Bioinformatics analysis on potential anti-viral targets against spike protein of MERS-CoV

Subtitle: Potential epitopes in MERS-CoV S Protein

Yan-Hua Li[a,b#], Hainv Gao[a,b,c#], Yunfeng Xiao[d#], Tianhao Weng[a,b], Dongshan Yu[a,b], Chenyu Hu[a,b], Hang-Ping Yao[a,b*], Lan-Juan Li[a,b,c*]

[a]State Key Laboratory for Diagnosis and Treatment of Infectious Diseases; [b]Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, the First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310031, China; [c]Shulan (Hangzhou) Hospital; [d]Department of Pharmacy, Tangdu Hospital, Xi'an, Shaanxi 710038, PR China

*Corresponding authors:

Lan-Juan Li, Professor, Tel/Fax: 86-571-87236582, Email: ljli@zju.edu.cn; Hang-Ping Yao, Professor. Email: yaohangping@zju.edu.cn

[#] These authors contribute equally to this paper.

Abstract **Objective** Middle East respiratory syndrome is caused by the Middle respiratory syndrome coronavirus (MERS-CoV) and the mortality is high. However, to date, there is no effective vaccine or antibody for human immunity and treatment as a putative therapeutic agent specific for MERS. The aim of this study was to obtain the bioinformatic characteristics of the MERS-CoV S protein antigen. **Methods** SOPMA Server software and the DiscoTope were used to predict the secondary and tertiary structures of the MERS-CoV S protein, respectively, whilst a number of online prediction software applications, including IEDB, Syfpeithi and other resources of IEDB, were used for the T- and B-cell epitope predictions. **Results** The prediction results indicated that the T-cell epitopes were located at positions 950-958, 317-325, 1309-1317, 480-488 and 388-396, whereas the B-cell epitopes were located at positions 520-528, 629-637, 659-667, 734-744, 1205-1212. 19-53, 300-309, 478-523, 528-550 and 622-632. **Conclusion** These regions were the potential dominant epitopes of the MERS-CoV S protein antigen. The results of our study provide experimental data for the identification and screening of epitopes and may be used for the development of epitope vaccines that have an enhanced safety and efficacy. This may result in the provision of improved regimens for the prevention and treatment of MERS.

Keywords MERS-CoV, spike protein, secondary structure, tertiary structure, anti-viral targets

INTRODUCTION

Coronavirus (CoV) is an enveloped, positive, single-stranded RNA virus that causes mild upper respiratory infections in humans. Middle East respiratory syndrome (MERS) is a viral respiratory disease caused by a de nove coronavirus (Middle East respiratory syndrome coronavirus, MERS-CoV). To date, about 36% (40.43% in Saudi Arabia) of patients with MERS have died[1]. Although rapid diagnostic and public health measures are currently being implemented, new cases of MERS-CoV infection are still being reported. Therefore, various effective measures should be taken to prevent the serious impact of similar epidemics in the future [2, 3]. SARS-CoV and MERS-CoV are highly pathogenic and lead to high mortality, especially in infants, the elderly, and immunocompromised patients [4, 5]. MERS has broken out in 27 countries around the world. The major outbreaks occurred in Saudi Arabia, the United Arab Emirates and South Korea, which accounted for more than 75% of the MERS prevalence. To date, in addition to the safety and tolerability of the fully human polyclonal IgG antibody (SAB-301) being described as a putative therapeutic agent specific for MERS, no effective vaccines or antibodies have been reported for human immunization and therapy [6]. The combination of computational biology and virology can accelerate advanced design and develop effective peptide therapeutics against MERS-CoV.

Many coronaviruses, including MERS-CoV, have a very large RNA genome of about 30 kb with at least 10 predicted open reading frames (ORFs) [3, 7]. 5′-replicase-structural proteins (spike-envelope-membrane-nucleocapsid)-poly(A)-3′ [i.e., 5′-ORF1a/b-S-E-M-N-poly(A)-3′] is similar to that of other CoVs and unambiguously distinguishes MERS-CoV from lineage A CoVs [8]. Many of these genes and their encoded proteins are useful diagnostic, therapeutic or vaccination targets (Table 1) [9, 10].

The S protein is a transmembrane glycoprotein sorting to type I that is expressed on the surface of the envelope of virus and forms spikes from the virus body. S contains 1353 amino acids, is heavily glycosylated, and consists of a large extracellular domain and a short cytoplasmic tail. S protein can be subdivided to S1 and S2 that take important effect in virus binding, fusion and entry.

