pISSN 2302-1616, eISSN 2580-2909
Vol 8, No. 2, December 2020, pp. 185-194
Available online http://journal.uin-alauddin.ac.id/index.php/biogenesis
DOI https://doi.org/10.24252/bio.v8i2.15696

Analysis of Amino Acid Sequence of SARS-CoV, SARS-CoV-2, and MERS-CoV Spike Glycoproteins: Preliminary Study for Obtaining Universal Peptide Vaccine Candidates

Yani Suryani*, Opik Taupiqurrohman
Department of Biology, Faculty of Science and Technology,
Universitas Islam Negeri Sunan Gunung Djati Bandung
Jl. A.H. Nasution No. 105A, Bandung, West Java, Indonesia. 40614
*Email: yani.suryani@uinsgd.ac.id

ABSTRACT. In the manufacture of universal peptide vaccines, it is necessary to analyze the amino acids of the various candidates. Therefore, this study aims to examine the amino acids of the spike glycoproteins of SARS-CoV, SARS-CoV-2, and MERS CoV. The method used is the alignment of the amino acid spike glycoprotein between SARS-CoV with SARS-CoV-2, MERS CoV with SARS-CoV-2, and SARS-CoV with MERS-CoV using web-based software water emboss. The analysis result showed that SARS and SARS CoV-2 were very similar with 87% similarity and 76.4% identity values. In contrast, SARS CoV-2 with MERS and SARS with MERS were very different, having similarity and identity values less than 70%. Therefore, it is reasonable to conclude that spike glycoprotein's peptide is only useful from attacks by the SARS-CoV and SARS-CoV-2 viruses.

Keywords: Coronavirus; COVID-19; MERS, peptide vaccine; SARS

Article History: Received 04 September 2020; Received in revised form 30 October 2020; Accepted 23 November 2020; Available online 30 December 2020
How to Cite This Article: Suryani Y, Taupiqurrohman O. 2020. Analysis of amino acid sequence of SARS-CoV, SARS-CoV-2, and MERS-CoV spike glycoproteins: preliminary study for obtaining universal peptide vaccine candidates. *Biogenesis: Jurnal Ilmiah Biologi*. vol 8(2): 185-194. doi: https://doi.org/10.24252/bio.v8i2.15696.

INTRODUCTION

Coronavirus is 65 to 125 nm in size and as a single-stranded RNA virus, it has the length of 26-32 kbp. Furthermore, it consists of four subfamilies of alpha (α), beta (β), gamma (ϒ), and delta (δ) (Shereen *et al*., 2020). Before now, the virus only infected animals, however, in the 2000s, there were known cases of SARS (Severe Acute Respiratory Syndrome), MERS (Middle East Respiratory Syndrome), and COVID-19 (Coronavirus Disease-2019) (Wang *et al*., 2013; Shereen *et al*., 2020). The Coronaviruses that cause SARS, MERS and COVID-19 are known as SARS-CoV (SARS-coronavirus), MERS CoV (MERS-coronavirus) and SARS CoV 2 (COVID-19) respectively.

SARS-CoV is an etiologic agent of acute respiratory syndrome (Liu *et al*., 2014), and was endemic in 2002-2003 (Peiris *et al*., 2003). MERS (Middle East Respiratory Syndrome) is an infectious disease caused by the MERS virus (MERS-CoV), and it has similar symptoms with SARS, that cause flu-like illnesses which is respiratory tract disruption (da Costa *et al*., 2020; Li *et al*., 2020). Both belong to the genus *Betacoronavirus* (Schoeman & Fielding, 2019; ICTV, 2020). Initially, MERS-CoV case was first identified in Saudi Arabia in September 2012 by camel-to-human transmission, and then by human-to-human transmission (Jeong *et al.*, 2017) and spread to countries in the Middle East (Zumla *et al*., 2015), precisely in the Arabian Peninsula and its surroundings such as the United Arab Emirates, Qatar, Oman, Jordan, Kuwait, Yemen, Iran, Egypt, and Lebanon (Rampengan, 2016; Shapiro *et al*., 2016; WHO, 2016). Importations of MERS had been reported in France (Mailles *et al*., 2013; Puzelli *et al*., 2013), Italy (Puzelli *et al*., 2013), Spain (Rashid *et al*., 2013), United Kingdom (Thomas *et al.*, 2014), Tunisia (Abroug *et al*., 2014), Malaysia (Cunha & Opal, 2014), Philippines (Racelis *et al*., 2015), and Korea (Jeong *et al*., 2017) in the following years. The SARS-CoV-2 attack resurfaced at the end of 2019 in Wuhan, China, and has spread to almost all countries worldwide, and this condition was given the term COVID-19 (Corona Virus Disease-2019) (ICTV, 2020).

Vaccines for MERS, SARS and COVID-19 are yet to be discovered (Slamet *et al*., 2013) but

the process continues to experience development. In addition, it consists of several amino acid residues (small proteins) (Subroto *et al*., 2013) and with the existence, it is possible to make a universal vaccine. This can serve a protective function against various types of antigens, which is primarily used for making peptide vaccines. Therefore, a strong similarity protein is needed to make a universal vaccine against coronaviruses (SARS, MERS and COVID-19) (Alouane *et al*., 2020; Khalaj-Hedayati, 2020; Wu *et al*., 2020).

