

Conformational Epitope Prediction of Envelope Protein for Novel Wuhan Pneumonia Coronavirus (2019-nCoV)

Pritesh Nilendu and Sandeep D Bansode*

Dr D Y Patil Biotechnology and Bioinformatics institute Dr.D Y Patil Vidyapeeth, India

*Corresponding Author: Sandeep D Bansode, Dr D Y Patil Biotechnology and Bioinformatics institute Dr.D Y Patil Vidyapeeth, Tathawade, Pune, India.

Received: May 11, 2020; Published: June 30, 2020

Abstract

Previously, there were a number of Coronavirus Outbreaks that took place in 2003 and 2012 but were not as lethal as current 2019-nCoV. As the Envelope Protein Plays an Important role in forming the outer protective cover, So we took a closer look into the mutation that took place in 2019-nCoV that makes it different from SARS And MERS-CoV.

Keywords: 2019-nCoV; SARS-CoV; MERS-CoV; BLAST; CLUSTAL W; ELLIPRO

Abbreviations

2019-nCoV: 2019 Novel Coronavirus; SARS-CoV: Severe Acute Respiratory Syndrome Coronavirus; MERS-CoV: Middle East Respiratory Syndrome Coronavirus; BLAST: Basic Local Alignment Search Tool

Introduction

In December 2019, suddenly there was an increase in the number of pneumonia cases in china [1,2]. On investigation by D.R Lee, who was a doctor in a hospital of Wuhan, Hubei province of china found that the disease was caused by previously existing virus with a modified and mutated version called as 2019 novel coronavirus (2019-nCoV) [3,4].

Corona viruses are a large group of RNA Viruses in the nidovirales order and family coronaviridae and have one of the largest RNA Genome. This virus has +ve single stranded RNA. Corona prefix came for the Latin origin and means "crown like appearance". Club like or the Crown like appearance is caused due to spike protein (S) present in the lipid bilayer protruding outwards [5]. There are four classification of corona viruses:

- 1. Alpha coronaviruses
- 2. Beta coronaviruses
- 3. Gamma coronaviruses
- 4. Delta coronaviruses.

Alpha and beta coronaviruses are the coronaviruses that infect humans and get transmitted from bats and live animals.

Early cases of coronaviruses occurred in 2002 - 2003 and was called severe acute respiratory syndrome (SARS-CoV) and was in southern china, Guangdong province of china. SARS coronaviruses were transmitted from bats to civets and then to humans, with a mortality rate of 9.6%. SARS-CoV has an incubation period of 8 to 14 days [6].

Another outbreak was in 2012 which occurred in Saudi Arabia which was called Middle East respiratory syndrome (MERS-CoV). It was transmitted from camels to humans and has an incubation period of 3 days [7].

Conformational Epitope Prediction of Envelope Protein for Novel Wuhan Pneumonia Coronavirus (2019-nCoV)

179

Another outbreak took place in 2019, Wuhan, Hubei Province of china, and was identified as novel beta corona virus (2019-nCoV) which originated from live and Dead Sea animals market of Wuhan city [8,9].

Corona viruses has 4 structural proteins:

- 1. Spike protein(S).
- 2. Envelope protein (E).
- 3. Membrane protein (M)
- 4. Nucleocapsid protein (N).

Spike protein(S) is a trimeric protein which is a receptor binding protein and helps in fusion of the envelope of virus to the host cell. This spike protein is responsible for the crown like appearance like a king's crown. Envelope protein (E) is a pentameric protein that functions as an ion channel protein and allows flow of ions from host to virus and vice-versa. Membrane glycoprotein (M) is the protein that forms a lipid bilayer membrane for the corona virus. Nucleocapsid protein (N) is the RNA binding protein and helps in synthesis of more number of RNA and in the process of translation [10,11].

Immune system

A substance that is recognized by the immune system as foreign and invokes or introduces an immune response is called as antigen [12]. Antigens are usually proteins that are too large to bind as a whole to any receptor, so only specific site on the antigen bind with a specific antibody. These segments are called epitopes. Likewise, it is only paratope of the antibody that comes in contact with the epitope.

