one epitope (YTSNETTDVTS<sup>175–185</sup>) was predicted within the S1 glycoprotein of M41 IBV strains using BepiPred epitope prediction server version 1, and three such epitopes (VSNASPNNSGGVD<sup>279–290</sup>, HPKCNFRPENI<sup>328–338</sup>, NETNNAGSVSDCTAGT<sup>54–69</sup>) were predicted in CR88 IBV strains [11].

The majority of B-cell epitopes are conformational (around a 90%) and only a minority of native antigens have linear B-cell epitopes [54]. Discotope server has been used for predicting discontinuous.

Epitopes from the 3D structure of the spike IBV reference sequence. Around 30 discontinuous epitopes with a specificity of 90% were recognized at different exposed surface areas. These epitopes have a significant



Peptide	Receptor	Energy	Attractive vdw
<b>YYITARDMY</b>	BF <sub>2</sub> 2101	- 64.68	- 35.26
	BF <sub>2</sub> 0401	- 37.57	- 23.93

**Fig. 8** Docking of YYITARDMY with BF<sub>2</sub> alleles



Peptide	Receptor	Energy	Attractive vdw
TARDMYMPR	BF <sub>2</sub> 2101	- 45.61	-30.25
	BF <sub>2</sub> 0401	- 35.47	-28.81

**Fig. 9** Docking of TARDMYMPR with BF<sub>2</sub> alleles

advantage in identifying the native well-structured protein Ag [55].

Cytotoxic T lymphocytes (CTL) provide a critical arm of the immune system in eliminating autologous cells expressing foreign antigen. Unlike humoral immunity, the specificity of CTL activation depends on membrane receptors rather than secreted molecules, and antigen receptors of CTL interact with peptide determinants only in association with matched major histocompatibility complex (MHC) molecules. Virus-specific CTL have been shown to be important, if not critical, for resolution of infection and elimination of viral shedding [1].

It is stated that, the major histocompatibility complex MHC restricted CTL response can be associated with decreases in viral load, and CD8<sup>+</sup> lymphocytes were mostly responsible for the observed protection [1, 56]. Responses to infectious bronchitis virus (IBV) with cytotoxic T-lymphocyte (CTL) were calculated at regular intervals between 3 and 30 days post infection [1].

However, MHCI prediction methods showed three conserved CTL epitopes <sub>985</sub>TARDMYMPR<sub>993</sub>, <sub>983</sub>YITA RD<sub>991</sub> and <sub>982</sub>YYITARDMY<sub>990</sub> as they linked to 7 and 3 human MHCI alleles respectively and showed high antigenicity, no allergenicity and no toxicity. Recent studies showed that vigorous cytotoxic T lymphocyte (CTL) responses that correlate with initial decrease in infection and illness can be detected after IBV infection. It has been identified that the CD8<sup>+</sup> T cells were exhausted without CD4<sup>+</sup> helper T cells. CD4<sup>+</sup> T cells do not seem important in the initial resolution of IBV infection in chickens [56].

In MHCII prediction method, several core peptides were predicted to interact with MHCII alleles, but surprisingly the top core peptides were also <sub>983</sub>YITARDMY<sub>991</sub> and <sub>982</sub>YYITARDMY<sub>990</sub> which were presented in

MHCI prediction methods. They linked with 52 and 38 human alleles respectively. These epitopes showed high antigenicity, no allergenicity and no toxicity.

Molecular docking was performed to display the interaction between BF alleles (BF2\*2101 & BF2\*0401) and MHCI epitopes (<sub>982</sub>YYITARDMY<sub>990</sub>, <sub>983</sub>YITARDMY<sub>991</sub> and <sub>985</sub>TARDMYMPR<sub>993</sub>). The 3D structures of MHCI epitopes were designed using PEPFOLD and docked with BF alleles via Patchdock server. Docking the epitope <sub>983</sub>YITARDMY<sub>991</sub> with both BF2 alleles produced strong binding affinity (- 72.11 and - 37.97 global energy respectively) followed by <sub>982</sub>YYITARDMY<sub>990</sub> (- 64.68 and - 37.57 global energy respectively). This indicates the strong interaction between the ligand and the receptor compared to other epitopes (see Figs. 5, 6 and 7).