There have been many research related to the manufacture of vaccine candidates, one of which is the HPV (Human Papillomavirus). This study shows that the promising vaccine peptide candidate of the E1 protein obtained from HPV genome is LLITSNINA, from E5 is VLLCVCLLI and from E7 is LLMGTLGIV (Aprilyanto & Sembiring, 2017). Furthermore, they have been tested in vitro, and the results are useful in activating the immune response.

One of the conditions for making a peptide vaccine is that the protein antigen should be located at the outer part of the virus in order to ease the purification process. The spike glycoprotein is used as a peptide vaccine candidate since its position is in the outer part, and it is possessed by all types of Coronavirus. Therefore, this protein is used as a candidate source for peptide vaccines for all kinds of Coronavirus.

Initially, the research was conducted to obtain an overview of its potential as a vaccine candidate. In addition, an explanation of their potential will be obtained by testing the similarity and identity for the amino acid sequence of spike glycoproteins.

MATERIALS AND METHODS

This research used the following materials; amino acid sequence of SARS-CoV-2, SARS-CoV (SARS), and MERS-CoV (MERS) spike glycoprotein with NCBI having the Accession Number YP_009724390.1 (COVID-19), P59594, and A0A140AYZ5 respectively.

Work Procedures. This research was conducted in the following stages: SARS CoV-2 (COVID-19) spike glycoprotein downloaded at https://www.ncbi.nlm.nih.gov/; SARS CoV spike glycoprotein downloaded at http://www.uniprot.org; MERS spike glycoprotein download at http://www.uniprot.org; alignment process between spike glycoproteins of COVID-19 with SARS in the https://www.ebi.ac.uk/Tools/psa/emboss_water/program; alignment between spike glycoproteins of COVID-19 with MERS in the https://www.ebi.ac.uk/Tools/psa/emboss_water/ program; alignment between spike glycoproteins of MERS with SARS in the program https://www.ebi.ac.uk/Tools/psa/emboss_water/ program.

RESULTS AND DISCUSSION

Alignment result between COVID-19 with SARS.

```
##ucu#####################################
# Program: water
# Aligned_sequences: 2
# 1: YP_009724390. (SARS CoV-2/Covid-19)
# 2: SPIKE_CVHSA   (SARS)
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
# Length: 1277
# Identity:     975/1277 (76.4%)
# Similarity:  1111/1277 (87.0%)
# Gaps:          26/1277 (2.0%)
# Score: 5230.0
#=======================================
YP_009724390.      1 MFVFLVLLPLVS-SQCVNLTTRTQL-PPAYT--NSFTRGVYYPDKVFRSS     46
                       ||:||:.|.|.| |.....||...: .|.||  .|..||||||::|||.
SPIKE_CVHSA        1 MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSD     50
```

```
YP_009724390.     47 VLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTE     96
                     .|:.||||||||:||||.||.|     |.|  |.|||:||.||:|||:||
SPIKE_CVHSA       51 TLYLTQDLFLPFYSNVTGFHTI-----NHT--FGNPVIPFKDGIYFAATE     93

YP_009724390.     97 KSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYH    146
                     |||::|||:||:|:::|:||::|:||:|||||:.|.|:.|::||..|
SPIKE_CVHSA       94 KSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAV---    140

YP_009724390.    147 KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKN    196
                      :.....::...::.:|.||||||:|..|.:|:..|.||||:||||||||
SPIKE_CVHSA      141 -SKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKN    189

YP_009724390.    197 IDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHR    246
                     .||:..:|..:.||::||||||.||:.|:|:..||:|||||.|:.:|
SPIKE_CVHSA      190 KDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAIL----    235

YP_009724390.    247 SYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPL    296
                     :..:|....  |...||||:||||:|.||:|||:||||||||||:.:||
SPIKE_CVHSA      236 TAFSPAQDI--WGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPL    283

YP_009724390.    297 SETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATR    346
                     :|.||::|||.::||||||||||||.|:..:||||||||||||||||||:
SPIKE_CVHSA      284 AELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATK    333

YP_009724390.    347 FASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVY    396
                     |.||||||.||:|||||||||||||..||||||||||.|||||||:|||
SPIKE_CVHSA      334 FPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVY    383

YP_009724390.    397 ADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG    446
                     |||||::||:||||||||||.||||||||||||.|||:|||:.|:|:...
SPIKE_CVHSA      384 ADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATST    433

YP_009724390.    447 GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYG    496
                     |||||.||..|...|:|||||||...:.....||. ....|||:||..||
SPIKE_CVHSA      434 GNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCT-PPALNCYWPLNDYG    482

YP_009724390.    497 FQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGL    546
                     |..|.|:|||||||||||||:|||||||||.|||:|:||:|||||||||
SPIKE_CVHSA      483 FYTTTGIGYQPYRVVVLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGL    532

YP_009724390.    547 TGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS    596
                     |||||||.|:|:|.||||||||::|.||:||||:|.|||||:||||||||
SPIKE_CVHSA      533 TGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCSFGGVS    582


YP_009724390.    597 VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTR    646
                     |||||||.|::||||||||||:|..|||||||||.||:||||:||||||:
SPIKE_CVHSA      583 VITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQ    632

YP_009724390.    647 AGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYT    696
                     ||||||||||:.|||||||||||||||.|.:    ..||.:.:||:|||
SPIKE_CVHSA      633 AGCLIGAEHVDTSYECDIPIGAGICASYHTVS----LLRSTSQKSIVAYT    678

YP_009724390.   1247 CCSCGSCCKFDEDDSEPVLKGVKLHYT   1273
                     .||||||||||||||||||||||||||
SPIKE_CVHSA     1229 ACSCGSCCKFDEDDSEPVLKGVKLHYT   1255
```

Fig. 1. Alignment result between SARS CoV-2 with SARS-CoV spike glycoproteins.

