



Proteome Analysis and Antigenic Peptide Prediction of Beta Corona Virus, A Cause of MERS

*¹Syeda Sograh Fatima, ²Desu Jyothisna

¹Princess Nourah Bint Abdul Rahman University -Riyadh, Saudi Arabia, ²BioAxis DNA Research Centre Pvt Ltd, Hyderabad

* Email Id: syedasougra@yahoo.in

Received: 10th August 2014, Accepted: 16th August 2014, Published: 1st September 2014

Abstract:

Middle East respiratory syndrome coronavirus is the a genus belonging to beta corona Virus, currently well known for its out break in the countries like Saudi Arabia, Malaysia, Jordan, Qatar, Egypt, the UAE etc causing severe illness. MERS CoV is a dreadful RNA virus infecting huge population especially from Saudi Arabia. There is not been any protection against this well known virus in spite of its outbreak and rapid spread. The current work is an approach to design a vaccine that can protect the people by making them immune to the virus. The important proteins of the beta corona virus were retrieved from NCBI public data base and the screening was performed to identify the Foreign proteins among the complete set of viral proteins. The methodology is based on the principle that the diseases are caused due to the viral proteins that are totally foreign to the human proteome so that the synthesis of antibodies does not occur. The tools like BLAST of NCBI and TFASTY of SDSC Biology Workbench have been used for this screening. Once the foreign proteins were collected they were further analyzed for the presence of the antigenic peptide on their sequences. The tool used for the purpose was Antigenic Emboss. The result contains several peptides for a single protein. All the set of peptides collected were further screened to identify the peptides having highest antigenic propensity using the PVS server. Later the Surface accessibility Area of the selected peptides was calculated to eliminate the peptides having lowest SAA so as to increase the accuracy and effectively of the vaccine thus generated. The final list of selected peptides was designed in Argus lab and their internal energy values were calculated. The peptides having lowest energy can be selected as the targets for the development of protein based vaccines. The work can further be extended to synthetically produce the antigenic peptides and attenuate them so as to check for its capability of immunization on lab animals. Once the result shows positive in the lab animals its can be processed for the production of large scale vaccines for MERS so as to eradicate the disease.

Keywords:

MERS, Beta Corona Virus, RNA based Virus, TFASTY, SDSC Biology work bench, NCBI, PVS.

Introduction:

The virus MERS-CoV is a new member of the beta group of coronavirus, Betacoronavirus, lineage C. It causes a dreadfull breathing disorder MERS whose onset was first reported in 2012 at Saudi Arabia. The Organism as isolated and identified for the first time by an Egyptian Virologist Dr. Ali Mohamed Zaki isolated from patients lungs [1]. This virus has spread from ill people to others through close contact, such as caring for or living with an infected person [2]. In addition to humans, MERS-CoV has been found in camels in Qatar, Oman, Egypt and Saudi Arabia, and a bat in Saudi Arabia. There is growing evidence that camels play a significant role in the virus transmission [3]. The symptoms of the disease include Fever, cold, cough, shortness of breath this may finally become severe and deadly [4]. There has not yet been any vaccine or medication designed to treat the symptoms or stop the onset of the disease. The biggest precaution one can take from contracting MERS is to maintain healthy habits that can prevent the individual from coming down with any sort of respiratory illness, such as: thoroughly washing hands with soap and water, covering mouth and nose when sneezing or coughing, refraining from touching the face with unwashed hands, avoiding close contact (no kissing or sharing utensils) with sick people, and keeping one's environment clean and disinfected.

Vaccine designing is one of the most advanced methodologies in the area of medication. Other than the regular vaccine development in case of Reverse vaccinology approach the vaccine molecules can be predicted without the use of human volunteers. Here in this method the screening starts from the proteome of the bacteria or any other pathogen that causes the disease. The identification of the foreign proteins is performed followed by the prediction of antigenic regions. Final step is to design the structure and calculate the energy values. Once the best peptide is proposed that can be further synthesized in the laboratory and used for vaccination.