Ligands' interaction with the receptor BF2\*2101 alleles was found to be better compared with BF2 \* 0401. However for both BF alleles, the docked molecules showed different groove binding site. Future studies should test the predicted epitopes for therapeutic potency to prove their safety and effectiveness.

## Conclusion

In this study, five epitopes were predicted from spike glycoprotein of IBV as the best B cell (<sub>1139</sub>KKSSYY<sub>1144</sub>, <sub>1140</sub>KSSYYT<sub>1145</sub> and <sub>1141</sub>SSYYT<sub>1145</sub>) and T cell epitopes (<sub>982</sub>YYITARDMY<sub>990</sub> and <sub>983</sub>YITARDMY<sub>991</sub>). They showed high antigenicity, no allergenicity and no toxicity as well as great linkage of MHC epitopes with their alleles. The suggested epitopes should be designed, incorporated and tested as multi-epitopes vaccine against IBV. This vaccine may serve as a possible peptide vaccine to control IBV infection in chicken by inducing humoral and cellular responses.

Peptide vaccination against IBV spike protein (S) can strongly replace traditional vaccines as it is designed to cover all strains in different serotypes, which can reduce recurring outbreaks and their associated massive economic losses.

#### Abbreviations

IB: Infectious bronchitis; IBV: Infectious Bronchitis Virus; IEDB: Immune Epitope Database; S: Spike; MHC: Major histocompatibility complex; BF: The genetic polymorphism of propeptidin factor B; refseq: Reference sequence; NCBI: National Central Biotechnology Information; MSA: Multiple sequence alignment; GRAYV: Grand average of hydropathicity; CDD: Conserved Domain Database; IC50: The half maximal inhibitory concentration; ANN: Artificial neural networks; NN-align: Artificial neural network-based alignment; HLA: The human leukocyte antigen; CTL: Cytotoxic T lymphocytes

#### Acknowledgements

Not applicable.

#### Authors' contributions

Eman, A. Awadelkareem and Sumaia A. Ali designed this study, accomplished the experiments and analyze the results. Sumaia A. Ali interpreted the data and wrote the manuscript. The authors read and approved the final manuscript.

#### Funding

Not applicable.

#### Availability of data and materials

All the data supporting the findings are contained within the manuscript.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.  
All manuscripts must contain the following sections under the heading 'Declarations':

#### Author details

<sup>1</sup>Faculty of Veterinary Medicine, University of Khartoum, Khartoum, Sudan.  
<sup>2</sup>Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, Sudan University of Science and Technology, Khartoum, Sudan.