Alignment result between COVID-19 with MERS

```
########################################
# Program: water
# Aligned_sequences: 2
# 1: YP_009724390.1   (Covid-19)
# 2: A0A140AYZ5_9BETC (MERS)
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1440
# Identity:     433/1440 (30.1%)
# Similarity:   654/1440 (45.4%)
# Gaps:         276/1440 (19.2%)
# Score: 1565.5
#=======================================
YP_009724390.      3 VFLVLLPL-------------VSSQCVNLTTRTQL-----PPAYTNSFTR     34
                       |||::..|             |.|.|:.:..:...     |.....|...
A0A140AYZ5_9B      5 VFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFFDKTWPRPIDVSKAD     54

YP_009724390.     35 GVYYPDKVFRSSVLHSTQDLFLPF---FSNVTWFHAIHVSGTNGTKRF--     79
                       |:.||.....|::..:.|.|| |:   ..:::..:.|.|.:||...|.|
A0A140AYZ5_9B     55 GIIYPQGRTYSNITITYQGLF-PYQGDHGDMYVYSAGHATGTTPQKLFVA    103

YP_009724390.     80 --DNPVLPF-------------NDGVYFASTEKSNIIR--------GWIF    106
                       ...|..|             :.|....|...|..||        |...
A0A140AYZ5_9B    104 NYSQDVKQFANGFVVRIGAAANSTGTVIISPSTSATIRKIYPAFMLGSSV    153

YP_009724390.    107 GTTLDSK-----TQSLLIVNNATNVVIKVCEFQFCNDPFLGVY-------    144
                       |...|.|     ..:|:::.:....:::.  |....:|..|.:
A0A140AYZ5_9B    154 GNFSDGKMGRFFNHTLVLLPDGCGTLLRA--FYCILEPRSGNHCPAGNSY    201

YP_009724390.    145 -----YH---------KNNKSWMESEFRVYSSANNCTFEY---VSQPFLM    177
                           ||         ..|::...:.|:.|.:..||||.|   :::..::
A0A140AYZ5_9B    202 TSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEIL    251

YP_009724390.    178 DLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLV    227
                       :..|.....:.:..|..:.:|.|
A0A140AYZ5_9B    252 EWFGITQTAQGVHLFSSRYVDLY-------------------------    274


YP_009724390.    228 DLPIGINITRFQTL---------LALHRSYLTPGDSSSGWTAGAAAYYVG    268
                           |.|:.:|.||         ..:..|..:.......|    ||:||.
A0A140AYZ5_9B    275 ----GGNMFQFATLPVYDTIKYYSIIPHSIRSIQSDRKAW----AAFYVY    316

YP_009724390.    269 YLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF    318
                       .|||.||||.::.:|.|..|:||..:.||:..|:.:||.||.|:|..|:|
A0A140AYZ5_9B    317 KLQPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSF    366

YP_009724390.    319 RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL    368
                       ..:|:.|:|..... |.|..:.:.|. ..||.:.|...:||..:.:.|
A0A140AYZ5_9B    367 EAKPSGSVVEQAEGVE-CDFSPLLSGTP-PQVYNFKRLVFTNCNYNLTKL    414

YP_009724390.    369 YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKI    418
                       .:..|.:.|.|..:||..:...|::::..|.|.........::....|.|
A0A140AYZ5_9B    415 LSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPLSMKSDLSVSSAGPI    464

YP_009724390.    419 ADYNYKLPDDFTGC-VIAWNSNNLDSKVGG-NYNYLYRLFRKSNLKPFER    466
                       :.:|||.......| ::|...:||.:.... .|:|:.:..|..:
A0A140AYZ5_9B    465 SQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRLLS------    508
```

```
YP_009724390.    467 DISTEIYQAGS----TPCNGV------EGFNCY----FPLQSYGFQPTNG    502
                     |..||:.|..:    :||...      |..:.|    .||:..|:...:|
A0A140AYZ5_9B    509 DDRTEVPQLVNANQYSPCVSTVPSTVWEDGDYYRKQLSPLEGGGWLVASG    558

YP_009724390.    503 VGYQPYRVVVLSFELLHAPAT----VCGPKKSTNLVK-----NKCVNFNF    543
                     ........:.:.|.:.....|    ||...:..|..|     ..||.::.
A0A140AYZ5_9B    559 STVAMTEQLQMGFGITVQYGTDTNSVCPKLEFANDTKIASQLGNCVEYSL    608

YP_009724390.    544 NGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFG    593
                     .|::|.||........:..|:|      ..||.:
A0A140AYZ5_9B    609 YGVSGRGVFQNCTAVGVRQQRF------VYDAYQ----------------    636

YP_009724390.   1175 SVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL   1224
                     :::::..|:..|.:|.|.||||.|||:|||.|..|.|||||||||||||
A0A140AYZ5_9B   1258 TLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNKWPWYIWLGFIAGL   1307

YP_009724390.   1225 IAIVMVTIMLCCMTSCCSCLKGCCSCGSCC-KFDEDDSEP   1263
                     :|:.:...:.|.|.|.:...|...|..|| ::::|.|.||
A0A140AYZ5_9B   1308 VALALCVFFILCCTGCGTNCMGKLKCNRCCDRYEEYDLEP   1347
```

Fig. 2. Alignment Result between SARS CoV-2 with MERS CoV Spike Glycoproteins.