MATERIALS AND METHODS

Amino acid sequence of the MERS-CoV spike protein(S protein). The nucleotide sequence of *S protein* was determined using GenBank (GenBank no. KT182957.1; http://www.ncbi.nih.gov/ genbank/). According to GenBank, the *S protein* is composed of 1353 amino acid residues, encoded by the 1-4062 region of the S protein mRNA.

Prediction of the secondary structure of the S protein. The secondary structure of the S protein was predicted by the improved self-optimized prediction method (SOPMA) software (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) (20). The protein sequence of the S protein was input, and four conformational states, including helices, sheets, turns and coils, were analyzed. The parameters of similarity threshold and window width were set to 8 and 17, respectively, whilst the remaining parameters were not adjusted.

Prediction of the T-cell epitopes for the S protein. The major histocompatibility complex (MHC)-I human leukocyte antigen (HLA)-A*0201-restricted T-cell epitopes were predicted using online prediction software from the Immune Epitope Database (IEDB; http://tools.immuneepitope.org/ main/index.html) (21) and Syfpeithi (http://www.syfpeithi.de). The protein sequence of the S protein was input, and the parameters were adjusted so that 'MHC allele(s)' was set at HLA-A*02:01, and 'length' was set at 9. The remaining parameters were not altered.

Prediction of the B-cell epitopes for the S protein. The B-cell epitopes of the S protein were predicted using IEDB. The surface accessibility was calculated with Emini surface accessibility scale[11] The accessibility profile was obtained using the formulae $Sn = (n+4+i)\ (0.37)^{-6}$ where Sn is the surface probability, dn is the fractional surface probability value, and i vary from 1 to 6. A hexapeptide sequence with Sn

greater than 1.0 indicates an increased probability for being found on the surface. Kolaskar and Tongaonkar antigenicity scale[12] was used to predict antigenic determinants.A semi-empirical method which makes use of physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes was developed to predict antigenic determinants on proteins. Application of this method to a large number of proteins has shown by the authors that the method can predict antigenic determinants with about 75% accuracy which is better than most of the known methods. *Bepipred Linear Epitope Prediction 2.0[13]* predicts the location of linear B-cell epitopes using a combination of a hidden Markov model and a propensity scale method. The residues with scores above the threshold (the average and threshold is 0.474) are predicted to be part of an epitope and colored in yellow on the graph (where Y-axes depicts residue scores and X-axes residue positions in the sequence). The values of the scores are not affected by the selected threshold.

Prediction of the tertiary structure of the S protein.
Predictive analysis of the S protein tertiary structure was conducted using DiscoTope, a method for predicting discontinuous epitopes from 3D structures of proteins in PDB format. After entering the structure either by entering its PDB id, then enter the Chain id for the protein chain of interest, click on submit. It is a useful tool for the analysis of protein tertiary structure. The site uses ligands from similar structures to make predictions, as well as providing details of conservation information. The default value for version 1.1 is -7.7 and version 2.0 is -3.7, which corresponds to a specificity of 75%. Higher values correspond to higher specificity. A specificity of 0.75 means that 25% of the nonepitope residues were predicted as part of epitopes. A sensitivity of 0.47 means that 47% of the epitope residues were predicted as part of epitopes.

RESULTS

Prediction of the secondary structure of the MERS-CoV S protein. In order to assess the antigenic features of the MERS-CoV S *protein, we predicted its secondary structure using SOPMA Server software. A greater proportion of extended strands and random coils present in the structure of the protein corresponded with an increased likelihood of the protein forming an antigenic epitope. The predicted secondary structure results for the* MERS-CoV S *prote*in are demonstrated in Figure 1. The results revealed that the proportion of random coils, β turns, α helices and extended strands (β folds) accounted for 34.89, 8.06, 35.40 and 21.66% of the secondary structure, respectively.

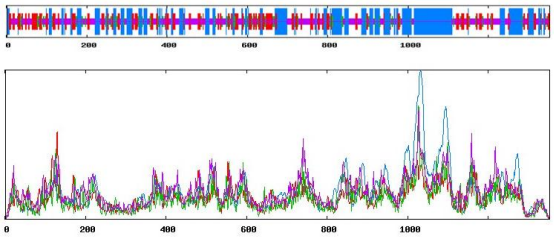

Figure 1. Secondary structure prediction results for the MERS-CoV S protein. The improved self-optimized prediction method (SOPMA) software (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) was used to predict the secondary structure of the MERS-CoV S protein. An increased number of extended strands and random coils in the protein corresponded with an increased likelihood of the protein forming an antigenic epitope. Lines in different colors represent different secondary structures: Blue, α helix; green, β turn; red, extended strand; and purple, random coil. Parameters: Window width 17, Similarity threshold 8, Number of states 4