Received: 29 April 2020 Accepted: 6 July 2020

Published online: 27 August 2020

#### References

- Seo SH, Collisson EW. Specific cytotoxic T lymphocytes are involved in vivo clearance of infectious bronchitis virus. *J Virol*. 1997;71(7):5173–7.
- Cavanagh D. Coronavirus avian infectious bronchitis virus. *Vet Res*. 2007; 38(2):281–97.
- Ignjatovic J, Sapats S. Avian infectious bronchitis virus. *Revue Sci Tech Office Int des Epizooties*. 2000;19(2):493–501.
- Raj GD, Jones R. Infectious bronchitis virus: immunopathogenesis of infection in the chicken. *Avian Pathology*. 1997;26(4):677–706.
- Yang T, Wang H-N, Wang X, Tang J-N, Lu D, Zhang Y-F, et al. The protective immune response against infectious bronchitis virus induced by multi-epitope based peptide vaccines. *Biosci Biotechnol Biochem*. 2009;0864:1–5.
- Abro SH, Ullman K, Belák S, Baule C. Bioinformatics and evolutionary insight on the spike glycoprotein gene of QX-like and Massachusetts strains of infectious bronchitis virus. *Virol J*. 2012;9(1):211.
- Abdel-Moneim A, Madbouly H, Gelb J, Ladman B. Isolation and identification of Egypt/Beni-Suef/01 a novel genotype of infectious bronchitis virus. *Vet Med J IZA*. 2002;50(4):1065–78.
- Mo M-L, Hong S-M, Kwon H-J, Kim I-H, Song C-S, Kim J-H. Genetic diversity of spike, 3a, 3b and e genes of infectious bronchitis viruses and emergence of new recombinants in Korea. *Viruses*. 2013;5(2):550–67.
- Miller MM, Taylor RL Jr. Brief review of the chicken major histocompatibility complex: the genes, their distribution on chromosome 16, and their contributions to disease resistance. *Poultry Sci*. 2016;95(2):375–92.
- Shina T, Hosomichi K, Hanzawa K. Comparative genomics of the poultry major histocompatibility complex. *Anim Sci J*. 2006;77(2):151–62.
- Bande F, Arshad SS, Hair Bejo M, Kadkhodaei S, Omar AR. Prediction and in silico identification of novel B-cells and T-cells epitopes in the S1-spike glycoprotein of M41 and CR88 (793/B) infectious bronchitis virus serotypes for application in peptide vaccines. *Adv Bioinforma*. 2016;2016:5484972.
- Dash R, Das R, Junaid M, Akash MF, Islam A, Hosen SZ. In silico-based vaccine design against Ebola virus glycoprotein. *Adv Appl Bioinform Chem*. 2017;10:11.
- Gaafar B, Ali SA, Abd-elrahman KA, Almofti YA. Immunoinformatics approach for multiepitope vaccine prediction from H, M, F, and N proteins of Peste des Petits ruminants virus. *J Immunol Res*. 2019;2019:1–18.
- Zheng J, Lin X, Wang X, Zheng L, Lan S, Jin S, Ou Z, Wu J. In silico analysis of epitope-based vaccine candidates against hepatitis B virus polymerase protein. *Viruses*. 2017;9(5):112.
- Ali SA, Almofti YA, Abd-elrahman KA. Immunoinformatics approach for multiepitope vaccine prediction against glycoprotein B of avian infectious laryngotracheitis virus. *Adv Bioinforma*. 2019;2019:1–23.
- Gasteiger EGA, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*. 2003;31:3784–8.
- Buchan DWMF, Nugent TC, Bryson K, Jones DT. Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res*. 2013;41:W349–57.
- Ferre F, Clote P. Disulfide connectivity prediction using secondary structure information and diresidue frequencies. *Bioinformatics*. 2005;21(10):2336–46.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389–402.
- Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DL, Jackson JD, Ke Z. CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res*. 2007;35:D237–40.
- Eddy SR. Profile hidden Markov models. *Bioinformatics (Oxford, England)*. 1998;14(9):755–63.
- El-Gebali S, et al. The Pfam protein families database in 2019. *Nucleic Acids Res*. 2019;47(D1):D427–32.
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 2009;25(9):1189–91.
- Bui H-H, Sidney J, Li W, Fusseder N, Sette A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 2007;8(1):361.
- Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33(7):1870–4.
- Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res*. 2014; 43(D1):D405–D12.
- Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res*. 2006;2(1):2.
- Emini EA, Hughes JV, Perlow D, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol*. 1985;55(3):836–9.
- Kolaskar A, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*. 1990;276(1–2): 172–4.
- Sun P, Ju H, Liu Z, Ning Q, Zhang J, Zhao X, et al. Bioinformatics resources and tools for conformational B-cell epitope prediction. *Comput Math Methods Med*. 2013;2013:943636.
- Chan WM, Rogers SE, Nash SM, Buning PG, Meakin R. User's manual for Chimera grid tools, version 1.8. NASA Ames Research Center, URL: <http://people.nas.nasa.gov/~rogers/cgt/doc/man.html>. 2003.
- Patronov A, Doychinova I. T-cell epitope vaccine design by immunoinformatics. *Open Biol*. 2013;3(1):120139.
- Abdelbagi M, Hassan T, Shihabeldin M, Bashir S, Ahmed E. Immunoinformatics prediction of peptide-based vaccine against African horse sickness virus. *Immunome Res*. 2017;13(135):2.
- Nielsen M, Lund O, et al. BMC Bioinformatics. 2009;10(1):296.

35. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics.* 2007;8(1):1–7.
36. Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v. 2—a server for in silico prediction of allergens. *J Mol Model.* 2014;20(6):2278 1–6.
37. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GP, et al. In silico approach for predicting toxicity of peptides and proteins. *PLoS One.* 2013;8(9):1–10.
38. Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc.* 2012; 7(8):1511–22.
39. Peng J, Xu J. RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins.* 2011;79(S10):161–71.
40. Peng J, Xu J. A multiple-template approach to protein threading. *Proteins.* 2011;79(6):1930–9.
41. Maupetit J, Derreumaux P, Tufféry P. A fast method for large-scale De novo peptide and miniprotein structure prediction. *J Comput Chem.* 2010;31(4): 726–38.
42. Beaufays J, Lins L, Thomas A, Brasseur R. In silico predictions of 3D structures of linear and cyclic peptides with natural and non-proteinogenic residues. *J Pept Sci.* 2012;18(1):17–24.
43. 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):4745–58.
44. Andrusier N, Nussinov R, Wolfson HJ. FireDock: fast interaction refinement in molecular docking. *Proteins.* 2007;69(1):139–59.
45. Shang J, Zheng Y, Yang Y, Liu C, Geng Q, Luo C, et al. Cryo-EM structure of infectious bronchitis coronavirus spike protein reveals structural and functional evolution of coronavirus spike proteins. *PLoS Pathog.* 2018;14(4): 1–9.
46. Verma SK, Yadav S, Kumar A. In silico prediction of B-and T-cell epitope on Lassa virus proteins for peptide based subunit vaccine design. *Adv Biomed Res.* 2015;4:201.
47. Reche PA, Fernandez-Caldas E, Flower DR, Fridkis-Hareli M, Hoshino Y. Peptide-based immunotherapeutics and vaccines. *J Immunol Res.* 2014; 2014:1–2.
48. Sylvester SA, Dhama K, Kataria J, Rahul S, Mahendran M. Avian infectious bronchitis: a review. *Indian J Comp Microbiol Immunol Infect Dis.* 2005;26:1–14.
49. Idris S, Salih S, Basheir M, Elhadi A, Kamel S, Abd-elrahman K, et al. In silico prediction of peptide based vaccine against Fowlpox virus (FPV). *Immunome Res.* 2018;14(2):1–11.
50. Badawi MM, Fadli Alla A, Alam SS, Mohamed WA, Osman D, Alraziq Ali S, et al. Immunoinformatics predication and in silico modeling of epitope-based peptide vaccine against virulent Newcastle disease viruses. *Am J Infect Dis Microbiol.* 2016;4(3):61–71.
51. Bashir S, Abd-elrahman KA, Hassan MA, Almofti YA. Multi epitope based peptide vaccine against Marek's disease virus serotype 1 glycoprotein H and B. *Am J Microbiol Res.* 2018;6(4):124–39.
52. Gen J. Cloning and sequencing of the gene encoding the spike protein of the coronavirus IBV. *J Gen Virol.* 1985;66:719–26.
53. Yao B, Zheng D, Liang S, Zhang C. Conformational B-cell epitope prediction on antigen protein structures: a review of current algorithms and comparison with common binding site prediction methods. *PLoS One.* 2013;8(4):1–4.
54. Sanchez-Trincado JL, Gomez-Perez M, Reche PA. Fundamentals and methods for T-and B-cell epitope prediction. *J Immunol Res.* 2017;2017: 1–14.
55. Demolombe V, de Brevern AG, Felicori L, NGuyen C, de Avila RAM. Valera L, et al. PEPOP 2.0: new approaches to mimic non-continuous epitopes. *BMC Bioinformatics.* 2019;387:1–14.
56. Pei J, Briles WE, Collisson EW. Memory T cells protect chicks from acute infectious bronchitis virus infection. *Virology.* 2003;306(2):376–84.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](http://biomedcentral.com/submissions)