Materials and Methods:

Collection of Viral Proteins from Public Databases:

The prerequisite for the Vaccine development is the proteome library of the target organism. Here the complete set of MERS CoV proteins is collected from NCBI public data base. All the proteins were collected in their FASTA format and were used for further analysis. One can also use other data bases like uniprot, swiss prot, JCVI CMR etc for the collection of the proteomes.

Primary Screening For the identification of Foreign Proteins:

The proteins that are immunogenic are the one that are unknown or foreign to the human proteome. To screen for the foreignness by comparison with the human proteome BLAST P tool has been used that is based on pair wise sequence alignment and comparison. All those viral proteins that show less than 30% identity or no identity with the human proteome can be selected in the screening step.

Secondary Screening using TFASTY:

The screening as performed by BLASTP may show some errors. To confirm the accuracy of the results of BLASTP, TFASTY can be used that would further screen the viral proteins against the human CDS to check for the similarity. Only those proteins that are totally foreign to the human proteome are selected as the final list of Foreign Proteins. TFASTY is a facility provided by SDSC Biology Workbench.

Antigenic Peptide Prediction:

From the selected list of foreign protein only few may be pathogenic and cause the disease. There are certain regions within the protein sequence called the Antigenic sites that are the actual regions of pathogenicity that can provoke the immune response in a host. These regions are identified for all the selected protein using Antigenic Emboss. Each protein may have more than one Antigenic site within its sequence. Total number of antigenic sites present in the selected sequences must be identified and collected in this step.

Identification of the peptides having highest antigenic Propensity:

Though all the above selected peptides are antigenic and can provoke an immune response, their antigenicity is measured in terms of antigenic propensity. The peptide having high antigenic propensity can be a good pathogenic peptide.

The competition of their antigenic propensity can be done using a tool “Protein Variability Server”, that would provide a graphical representation for each protein and its antigenic sites along with their propensity. From the results the peptide having maximum propensity can be considered.

Calculation of Surface Accessibility Area of the selected Peptides:

Surface Accessibility Area is the total area on the protein structure that is easily available for the external ligand to bind to it. It is that area that is exposed and has a cavity or the binding space. The area depends on the concentration of the Valine, Cystine and Leucine amino acids in the peptides. The formula for the calculation of the SAA is as follows:

$$\text{SAA} = \frac{\text{Total no of V} + \text{L} + \text{C}}{\text{Length of the peptide}}$$

The SAA for all the selected foreign Pathogenic peptides has been calculated and the one having maximum SAA are selected for the further step.

Construction of the Peptides using Argus lab:

Argus lab is a downloadable software package that is used for model building of proteins, chemicals, peptides etc. It can also be used for the docking studies and energy evaluation of the proteins. Here Argus lab has been used to design the 3D Structures of the selected peptides. The structures thus designed were used for the development of the Vaccine.

Results and Discussion:

Collection of the Viral Proteins from the Data bases:

A total of 70 known viral proteins were collected from NCBI database and were screened to eliminate the proteins that are Putative, Hypothetical, predicted etc. After screening and eliminating the above protein 54 protein sequences were retained that were processed for the further analysis.

Screening based on BLAST P^[5] and TFASTY:

Screening was done for the elimination of the sequence that share 30 % or more similarity to the human proteome. The proteins that are either totally new or share less than 30% similarity to the human proteome are selected and are termed as Foreign Proteins. As per the principles some of these proteins of the MERS CoV are capable of causing a disease as they can provoke an immune response. A total of 45 proteins are selected as foreign.