Alignment result between SARS with MERS spike glycoproteins

```
########################################
# Program: water
# Aligned_sequences: 2
# 1: SPIKE_CVHSA      (SARS)
# 2: A0A140AYZ5_9BETC (MERS)
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
# Length: 1400
# Identity:     443/1400 (31.6%)
# Similarity:   662/1400 (47.3%)
# Gaps:         214/1400 (15.3%)
# Score: 1561.0
#=======================================
SPIKE_CVHSA        3 IFLLFLTLTS-------GSD--------LDRCTTFDDVQAPNYTQHTSSM     37
                     :|||...||.       |.|        :|...||.|...|.... .|..
A0A140AYZ5_9B      5 VFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFFDKTWPRPID-VSKA     53

SPIKE_CVHSA       38 RGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGI     87
                     .|:.||.....|:.....|.|| |:..:....:..:........|.|
A0A140AYZ5_9B     54 DGIIYPQGRTYSNITITYQGLF-PYQGDHGDMYVYSAGHATGTTPQK---     99

SPIKE_CVHSA       88 YFAATEKSNV---VRGWV--FGSTMNNKSQSVIIINNSTNVVIRACNFEL    132
                     .|.|....:|   ..|:|  .|:..|  |...:||:.||:..||..
A0A140AYZ5_9B    100 LFVANYSQDVKQFANGFVVRIGAAAN--STGTVIISPSTSATIRKI----    143

SPIKE_CVHSA      133 CDNPFFAVSKPMGTQT---------HTMIF--D------NAFNCTFE---    162
                       .|.|.:...:|..:         ||::.  |      .||.|..|
A0A140AYZ5_9B    144 --YPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILEPRS    191

SPIKE_CVHSA      163 ---------YISDAF----SLDVSEKSGNFKH------LREFVFKNKDGF    193
                             |.|.|.    :.|.|:  ||:..      .:|:.......|
A0A140AYZ5_9B    192 GNHCPAGNSYTSFATYHTPATDCSD--GNYNRNASLNSFKEYFNLRNCTF    239

SPIKE_CVHSA      194 LYVY--------------KGYQPIDVVR----DLPSG----------FNT    215
                     :|.|              :..|.:.:..    ||..|          ::|
```

```
A0A140AYZ5_9B    240 MYTYNITEDEILEWFGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDT    289

SPIKE_CVHSA      216 LKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLK    265
                     :|....:|..|.         :....:..|    ||::|..|:|.||:|.
A0A140AYZ5_9B    290 IKYYSIIPHSIR---------SIQSDRKAW----AAFYVYKLQPLTFLLD    326

SPIKE_CVHSA      266 YDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVR    315
                     :..:|.|..|:||..|.|::|.||.:||::::.|:|..|:|...|||.||.
A0A140AYZ5_9B    327 FSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSFEAKPSGSVVE    376

SPIKE_CVHSA      316 FPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFST--    363
                     ...... |.|..:.:.|. |.||.::|...:||  :|::....:.||.
A0A140AYZ5_9B    377 QAEGVE-CDFSPLLSGTP-PQVYNFKRLVFTNC--NYNLTKLLSLFSVND    422

SPIKE_CVHSA      364 FKCYGVSATKLNDLCFSNVYADSF----VVKGDDVRQIAPGQTGVIADYN    409
                     |.|..:|...:...|:|::..|.|    .:|.|    ::....|.|:.:|
A0A140AYZ5_9B    423 FTCSQISPAAIASNCYSSLILDYFSYPLSMKSD----LSVSSAGPISQFN    468

SPIKE_CVHSA      410 YKLPDDFMGC-VLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNV    458
                     ||.......| :||....|:   :|.....||.|:.... |....|.:.|
A0A140AYZ5_9B    469 YKQSFSNPTCLILATVPHNL---TTITKPLKYSYINKCS-RLLSDDRTEV    514

SPIKE_CVHSA      459 PFSPDG---KPCTPPALNCYWPLNDY-----------GFYTTTGIGYQPY    494
                     |...:.   .||.....:..|...||           |:...:|......
A0A140AYZ5_9B    515 PQLVNANQYSPCVSTVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMT    564

SPIKE_CVHSA      495 RVVVLSFELLNAPAT----VCGPKL---STDLIKNQ---CVNFNFNGLTG    534
                     ..:.:.|.:.....|    || |||   :...|.:|   ||.::..|::|
A0A140AYZ5_9B    565 EQLQMGFGITVQYGTDTNSVC-PKLEFANDTKIASQLGNCVEYSLYGVSG    613