Prediction of the T-cell epitopes for the MERS-CoV S protein. In order to develop an epitope vaccine, it is essential to determine the precise location of the epitope. In the current study, the MHC I HLA-A*0201-restricted T-cell epitopes were predicted using two online prediction software applications, IEDB and Syfpeithi, which represented the probability of a particular region forming a T-cell epitope by a score, Table 1. It is worthy of note that the two software solutions utilized different scoring systems. As predicted by the IEDB software, each row in this table corresponds to one peptide binding prediction. The predicted output is given in units of IC50nM. Therefore, a lower number indicates higher affinity. As a rough guideline, peptides with IC50 values <50 nM are considered high affinity, <500 nM intermediate affinity. Most known epitopes have high or intermediate affinity. However, the high scores predicted by the Syfpeithi software ranged between 20 and 35. Despite the difference, these high-scoring regions all had strong potential as epitope regions. The results from the IEDB and Syfpeithi software indicated that the T-cell epitopes were listed below. The five highest-scoring regions, selected from the combined results of the two software applications, were 950-958, 317-325, 1309-1317, 480-488 and 388-396.

Prediction of the B-cell epitopes for the MERS-CoV S protein antigen. Emini surface accessibility scale was used to calculate the surface accessibility. A hexapeptide sequence with Sn greater than 1.0 indicates an increased probability for being found on the surface(Table 2 and Figure 2A). Kolaskar and Tongaonkar antigenicity scale was used to predict antigenic determinants. A semi-empirical method which makes use of physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes was developed to predict antigenic determinants on proteins(Table 2 and Figure 2B). We combined the two methods and got positions with high possibility in forming epitopes: 511-518, 520-528, 629-637, 659-667, 734-744 and 1205-1212.

Bepipred Linear Epitope Prediction 2.0 were used for predicting linear B-cell epitopes with a combination of a hidden Markov model and a propensity scale method. The residues with scores above the threshold (the average and threshold is 0.474) are predicted to be part of an epitope. The results are displayed in Figure 2C. A combination of the results predicted by the different methods indicated that the potential B-cell epitopes of the MERS-CoV S protein antigen were located at positions 19-53 (VDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKA), 300-309 (IRSIQSDRKA), 478-523 (CLILATVPHNLTTITKPLKYSYINKCSRLLSDDRTEVPQ LVNANQY), 528-550 (STVPSTVWEDGDYYRKQLSPLEG) and 622-632 (AVGVRQQHFVY).

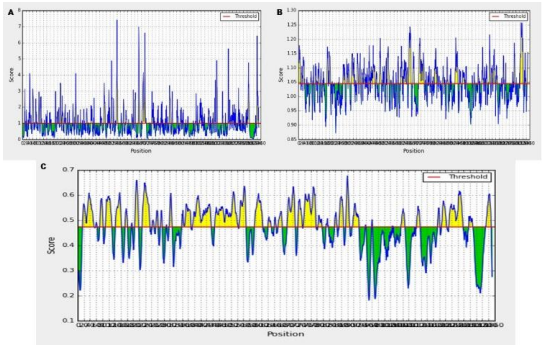

Figure 2. The residues with scores above the threshold are predicted to be part of an epitope, Y-axes depicts residue scores and X-axes residue positions in the sequence. The larger score for the residues might be interpreted as that the residue might have a higher probability to be part of epitope (those residues are colored in yellow).

Prediction of the tertiary structure of the MERS-CoV S protein epitopes. The tertiary structure of the MERS-CoV S protein was obtained using the Disco Tope online software. The results of the predicted conformations of the epitopes are displayed in Figure 3. Contact Number is the number of Cα atoms in the antigen within a distance of 10 Å of the residue's Cα atom. A low contact number correlates with localization of the residue close to the surface or in protruding regions of the antigen's structures. Propensity Score showed the probability/tendency of being part of an epitope for that particular residue. Then the scores are summed up based on the proximity in the 3D structure of the antigen. For any given residue, the sequentially averaged log-odds scores from all residues within 10Å are summed to give the propensity score. Discotope Score is

calculated by combining the contact numbers with propensity score. DiscoTope score above the threshold value indicates positive predictions.