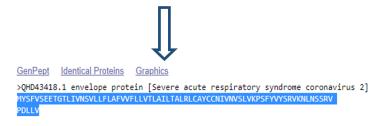
A linear or a sequential epitope is an epitope that is recognized by antibodies by its linear or straight sequence of amino acids. These have primary structure present in them. In contrast, most antibodies recognize a conformational epitope that has a specific three-dimensional shape as its protein structure.

Conformational epitope

A conformational epitope consists of amino acid residues that are discontinuous and not continuous which will be recognized by the antibody and which are present in an antigen that come in direct contact with a receptor of the immune system and binds to an antibody. These generally exist as tertiary or quaternary structure. There were some studies that led to the conclusion that envelope protein (E) of Coronavirus serves as an Epitope [13].

Materials and Methods

a. Initially 2019-nCoV corona virus was thoroughly studied and it's properties of the epitope of the envelope protein was studied.



b. Fasta sequence of the envelope (E) protein [Wuhan Seafood Market Pneumonia Virus] was studied and extracted.



c. Structure of the above protein was searched in Protein Data Bank.





d. Peptide Sequence was searched in Uniprot Yielded no results.

NIH U.S. National Library of Medicine	NCBI National Center for Biotechi	nology Information	
BLAST $^{\circ}$ » blastn suite			
		s	andard Nucleotide BLAST
blastn blastp blastx tblastn tblastx			
Enter Query Sequence		BLASTN programs sear	rch nucleotide databases using a nuc
Enter accession number(s), gi(s), or FAS	TA sequence(s) 😡	Clear Query subrange 😡	
MN988669.1 Severe acute respiratory syndro nCov WHU02, complete genome	me coronavirus 2 isolate 2019- 🔺	From]
TTAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTC GAACTTTAAAATCTGTGTGGGCTGTCACTCGGCTGCATGCTTAG			
AATTACTGTCGTTGACAGGACACGAGTAACTCGTCTATCTTT		То	
Or, upload file Choose file No file	chosen		
Job Title			
Enter a descriptive title	for your BLAST search 😡		
Align two or more sequences 😣			

Citation: Pritesh Nilendu and Sandeep D Bansode. "Conformational Epitope Prediction of Envelope Protein for Novel Wuhan Pneumonia Coronavirus (2019-nCoV)". *EC Clinical and Medical Case Reports* 3.7 (2020): 176-187.

180

Conformational Epitope Prediction of Envelope Protein for Novel Wuhan Pneumonia Coronavirus (2019-nCoV)

181

Blast was dor	ne for the 2019-nCoV.
	Clustal Omega
	Input form Web services Help & Documentation Bioinformatics Tools FAQ
	Tools > Multiple Sequence Alignment > Clustal Omega
	Multiple Sequence Alignment Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more sequences. For the alignment of two sequences please instead use our <u>pairwise sequence alignment tools</u> .
	Important note: This tool can align up to 4000 sequences or a maximum file size of 4 MB.
	STEP 1 - Enter your input sequences
	Enter or paste a set of
	PROTEIN
	sequences in any supported format
	OF, upload à file: Choose file No file chosen Use a example sequence Clear sequence See more example inputs

f. Multiple sequence alignment of the protein with existing SARS corona virus and Bat Coronavirus was done for identifying the degree of similarity.



g. 3d model of the envelope protein model was built using swiss model server.

SIB UDZENTRUM University of Basel The Center for Molecular Life Sciences	WISS-MODEL
Start a New Modelling	g Project o
Target Sequence(s): (Format must be FASTA, Clustal, plain string, or a valid UniProtKB AC)	Targer MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTAIRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV
	Ų

h. Model generated through Swiss Model was uploaded in Ellipro.