Figure 1: The BLAST page showing no similarity of the viral protein to the Human Proteome:

RID YRFZHDGY01R (Expires on 08-15 03:43 am)

Query ID Id|159764

Description gi|657204379|gb|AID50426.1| ORF8b [Middle East respiratory syndrome coronavirus]

Molecule type amino acid

Query Length 112

① No significant similarity found. For reasons why,[click here](#)

Figure 2: The result of TFASTY showing no similarity of the MERS CoV Protein to the human proteome

ORF8b [Middle East respiratory syndrome coronavirus]

Databases chosen:

H. sapiens CDS (hs_rna.na)

Match Results:

No library sequences with E() < 5.00

Identification of the Antigenic Sites on the Foreign proteins:

All those proteins that showed no similarity to the human proteome was considered and foreign and the antigenic sites on these proteins were calculated using EMBOSS Antigenic. The results were shown below. A total of 5 proteins were selected to be foreign and the antigenic sites were identified in them.

Table 1: Total number of Antigenic Peptides found for each foreign Protein

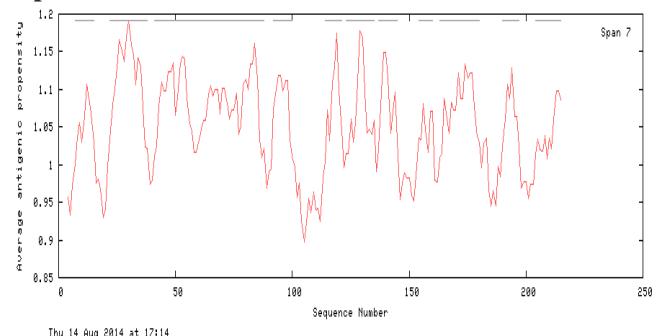
Protein Name	No of Antigenic Sites
ORF4b	11
ORF5	6
ORF8b	5
M protein	11
S protein	11

As per the above table 1 it can be summarized that out of the total proteome of MERS CoV the foreign proteins with good antigenic potential identified were only 5 whose antigenic peptides were 44 in number. These peptides must be further compared with the PVS server to identify the peptides having maximum antigenic propensity.

Calculation of Antigenic Propensity using PVS:

The five foreign proteins were further submitted to PVS server for the calculation of the antigenic Propensities of their peptides and to select the best ones.

Figure 3: The Antigenic Propensity Curve for the M protein



There are 11 antigenic determinants in your sequence:

n	Start Position	Sequence	End Position
1	7	LTEAQIIAI	15
2	22	AWSLIFLLITIVLQYGY	38
3	41	RSMTVYVFKMFVLWLLPSSMALSIFSAIYPIDLASQIISGVAAVSA	88
4	92	ISYFVQSIR	100
5	114	ETNCLLN	121
6	123	FGGTTVVRPLVED	135
7	137	TSVTAWTN	145
8	154	HFGACDY	160
9	163	LPNEVTVAKPVLIAALK	180
10	190	SGVAIYHR	197
11	204	RSPPITADIELA	215

Similar Curves were generated for the other 4 proteins and the best peptides out of all the peptides as shown above were selected. The summary of the work can be illustrated below.

Table 2: SAA Calculation for the best peptides:

Protein	Peptide	SAA
ORF8b	APLMLLQLLG	0.5
ORF4b	KKLRYVKRRFSLLRPEDLSVIVQ	0.28
	PTHYVRVTF	
ORF4b	NKQIVHSYPILH	0.16
ORF5	TDSIVFTYIPASGYVAALAVNVC	0.3
	LIPLLLLLRQ	
ORF5	IRTMVLYFLVLYNFLLAIVLVNG	0.35
	VHYPTGSCLIAFLVILILWFVDRI	
	RFCLMLNSYIPL	
M protein	AWSLIFLLITIVLQYGY	0.29

M protein	ETNCLLNV	0.5
M protein	FGGTTVVRPLVED	0.3
S protein	SVFLLMFLLPTESYVDVGPDSV KSACIEVDIQQ	0.29
S protein	FNHTLVLLPDGCGTLLGAFYCIL E	0.33

From the above table 2 it can be inferred that out of the total 10 peptides selected, only 2 peptides can be finalized for 3D structure construction step, as per the results of their SAA. The peptide having high SAA would have high ligand binding property when compared to the one with low SAA.