SPIKE_CVHSA      535 TGVL---TPSSKRFQPF--QQFGRDVSDFTDSVRDPKTSEILDISPCSFG    579
                     .||.   |....|.|.|  ..:...|..::|.      .....:..|...
A0A140AYZ5_9B    614 RGVFQNCTAVGVRQQRFVYDAYQNLVGYYSDD------GNYYCLRACVSV    657

SPIKE_CVHSA      580 GVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWR-IYSTGNNV    628
                     .||||.  ...:...|.|:..|.|..:|:.:  .|.:.:.| :....::.
A0A140AYZ5_9B    658 PVSVIY--DKETKTHATLFGSVACEHISSTM--SQYSRSTRSMLKRRDST    703

SPIKE_CVHSA      629 F---QTQAGCLIGAEHVDTSY---ECDIPIGAGICASYHTVSLLRSTSQK    672
                     :   ||..||::|.  |::|.   :|.:|:|..:||...|.|.|...|.:
A0A140AYZ5_9B    704 YGPLQTPVGCVLGL--VNSSLFVEDCKLPLGQSLCALPDTPSTLTPRSVR    751

SPIKE_CVHSA      673 SI--------VAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMA    714
                     |:        :|:...:..| .:..|...::||||||..:|.|.:.:.::.
A0A140AYZ5_9B    752 SVPGEMRLASIAFNHPIQVD-QLNSSYFKLSIPTNFSFGVTQEYIQTTIQ    800

SPIKE_CVHSA      715 KTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVF    764
                     |.:|||..|:|....:|..||.:||.||:::|:||.|....||.:.|.:|
A0A140AYZ5_9B    801 KVTVDCKQYVCNGFQKCEQLLREYGQFCSKINQALHGANLRQDDSVRNLF    850

SPIKE_CVHSA      765 AQVKQMYKTPTLKYFGG-FNFSQILPDPLKP---TKRSFIEDLLFNKVTL    810
                     |.||....:|.:..||| ||.:.:.|..:..   :.||.||||||:|||:
A0A140AYZ5_9B    851 ASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSARSAIEDLLFDKVTI    900

SPIKE_CVHSA      811 ADAGFMKQYGECL--GDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAA    858
                     ||.|:|:.|.:|:  |..:||||||||...|..|||||:..:|.|||||::
A0A140AYZ5_9B    901 ADPGYMQGYDDCMQQGPASARDLICAQYVAGYKVLPPLMDVNMEAAYTSS    950

SPIKE_CVHSA      859 LVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQ    908
                     |:...|..|||.|..:...||||..:.||.||:|:||.||.||||.|||:
A0A140AYZ5_9B    951 LLGSIAGVGWTAGLSSFAAIPFAQSIFYRLNGVGITQQVLSENQKLIANK   1000
```

```
SPIKE_CVHSA       909 FNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSV    958
                      ||:|:..:|...|||:.|..|:||.||.|||||:.|..:||:.|||||:.
A0A140AYZ5_9B    1001 FNQALGAMQTGFTTTNEAFRKVQDAVNNNAQALSKLASELSNTFGAISAS   1050

SPIKE_CVHSA       959 LNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAA   1008
                      :.||:.|||.:|.:.||||||.|||.:|..:|.|||:|:.....||.||.
A0A140AYZ5_9B    1051 IGDIIQRLDVLEQDAQIDRLINGRLTTLNAFVAQQLVRSESAALSAQLAK   1100

SPIKE_CVHSA      1009 TKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFT   1058
                      .|::|||..||||..|||:|.|::||...||:|:.|:||.|.||......
A0A140AYZ5_9B    1101 DKVNECVKAQSKRSGFCGQGTHIVSFVVNAPNGLYFMHVGYYPSNHIEVV   1150

SPIKE_CVHSA      1059 TAPAIC---HEGKAYFPREGVFV-FNGT----SWFITQRNFFSPQIITTD   1100
                      :|..:|   :.....|..|.|: .|.|    .|..|..:|::|:.||:.
A0A140AYZ5_9B    1151 SAYGLCDAANPTNCIAPVNGYFIKTNNTRIVDEWSYTGSSFYAPEPITSL   1200

SPIKE_CVHSA      1101 NT-FVSGNCDVVIGIINNTVYDPL---QPELDSFKEELDKYFKNHTSPDV   1146
                      || :|:.  .|....|:..:..||   ...:| |::|||::|||.::...
A0A140AYZ5_9B    1201 NTKYVAP--QVTYQNISTNLPPPLLGNSTGID-FQDELDEFFKNVSTSIP   1247

SPIKE_CVHSA      1147 DLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPW   1196
                      :.|.::.||.::::::..|:..|.:|.|.||||.|||:|||.|..|.||||
A0A140AYZ5_9B    1248 NFGSLTQINTTLLDLTYEMLSLQQVVKALNESYIDLKELGNYTYYNKWPW   1297

SPIKE_CVHSA      1197 YVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCC-KFDEDDSEP   1245
                      |:|||||||:|:.:....:.|.|.|.:...|...|..|| :::|.|.||
A0A140AYZ5_9B    1298 YIWLGFIAGLVALALCVFFILCCTGCGTNCMGKLKCNRCCDRYEEYDLEP   1347
```