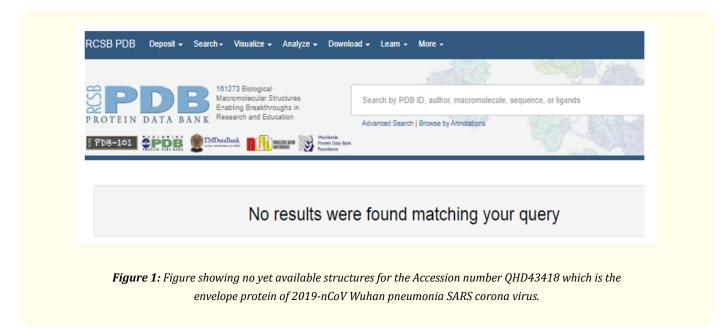
IEDB Analysis	Resource
Home Help Example Refe	erence Download Contact
ElliPro: Antibody	Epitope Prediction
Spec	tify Sequence(s)
Enter PDB ID(s) or upload PDB file	Choose file query.pdb
Select Epitop	e Prediction Parameters
Minimum score:	0.5 ▼ (Default is 0.5)
Maximum distance (Angstrom):	6 ▼ (Default is 6)
	Submit
	Ţ

i. Epitope Sequences for the Envelope protein were determined through Ellipro.



- j. Epitope sequence was searched in Scanprosite which yielded no hits.
- k. Amino acids for the epitope were analyzed by Discovery Studio.
- I. Secondary structure was analyzed through PROTEIN MODEL CHECK Database that yielded composition of helix and turns.

Results and Discussion



Citation: Pritesh Nilendu and Sandeep D Bansode. "Conformational Epitope Prediction of Envelope Protein for Novel Wuhan Pneumonia Coronavirus (2019-nCoV)". *EC Clinical and Medical Case Reports* 3.7 (2020): 176-187.

182

183

CLUSTAL O(1.2.4) multiple sequence alignment

QHD43418.1 ACZ72287.1 ABD75324.1	MYSFVSEETGTLIVNSVLLF MYSFVSEETGTLIVNSVLLF	LAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYS LAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPTVYVYS FAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPTVYVYS :************************************	60 60 60
QHD43418.1 ACZ72287.1 ABD75324.1	RVKNLNSSR-VPDLLV RVKNLNSSEGVPDLLV RVKNLNSSEGVPDLLV ********. ******	75 76 76	

Figure 2: Figure showing the multiple sequence alignment of envelope protein of Wuhan pneumonia corona virus (2019-nCoV), existing SARS corona virus and bat SARS corona virus.

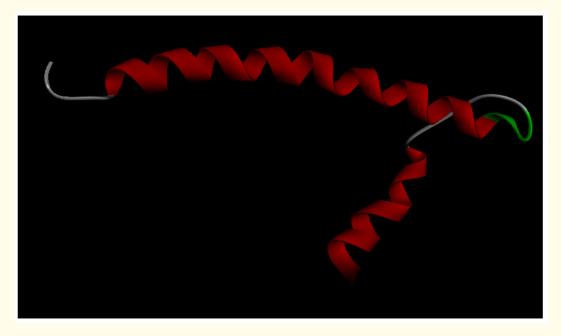


Figure 3: Figure showing the 3d structure of envelope protein generated using Swiss model database by choosing appropriate templates which was viewed in Discovery Studio.

184

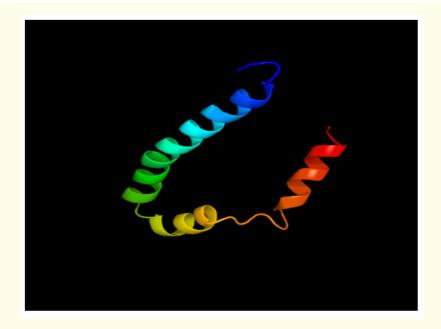


Figure 4: Figure showing the predicted secondary structure of the envelope protein using phyre2. In a colour picture, the residues that are part of a helix are shown in blue, strand residues in red. Preferred regions for helical residues are drawn in blue, for strand residues in red, and for all other residues in green.