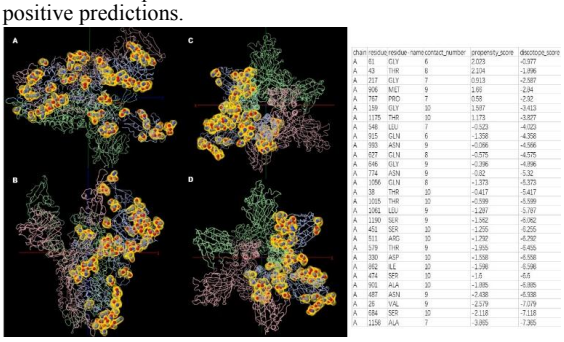

Figure 3. The 3D view uses Jmol to display the structure with positive predictions highlighted in yellow. The side chain of each predicted residue is shown. You can rotate, zoom and manipulate the structure by using different buttons on the mouse. The table lists the predicted epitope residues along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

Table 1 Analysis of the MHC I HLA-A*0201-restricted T-cell epitopes using IEDB and Syfpeithi online prediction software

IEDB MHC I HLA-A*0201, 9aa					Syfpeithi HLA-A*02:01 (9aa)		
start	end	peptide	percentile-rank	ann_ic50	Pos	1 2 3 4 5 6 7 8 9	Score
442	450	ILDYFSYPL	0.3	4.33	950	S L L G S I A G V	33
950	958	SLLGSIAGV	0.3	5.91	1070	L I N G R L T T L	29
1258	1266	TLLDLTYEM	0.4	6.27	171	L L P D G C G T L	27
317	325	KLQPLTFLL	0.4	7.83	317	K L Q P L T F L L	26
1309	1317	ALALCVFFI	0.2	14.62	794	Y I Q T T I Q K V	26
7	15	LLMFLLTPT	0.8	16.26	975	S I F Y R L N G V	26
567	575	LQMGFGITV	1.2	18.48	1268	S L Q Q V V K A L	26
716	724	GLVNSSLFV	0.8	18.63	1309	A L A L C V F F I	26
1074	1082	RLTTLNAFV	0.8	19.74	388	L L S G T P P Q V	25
893	901	LLFDKVTIA	0.8	21.91	480	I L A T V P H N L	25
1303	1311	FIAGLVALA	1.2	24.38	506	L L S D D R T E V	25
480	488	ILATVPHNL	1.2	30.1	721	S L F V E D C K L	25
388	396	LLSGTPPQV	0.9	36.02	737	A L P D T P S T L	25
630	638	FVYDAYQNL	2	36.76	1032	A L S K L A S E L	25
110	118	KQFANGFVV	1.8	43.52	1050	S I G D I I Q R L	25
178	186	TLLRAFYCI	1.5	43.87	1192	Y A P E P I T S L	25
11	19	LLTPTESYV	0.5	46.24	1275	A L N E S Y I D L	25
321	329	LTFLLDFSV	2.1	51.45	295	I I P H S I R S I	24

Table 2. High possibility of the surface accessibility and antigenic determinants

Emini surface accessibility scale						Kolaskar and Tongaonkar antigenicity scale					
Position	Residue	Start	End	Peptide	Score	Position	Residue	Start	End	Peptide	Score
541	Y	539	544	DYYRKQ	7.419	1311	A	1308	1314	VALALCV	1.258
665	K	663	668	YDKETK	6.977	1316	F	1313	1319	CVFFILC	1.256
700	R	698	703	KRRDST	6.609	656	S	653	659	ACVSVPV	1.243
1340	Y	1338	1343	DRYEEY	6.425	652	R	649	655	YCLRACV	1.222
1176	N	1174	1179	KTNNTR	5.628	1125	S	1122	1128	HIVSFVV	1.216
1110	Q	1108	1113	KAQSKR	4.899	926	A	923	929	LICAQYV	1.205
510	D	508	513	SDDRTE	4.88	714	V	711	717	VGCVLGL	1.204
1291	Y	1289	1294	YTYYNK	4.763	481	L	478	484	CLILATV	1.203
540	Y	538	543	GDYYRK	4.239	1320	C	1317	1323	ILCCTGC	1.203
691	R	689	694	YSRSTR	4.156	1270	Q	1267	1273	LSLQQVV	1.187
43	T	41	46	DKTWPR	4.094	417	L	414	420	LLSLFSV	1.178
306	D	304	309	QSDRKA	4.091	440	S	437	443	CYSSLIL	1.178

DISCUSSIONS

The secondary structure of a protein is closely correlated with its epitope distribution. Hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and antigenic propensity of polypeptides chains have been correlated with the location of continuous epitopes. This has led to a search for empirical rules that would allow the position of continuous epitopes to be predicted from certain features of the protein sequence. In order to avoid the risk of the recognition region being hidden inside the protein, we usually recommend selecting the corresponding antibody at both N terminal and C terminal of proteins. Because in intact proteins, both ends of N and C are usually exposed to the surface of the protein. However, it is important to note that the C-terminal hydrophobicity of membrane proteins is too strong to be suitable as an antigen. In our study we mainly analyzed surface accessibility, antigenic propensity and liner epitopes of MERS-CoV S protein.