Fig. 3. Alignment result between MERS CoV with SARS CoV spike glycoproteins.

Table 1 represents the results of the alignment shown by the value of identity and similarity. The Identity is the percentage of identical matches between the two sequences over the reported aligned region (including any gaps in the length). The following are the results of alignment among the three viruses.

Table 1. Conclusions on the alignment results of the coronavirus spike glycoprotein.

No	Aligned Type	Identity (%)	Similarity (%)
1	Covid-19 x SARS	76.4	87.0
2	Covid-19 x MERS	30.1	45.4
3	SARS x MERS	31.6	47.3

The similarity is the percentage of matches between the two sequences over the reported aligned region (including any gaps in the length) (Taupiqurrohman *et al.*, 2016). In addition, Identity value indicates the identical equation of the compared amino acids, while the similarity value indicates the conformity on chemical properties (Hui *et al*., 2020). Table 1 also shows that Coronaviruses of SARS and COVID-19 have a high similarity with an identity value of 76.4% and 87%. On the contrary, the comparison of MERS and COVID-19 is relatively not similar because the alignment results are below 70%. The low result is also shown by the comparison between SARS with MERS having 31.6% identity, and 47.3% similarity. This is consistent with the explanation of Andriani (2016) and Li *et al*. (2020), where it was stated that the coronaviruses of SARS and COVID-19 are very close based on evolution tree. According to Rice *et al.* (2000), phylogenetic results (evolutionary kinship) cannot be concluded because of the type of protein being compared. Therefore, this research has illustrated the great potential of spike glycoprotein to be the source of peptide vaccine candidates for the SARS and COVID-19 diseases. Below is the structure of the spike glycoprotein of SARS-CoV, SARS-CoV-2, and MERS based on the database (pdb.org).

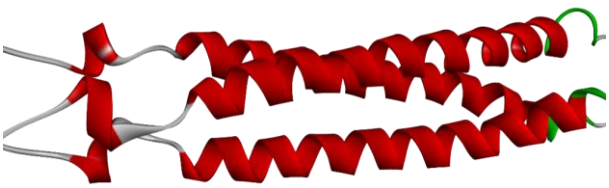

Fig. 4. Structure of SARS-CoV spike glycoproteins.

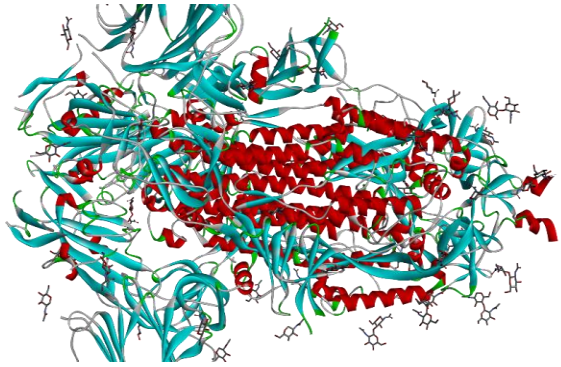

Fig. 5. Structure of SARS CoV-2 spike glycoproteins.

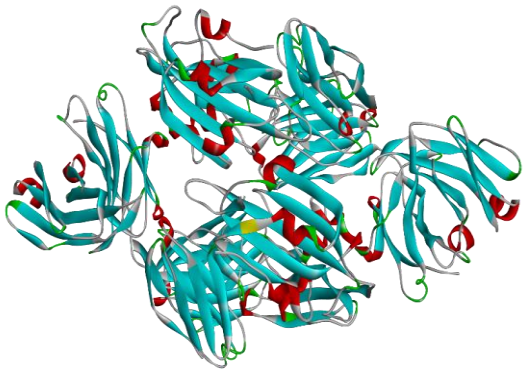

Fig. 6. Structure of SARS CoV spike glycoproteins

Work Principles of Universal Peptide Vaccine. The sequential analysis shows that spike glycoprotein can only be used as the source of peptide vaccine candidates for SARS and COVID-19. This should be properly conducted since the working principle of peptide vaccine is based on the immune system. The two common parts when a virus infects are the outside (specific body tissue) and the inside of an infected cell (body cell). When a part of the tissue is infected, the immune cells in the region begin to respond (Mothes *et al*., 2010; Mallapaty, 2020). This is evident in macrophages, which is one type of immune cell that is responsible for initiating the formation of antibodies through the activation of helper T cells. To activate this cells, macrophages will phagocytize the incoming antigen protein. Furthermore, the results of phagocytosis (small peptides) are raised to the surface of the body by major histocompatibility (MHC) class II protein to be recognized by helper T cell receptors (Li *et al*., 2020). Andriani (2016) stated the predicted part and made into a peptide vaccine.

During an internal cellular infection, the cell responds through a series of reactions (Fig. 4). An important part of this response in relation to the peptide vaccine is that the cell will attempt to bring the virus part to the surface. This is conducted by the MHC class I and recognized by cytotoxic T cells, which functions to reduce infection (Li *et al*., 2020). The part of the virus raised by MHC I and II is another peptide vaccine candidate that is predicted by using the spike glycoprotein (marked in the box in the picture). This protein is a potential candidate for peptide vaccine since it is found on the outer part of the virus spike glycoprotein is also in the outer part of the virus, thus it is a potential candidate for peptide vaccine. Initially, it is recognized or attached to the cell surface, and the location is given below.

Every disease has a cure. If the right medicine is found for a disease, the disease will be cured with the permission of Allah *Azza wa Jalla* (Sahih Muslim No. 4084). Based on this hadith, we can learn that there is no disease on this earth that was created by Allah swt. without a cure. As at present, many kinds of research have been carried out by scientists to find the most appropriate vaccine candidates for use in the prevention of infectious diseases caused by the coronavirus. The success of finding a vaccine candidate with the highest level of effectiveness is also inseparable from the power of Allah Almighty, as His word in QS. Ash-Shu'ara verse 80 (Kementerian Agama RI, 2019). This verse explains that it is Allah swt. who heals a man when he is sick. Allah has the power to heal any disease that a person has. But man, through the use of the mind by studying science, must also find out how to obtain this healing. Through science, humans can find out

the types of amino acids from the glycoprotein spike of various types of coronaviruses that are most appropriate to be used in the production of universal peptide vaccines for various types of infectious diseases caused by various types of coronaviruses. The lesson that can be taken from this verse is that diseases experienced by humans are the result of human actions themselves, including infection with diseases caused by the coronavirus, one of which is the lack of a clean lifestyle. Through the efforts made by humans and by the will of Allah swt. diseases suffered by humans can be cured. Diseases that occur in humans can also be a reminder to always be grateful for the various blessings from Allah swt. One of which is the favor of healing from an illness.