Sec	condarv	structure	assion	nent				
			10	20	30	40	50	60
			1	I	I		1	1
	1 -	60 MYSFVSE	ETGTLIV	NSVLLFLAFV	VFLLVTLAIL	TALRLCA <mark>YCCI</mark>	NIVNVS <mark>LVKP</mark>	SFY <mark>VYS</mark>
(1)-(60) TT	TTHHHH	нннннннтн	нннт ттннн	іннннннн <mark>ннт:</mark>	<mark>г Т</mark> ТТТ	Т
			70					
	61 -	75 <mark>RVKNLNS</mark>	SRVPDLI	<mark>.v</mark>				
(61)-(75) ННННТТ						

Figure 5: Figure showing the quality of the model generated by swiss-model server, Only helix (H), overwound or 3/10-helix (3), strand (S), turn (T) and coil (blank) are shown. This data is generated by PROTEIN MODEL CHECK DATABASE.

Predicted linear epitope(s)

No.	Chain	Start	End	Peptide	Number of residues	Score
1	-	23	35	FVVFLLVTLAILT	13	0.664
2	_	5	10	VSEETG	6	0.66
3	_	42	49	YCCNIVNV	8	0.628
4	_	58	73	VYSRVKNLNSSRVPDL	16	0.608

Table: Showing the predicted peptide sequences of the epitope of 2019-nCoV.

ScanProsite Results Viewer
Ouput format: Graphical view - this view shows ScanProsite results together with ProRule-based predicted intra-domain features [hei
Hits for USERPAT1{F-V-V-F-L-L-V-T-L-A-I-L-T} motif on all PDB database sequences :
found: 6 hits in 6 sequences
Legend:
• • •
disulfide bridge active site other 'ranges' other sites
Please note that the graphical representations of domains displayed hereafter are for illustrative purposes only, and that their colors a For more information about how these graphical representations are constructed, go to https://prosite.expasy.org/mydomains/. hits by patterns: [6 hits (by 1 pattern) on 6 sequences]
ruler: 1 100 200 300 400 500 600 700 800 900 1000
2MM4-A (58 aa) View all PROSITE motifs hits on sequence
USERPAT1 : Hits of F-V-V-F-L-L-V-T-L-A-I-L-T on 2MM4-A 3D structure Pattern: F - V - V - F - L - L - V - T - L - A - I - L - T Approximate number of expected random matches [Ref: PMID 11535175] in ∼ 100'000 sequences (50'000'000 residues): 2.549912€
16 - 28: FVVFLLVTLAILT

Figure 6: Figure showing the prosite search result of epitope peptide sequence "FVVFLLVTLAILT".

Hits for USER	RPAT1{V-S-E-E-T-G} motif on all PDB database sequences :
found: 19 hits in 19	sequences
Legend:	
	• • •
disulfide bridge	active site other 'ranges' other sites
hits by patter	1 100 200 300 400 500 600 700 800 900 1000
ruler:	
2FB5-A	(204 aa) View all PROSITE motifs hits on sequence
Pattern: V - S - E	its of V-S-E-E-T-G on 2FB5-A 3D structure -E - T- G bør of expected random matches [Ref: PMID 11535175] in ∼ 100'000 sequences (50'000'000 residues): 3.1
Approximate num	

Figure 7: Figure showing the prosite motif search result of epitope peptide sequence motif "VSEETG".

Citation: Pritesh Nilendu and Sandeep D Bansode. "Conformational Epitope Prediction of Envelope Protein for Novel Wuhan Pneumonia Coronavirus (2019-nCoV)". *EC Clinical and Medical Case Reports* 3.7 (2020): 176-187.