HR1P positioning at 998-1038 and HR2P positioning at 1251-1286 were found by Lu et al respectively, to form a stable six-helix bundle (6-HB) as fusion core[14] [15]. The HR2P peptide shows activity against MERS-CoV by inhibiting viral replication in calu-3 and HFL cells [14]. HR2P inhibits MERS-CoV replication and its stimulatory cell-cell fusion. Although there is not much knowledge of the assembly and release, inhibiting these virus process may be a good goal for the future [16, 17]. Recently, using low-temperature electron microscope, revealed a global prefusional structure of the full-length S-exodomain, indicating that the NTD and C-terminal domains (CTDs) form a "V" shape, contributing to the overall triangular appearance of the S-trimer [18-20]. The S2 subunit is linked to the viral membrane and is characterized by the presence of a long alpha-helix. This structural information might be a basis for the structure-based immunogen design of β-CoV vaccines. The combination of computational biology and virology could definitely accelerate advanced designs against MERS-CoV, resulting in higher efficacy [17].

The research and development of epitope vaccines is a difficult and highly targeted technology, which comprehensively utilizes molecular biology and immunology. A key step in the preparation of the vaccines is obtaining the necessary information concerning the epitope. In recent years, with the development of bioinformatics, epitope prediction has improved in simplicity and significance. Performing predictions with a multi-parameter and -method analysis greatly enhances the accuracy of the epitope prediction.

Antigenicity are the primary factors involved in epitope formation, although interrelated factors, such as the exposed surface area and the conformation of the secondary structure, are also important. Thus, we anaalyzed the secondary structure of the MERS-CoV S protein in order to obtain the antigenic features of the protein. However, as the combined effects of T- and B-cells are necessary for the elimination of antigens, it was also important to analyze the T- and B-cell epitopes of the MERS-CoV S protein. Therefore, we predicted the T- and B-cell epitopes using a multi-parameter and -method analysis. A comprehensive analysis of this nature may, in the future, improve the accuracy and specificity of epitope prediction. The α helices and β sheets, in the secondary structure of proteins, are very regular structures, and are not readily deformed. This is due to the presence of hydrogen bonds, which act to maintain structural stability. However, α helices and β sheets are usually located inside the protein, which is difficult for ligand binding. By contrast, the β turn and the random coil regions are located on the surface of the protein, where it is necessary for the surface structure to make appropriate changes to meet the functional needs of the protein. Therefore, these structures are suitable for binding ligands, and have a high possibility of forming epitopes. As analyzed by SOPMA Server software, the proportions of α helices and β sheets were 35.40% and 21.66%, respectively. This result indicates that the MERS-CoV S protein had a good stability. Random coils and β turns, which represented the potential epitope regions, accounted for 34.89% and 8.06% of the protein, respectively. high proportions of random coils suggested strong antigenicity.

The accuracy rate of the MHC I epitope prediction in the prediction of T-cell epitopes has been demonstrated to be up to 90% (27). Therefore, in the present study, the HLA-A*0201-restricted epitopes of the MERS-CoV S protein were analyzed using the IEDB and Syfpeithi online software applications. The T-cell epitopes of the MERS-CoV S protein were predicted to be located at positions 950-958, 317-325, 1309-1317, 480-488 and 388-396, as these regions exhibited high scores. In order to improve the accuracy of the B-cell epitope prediction for the MERS-CoV S protein, a multi-method and-parameter analysis was utilized. The antigenic propensity analysis demonstrated the immunogenic regions of the antigen. Potentially dominant epitopes are particularly likely to be located in regions with a high antigenic propensity. The exposed surface areas of the antigen have an enhanced probability of coming into contact with the solvent molecules. In combination with these parameters, two online software applications were used to predict the B cell epitopes. Ten potential epitope regions were revealed, located at positions 520-528, 629-637, 659-667, 734-744, 1205-1212. 19-53, 300-309, 478-523, 528-550 and 622-632. Protein tertiary structure, one of the higher-order structures of the protein, is a three-dimensional conformation of the naturally folded protein. Tertiary structure has a globular conformation that is formed by the further coiling and folding of the secondary structure. Therefore, the prediction of the tertiary structure was a useful supplement to the prediction of the MERS-CoV S protein antigenic epitopes. The results of our study provide experimental data for the identification and screening of epitopes and may be used for the development of epitope vaccines that have an enhanced safety and efficacy. This may result in the provision of improved regimens for the prevention and treatment of MERS.