CONCLUSION

SARS CoV-2 (COVID-19) and SARS are very similar with 87% similarity and 76.4% identity values. In contrast, covid-19 with MERS and SARS with MERS are very different because of their reduced similarity and identity values below 70%. Therefore, the spike glycoprotein can only be used as the peptide vaccine candidate for COVID-19 and SARS.

REFERENCES

Abroug F, Slim A, Ouanes-Besbes L, Kacem MAH, Dachraoui F, Ouanes I, Lu X, Tao Y, Paden C, Caidi H, Miao C, Al-Hajri MM, Zorraga M, Ghaouar W, BenSalah A, Gerber SI. 2014. Family cluster of Middle East respiratory syndrome coronavirus infections, Tunisia, 2013. *Emerging Infectious Diseases*. vol 20(9): 1527–1530. doi: https://dx.doi.org/10.3201%2Feid2009.140378.

Alouane T, Laamarti M, Essabbar A, Hakmi M, Bouricha EM, Chemao-Elfihri MW, Kartti S, Boumajdi N, Bendani H, Laamarti R, Ghrifi F, Allam L, Aanniz T, Ouadghiri M, Hafidi NE, Jaoudi RE, Benrahma H, Attar JE, Mentag R, Sbabou L, Nejjari C, Amzazi S, Belyamani L, Ibrahimi A. 2020. Genomic diversity and hotspot mutations in 30,983 SARS-CoV-2 genomes: moving toward a universal vaccine for the "confined virus"?. *Pathogens*. vol 9(10): 1–19. doi: https://doi.org/10.3390/pathogens9100829.

Andriani T. 2016. Application of Upgma method to identify kinship types of viruses and spread of Ebola epidemic through the formation of phylogenetic trees. [Thesis]. Surabaya: Magister Program in Mathematics Faculty of Mathematics and Natural Sciences Institute of Technology ten November.

Aprilyanto V, Sembiring L. 2017. Bionformatics. Yogyakarta: Innosains.

Cunha CB, Opal SM. 2014. Middle East respiratory syndrome (MERS) a new zoonotic viral pneumonia. *Virulence.* vol 5(6): 650–654. doi: https://doi.org/10.4161/viru.32077.

da Costa VG, Moreli ML, Saivish MV. 2020. The emergence of SARS, MERS and novel SARS-2 coronaviruses in the 21st century. *Archives of Virology*. vol 165(7): 1517–1526. doi: https://doi.org/10.1007/s00705-020-04628-0.

Hui DS, Azhar EI, Madani TA, Ntoumi F, Kock. 2020. The continuing 2019-nCoV Epidemic Threat of Novel Coronaviruses to Global Health — The Latest 2019 Novel Coronavirus Outbreak in Wuhan, China. *International Journal of Infectious Diseases*. vol 91: 264–266. doi: https://doi.org/10.1016/j.ijid.2020.01.009

ICTV. 2020. Naming the 2019 Coronavirus. London: International Committee on Taxonomy of Viruses. https://talk.ictvonline.org/.

Jeong SY, Sung SI, Sung JH, Ahn SY, Kang ES, Chang YS, Park WS, Kim JH. 2017. MERS-CoV infection in a pregnant woman in Korea. *Journal of Korean Medical Science*. vol 32(10): 1717–1720. doi: https://doi.org/10.3346/jkms.2017.32.10.1717.

Kementerian Agama RI. 2019. Terjemahan Al Qur'an Surat Ash-Shu'ara ayat 80. https://quran.kemenag.go.id/.

Khalaj-Hedayati A. 2020. Protective immunity against SARS subunit vaccine candidates based on spike protein: lessons for coronavirus vaccine development. *Journal of Immunology Research.* vol 2020: 1–11. doi: https://doi.org/10.1155/2020/7201752.

Li B, Si H-R, Zhu Y, Yang X-L, Anderson DE, Shi Z-L, Wang L-F, Zhou P. 2020. Discovery of bat coronaviruses through surveillance and probe capture-based next-generation sequencing. *Msphere*. vol 5(1): 1-11. doi: https://doi.org/10.1128/mSphere.00807-19.

Li G, Fan Y, Lai Y, Han, T, Wang HD. 2020. Coronavirus infections and immune responses. *Journal of Medical Virology***.** vol 92(4): 424–432. doi: https://doi.org/10.1002/jmv.25685.

Liu D X, Fung, T S, Kian K, Chong L, Shukla A, Hilgenfeld, R. 2014. Accessory proteins of SARS-CoV and other coronaviruses. *Antiviral Research*. vol 109: 97–109. doi: https://doi.org/10.1016/j.antiviral.2014.06.013.

Mailles A, Blanckaert K, Chaud P, van der Werf S, Lina B, Caro V, Campese C, Guéry B, Prouvost H, Lemaire X, Paty M C, Haeghebaert S, Antoine D, Ettahar N, Noel H, Behillil S, Hendricx S, Manuguerra J C, Enouf V, La Ruche G, Semaille C, Coignard B, Lévy-Bruhl D, Weber F, Saura C, Che D. 2013. The investigation team collective. first cases of middle east respiratory syndrome coronavirus (MERS-CoV) infections in France, investigations and implications for the prevention

of human-to-human transmission, France, May 2013. *Eurosurveillance*. vol 18(24): 1–5. doi: https://doi.org/10.2807/ese.18.24.20502-en.