185

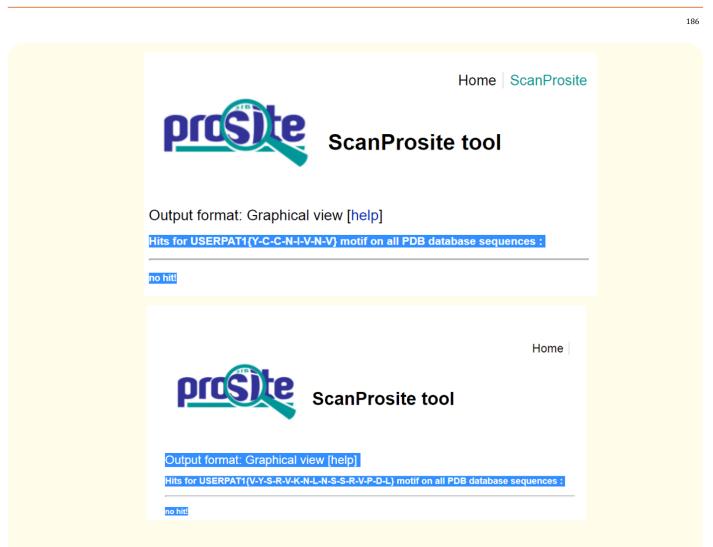


Figure 8: Figure showing the prosite search result of epitope peptide sequence motif "YCCNIVNV" and "VYSRVKNLNSSRVPDL" which shows no hits indicating the presence of a unique sequence motif in the proteins compared in the database.

Conclusion

Two unique conformational peptides were identified which were not there in the previous identified coronaviruses. The sequence is as follows:

- 1. YCCNIVNV
- 2. VYSRVKNLNSSRVPDL.

These two epitopes were predicted to be unique conformational epitope in Wuhan Pneumonia Coronavirus (2019-nCoV) also called SARS CoV-2.

Conflict of Interest

The authors have no conflict of interest.

Bibliography

- Kwok KO., et al. "Novel coronavirus (2019-nCoV) cases in Hong Kong and implications for further spread". Journal of Infection 80.6 (2020): 671-693.
- 2. Li JY., *et al.* "The epidemic of 2019-novel-coronavirus (2019-nCoV) pneumonia and insights for emerging infectious diseases in the future". *Microbes and Infection* 22.2 (2020): 80-85.
- 3. She J., *et al.* "2019 novel coronavirus of pneumonia in Wuhan, China: emerging attack and management strategies". *Clinical and Translational Medicine* 9.1 (2020): 19.
- Sohrabi C., et al. "World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)". International Journal of Surgery 76 (2020): 71-76.
- 5. Chen Y., *et al.* "Emerging coronaviruses: Genome structure, replication, and pathogenesis". *Journal of Medical Virology* 92.4 (2020): 418-423.
- 6. Cherry JD. "The chronology of the 2002-2003 SARS mini pandemic". Paediatric Respiratory Reviews 5.4 (2004): 262-269.
- 7. Khuri-Bulos N., *et al.* "Middle East respiratory syndrome coronavirus not detected in children hospitalized with acute respiratory illness in Amman, Jordan, March 2010 to September 2012". *Clinical Microbiology and Infection* 20.7 (2014): 678-682.
- 8. Carlos WG., et al. "Novel Wuhan (2019-nCoV) Coronavirus". American Journal of Respiratory and Critical Care Medicine 201.4 (2020): P7-P8.
- 9. Ralph R., *et al.* "2019-nCoV (Wuhan virus), a novel Coronavirus: human-to-human transmission, travel-related cases, and vaccine readiness". *Journal of Infection in Developing Countries* 14.1 (2020): 3-17.
- 10. Spaan W., et al. "Coronaviruses: structure and genome expression". Journal of General Virology 69.12 (1988): 2939-2952.
- 11. Liu DX., et al. "Accessory proteins of SARS-CoV and other coronaviruses". Antiviral Research 109 (2014): 97-109.
- 12. Nossal GJ., et al. "Antigens in immunity. VII. Analysis of immunological memory". Immunology 9.4 (1965): 333-348.
- 13. Hui Liu., *et al.* "Recombinant scFv Antibodies against E Protein and N Protein of Severe Acute Respiratory Syndrome Virus". *Acta Biochimica et Biophysica Sinica* 36.8 (2004): 541-547.

Volume 3 Issue 7 July 2020 © All rights reserved by Pritesh Nilendu and Sandeep D Bansode.