Conflicts of interest

The authors have no conflicts of interest to declare.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (grant number 2016YFC1200204) and the Major Program of National Natural Science Foundation of China (grant number 81590763).

References

1 Hui DS, Azhar EI, Kim YJ, Memish ZA, Oh MD, Zumla A. Middle East respiratory syndrome coronavirus: risk factors and determinants of primary, household, and nosocomial transmission. The Lancet Infectious diseases 2018,

2 Lee H, Lei H, Santarsiero BD, Gatuz JL, Cao S, Rice AJ, Patel K, *et al.* Inhibitor recognition specificity of MERS-CoV papain-like protease may differ from that of SARS-CoV. ACS Chem Biol 2015, 10: 1456-1465

3 Choi J, Kim MG, Oh YK, Kim YB. Progress of Middle East respiratory syndrome coronavirus vaccines: a patent review. Expert Opin Ther Pat 2017, 27: 721-731

4 Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 2012, 367: 1814-1820

5 Pallesen J, Wang N, Corbett KS, Wrapp D, Kirchdoerfer RN, Turner HL, Cottrell CA, *et al.* Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. Proc Natl Acad Sci U S A 2017, 114: E7348-E7357

6 Beigel JH, Voell J, Kumar P, Raviprakash K, Wu H, Jiao JA, Sullivan E, *et al.* Safety and tolerability of a novel, polyclonal human anti-MERS coronavirus antibody produced from transchromosomic cattle: a phase 1 randomised, double-blind, single-dose-escalation study. The Lancet Infectious diseases 2018,

7 van Boheemen S, de Graaf M, Lauber C, Bestebroer TM, Raj VS, Zaki AM, Osterhaus AD, *et al.* Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. MBio 2012, 3

8 Frey KG, Redden CL, Bishop-Lilly KA, Johnson R, Hensley LE, Raviprakash K, Luke T, *et al.* Full-genome sequence of human betacoronavirus 2c jordan-n3/2012 after serial passage in Mammalian cells. Genome Announc 2014, 2

9 Chan JF, Lau SK, To KK, Cheng VC, Woo PC, Yuen KY. Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. Clin Microbiol Rev 2015, 28: 465-522

10 Cotten M, Lam TT, Watson SJ, Palser AL, Petrova V, Grant P, Pybus OG*, et al.* Full-genome deep sequencing and phylogenetic analysis of novel human betacoronavirus. Emerg Infect Dis 2013, 19: 736-742B

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

12 Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 1990, 276: 172-174

13 Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. Immunome research 2006, 2: 2

14 Lu L, Liu Q, Zhu Y, Chan KH, Qin L, Li Y, Wang Q*, et al.* Structure-based discovery of Middle East respiratory syndrome coronavirus fusion inhibitor. Nat Commun 2014, 5: 3067

15 Liu S, Wu S, Jiang S. HIV entry inhibitors targeting gp41: from polypeptides to small-molecule compounds. Curr Pharm Des 2007, 13: 143-162

16 Fehr AR, Perlman S. Coronaviruses: an overview of their [illegible]sis. Methods Mol Biol 2015, 1282: 1-23

17 Mustafa S, Balkhy H, Gabere MN. Current treatment options and the role of peptides as potential therapeutic components for Middle East Respiratory Syndrome (MERS): A review. J Infect Public Health 2018, 11: 9-17

18 Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, Corbett KS*, et al.* Pre-fusion structure of a human coronavirus spike protein. Nature 2016, 531: 118-121

19 [illegible]MA, Frenz B, Snijder J, Li W, Rey FA, DiMaio F*, et al.* Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. Nat Struct Mol Biol 2016, 23: 899-905

20 Walls A, Tortorici MA, Bosch BJ, Frenz B, Rottier PJ, DiMaio F, Rey FA*, et al.* Crucial steps in the structure determination of a coronavirus spike glycoprotein using cryo-electron microscopy. Protein Sci 2017, 26: 113-121