The selected 2 peptides of M protein and ORF8b having 0.5 SAA are further constructed in Argus lab [6].

Peptide Designing in Argus Lab:

Both the Peptides that were finalized in the previous step were now synthesized in Argus Lab.

Figure 4: The peptide APLMLLQLLG designed in Argus Lab.

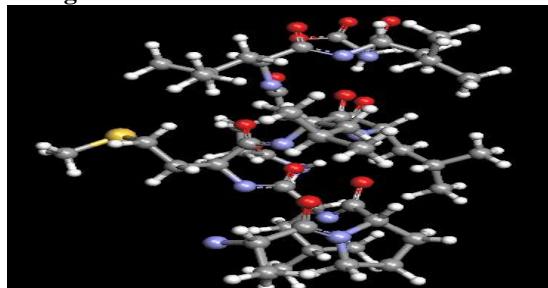
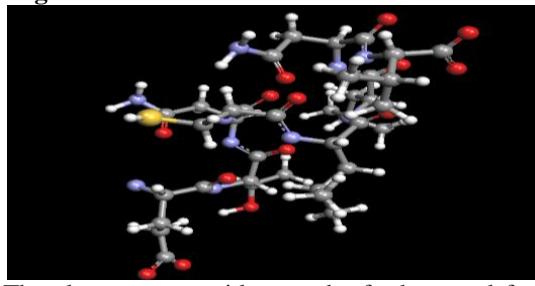


Figure 5: The peptide ETNCLLNV as designed in Argus Lab:



The above two peptides can be further used for the development of vaccine so as to inhibit the onset of MERS.

Conclusion:

The complete work involved the screening of the proteome of MERS CoV so as to identify a potent protein that was foreign to the human proteome followed by the identification of the Antigenic

regions within the selected peptides. Once all the antigenic peptides were clustered they were further compared for their antigenic propensity values using PVS server. The peptides with good antigenic propensities were further selected and their SAA (Surface Accessibility Area) were calculated. The peptides with High SAA and good antigenic propensity were finalized for Vaccine development. Here “APLMLLQLLG” and “ETNCLLNV” were selected and considered for the peptide model development in Argus Lab. The peptides are shown in the figures 4 and 5. The work concludes that these two peptides are identified to be the most antigenic and foreign proteins of MERS CoV and can be further used for the synthesis of viral vaccine.

Acknowledgement:

The authors would like to express my great appreciation to the university Dean Dr. Ghadeer Al Shaikh and Vice Dean Dr. Emaan Al Musaaed for their valuable and constructive support and motivation, their beloved parents and siblings for being there for them always especially for her brother for his patience and encouragement which he has shown throughout the project. Thank you all from the bottom of my heart

References:

- 1) <http://virologydownunder.blogspot.in/2014/04/professor-ali-mohamed-zaki-on-mers-cov.html>,
SUNDAY, 13 APRIL 2014
- 2) <http://www.cdc.gov/CORONAVIRUS/MERS/>
CDC : Centers for Disease Control and Prevention
- 3) Middle East Respiratory Syndrome Coronavirus (MERS-CoV), Travel Health notice: Updated: July 24, 2014
<http://www.phac-aspc.gc.ca/tmp-pmv/notices-avis/notices-avis-eng.php?id=108>
- 4) Middle East Respiratory Syndrome Virus (MERS CoV) Travel Health Notice.-**July 24, 2014.**
<http://www.phac-aspc.gc.ca/tmp-pmv/notices-avis/notices-avis-eng.php?id=108>
- 5) Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410. [PubMed](#)
- 6) Thompson, M. A. (2004). Molecular docking using ArgusLab, an efficient shape-based search algorithm and the AScore scoring function.- ACS meeting, Philadelphia, 172, CINF 42, PA