Mallapaty S. 2020. Mini organs reveal how the coronavirus ravages the body. *Nature*. vol 583(7814): 15–16. doi: https://doi.org/10.1038/d41586-020-01864-x.

Mothes W, Sherer NM, Jin J, Zhong P. 2010. Virus cell-to-cell transmission. *Journal of Virology*. vol 84(17): 8360–8368. doi: https://doi.org/10.1128/JVI.00443-10.

Peiris JSM, Yuen KY, Osterhaus ADME, Stöhr K. 2003. The Severe Acute Respiratory Syndrome. *The New England Journal of Medicine*. vol 349(25): 2431–2441. doi: https://doi.org/10.1056/NEJMra032498.

Puzelli S, Azzi A, Santini MG, Di Martino A, Facchini M, Castrucci MR, Meola M, Arvia R, Corcioli F, Pierucci F, Baretti S, Bartoloni A, Bartolozzi D, de Martino M, Galli L, Pompa MG, Rezza G, Balocchini E , Donatelli I. 2013. Investigation of an imported case of middle east respiratory syndrome coronavirus (MERS-CoV) infection in Florence, Italy, May to June 2013. *Eurosurveillance*. vol 18(34): 1–4. doi: https://doi.org/10.2807/1560-7917.ES2013.18.34.20564.

Racelis S, de los Reyes VC, Sucaldito MN, Deveraturda I, Roca JB, Tayaga E. 2015. Contact tracing the first Middle East respiratory syndrome case in the Philippines. *Western Pacific Surveillance and Response Journal*. vol 6(3): 3–7. doi: https://dx.doi.org/10.5365%2FWPSAR.2015.6.2.012.

Rampengan NH. 2016. Middle East Respiratory Syndrome Coronavirus. *Jurnal Biomedik*. vol 8(1): 17–26. doi: https://doi.org/10.1177/2165079915607497.

Rashid H, Azeem MI, Heron L, Haworth E, Booy R, Memish ZA. 2013. Has Hajj-associated MERS-CoV transmission occurred? The case for effective post-Hajj surveillance for infection. *Clinical Microbiology Infection*. vol 20(4): 271–276. doi: https://doi.org/10.1111/1469-0691.12492.

Rice P. Longden I, Bleasby A. 2000. EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics*. vol 16(6): 276–277.

Sahih Muslim No. 4084/Sahih Muslim 2204; In-book reference: Book 39, Hadith 95; USC-MSA reference: Book 26, Hadith 5466.

Schoeman D, Fielding BC. 2019. Coronavirus envelope protein: current knowledge. *Virology Journal*. vol 16(1): 1–22. doi: https://doi.org/10.1186/s12985-019-1182-0.

Shapiro M, London B, Nigri D, Shoss A, Zilber E, Fogel I. 2016. Middle East respiratory syndrome coronavirus: review of the current situation in the world. *Disaster and Military Medicine.* vol 2(1): 1–5. doi: https://doi.org/10.1186/s40696-016-0019-2.

Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R. 2020. COVID-19 Infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*. vol 24: 91–98. doi: https://doi.org/10.1016/j.jare.2020.03.005.

Slamet, Bratasena, A., Sitorus, M., Rizkiyat, N., Samoedro, E., Wignjadiputro. 2013. General guidelines on readiness to face middle east respiratory syndrome-corona virus (MERS-CoV). Jakarta: *Directorate General of Environmental Control and Investigation of Ministry of Health Republic of Indonesia*.

Subroto T, Hardianto A, Abdul Alim Kahari, Tika Pradnjaparamita. 2013. Protective group peptide synthesis as a component precursor to Universal Influenza vaccines. Bandung: PTNBR – BATAN.

Taupiqurrohman O, Yusuf M, Nuswantara S, Subroto T. 2016. Potensi Gen Oncoprotein Human Papillomavirus Tipe 16 Sebagai Kandidat Vaksin Kanker Serviks. *Majalah Kedokteran Bandung*. vol 48(2): 84–91. doi: https://doi.org/10.15395/mkb.v48n2.761

Thomas HL, Zhao H, Green HK, Boddington NL, Carvalho CF, Osman HK, Sadler C, Zambon M, Bermingham A, Pebody RG. 2014. Enhanced MERS coronavirus surveillance of travelers from the Middle East to England. *Emerging Infectious Diseases*. vol 20(9): 1562–1564. doi: https://dx.doi.org/10.3201%2Feid2009.140817.

Wang N, Shi X, Jiang L, Zhang S, Wang D, Tong P. 2013. Structure of MERS-CoV Spike Receptor-Binding Domain Complexed with Human Receptor DPP4. *Cell Research*. vol 23(8): 986.

WHO. 2016. Middle East Respiratory Coronavirus (MERS-CoV). Geneva: World Organization Health. http://www.who.int/emergencies/mers-cov/en/.

Wu F. Zhao S, Yu B. 2020. A new coronavirus associated with human respiratory disease in China. *Nature*. vol 579: 265–269. doi: https://doi.org/10.1038/s41586-020-2008-3.

Zhong N, Zheng B, Li Y, Poon L, Xie Z, Chan K. 2003. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February. *The Lancet*: vol 362(9393): 1353–1358.

Zumla A, Hui DS, Perlman S. 2015. Middle East respiratory syndrome. *The Lancet*. vol 386(9997: 995–1007. doi: 10.1016/S0140-6736(15)60454-8.