Content available at: https://www.ipinnovative.com/open-access-journals

IP International Journal of Medical Microbiology and Tropical Diseases

Journal homepage: www.ipinnovative.com

Short Communication

In-silico design of epitopes against SARS-CoV-2 based on membrane protein

Rudi Emerson De Lima Procopio[1,*]

[1] *Dept. of Microbiology, University of the State of Amazonas (Universidade Do Estado Do Amazonas), Brazil*

ARTICLE INFO

Article history:
Received 23-06-2020
Accepted 15-07-2020
Available online 13-10-2020

Keywords:
SARS-CoV-2

ABSTRACT

The selection of epitopes has become easier and faster with the aid of bioinformatics. SARS-CoV-2 is a single-stranded positive RNA virus (+ ssRNA) with approximately 30 kilobases, containing four main structural proteins: spike glycoprotein (S); small envelope protein (E); membrane protein (M); and nucleocapsid protein (N), which can generate epitopes. The protein chosen in this study was membrane glycoprotein ORF5, with 669 nt forming 222 amino acids. The analysis of the interaction of the protein with the membrane showed that most of the protein is inside the viral envelope and approximately 25 amino acids outside, where the initial and final regions of the protein are ideal for generating epitopes. The first 20 amino acids (MADSNGTITVEELKKLLEQW) were chosen because they are on the outer surface of the membrane of the viral envelope. For the production of heterologous protein in *Escherichia coli*, two expression systems (pET-28a and pFLAG-ATS) were used to carry out the various immunological tests necessary to validate the results obtained.

1. Introduction

Advances in bioinformatics have facilitated the selection of possible epitopes without the need for cultivation and the risks inherent in the manipulation of pathogens. This type of method is a great advantage over conventional vaccinology techniques, including faster results and lower costs (Soria-Guerra et al., 2015).[1] SARS-CoV-2 is a single-stranded positive sense RNA virus (+ ssRNA) belonging to the genus Betacoronavirus of the Coronaviridae family, with a genome of approximately 30 kilobases and that codes for several structural and non-structural proteins. The four main structural proteins of CoVs are the spike glycoprotein (S), small envelope protein (E), membrane protein (M) and nucleocapsid protein (N) (Ujike and Taguchi, 2015, Schoeman and Fielding 2019, Rahman et al., 2020),[2–4] all of which are encoded at the end of the genome. The focus of this work was thus the *in-silico* design and construction of epitopes for generation of antibodies against SARS-CoV-2, based on membrane protein. This viral proteins interact with receptors on the host cell membrane and are required for penetration of the viral genome into cells by fusion of the viral envelope with the plasma membrane or endosomal membranes (Holmes and Williams 1990).[5]

The SARS-CoV-2 sequence used for analysis was NC_045512.2, deposited with the NCBI (www.ncbi.nlm.nih.gov). The protein chosen was membrane protein ORF5, located in region 26,523 - 27,191 with 669 nt forming 222 aa (Figure 1). The similarity analysis of the membrane glycoprotein gene with other sequences deposited in the NCBI was 100% for MT585079, MT479226, LC553259, MT598178, MT577621, MT601295, MT577636 and MT577643, making it a relatively conserved gene. The analyses were performed in the Topology Data Bank of Transmembrane Proteins (TOPDB), Constrained Consensus TOPology Prediction Server (CCTOP) and Predict Secondary Structure (PSIPRED 4.0). The analysis of the interaction of the protein with the membrane showed that most of the protein is inside the viral envelope, as can be seen in Figure 1. The protein showed three regions inserted in the envelope membrane (underlined) and two

* *Corresponding author.*
E-mail address: rudiprocopio@gmail.com (R. E. D. L. Procopio).

https://doi.org/10.18231/j.ijmmtd.2020.043

regions outside the envelope (blue). In all, three analyses found the same topology, with approximately 120 amino acids inside the envelope (red) and 25 amino acids outside, as shown in blue in Figure 1.

MADSNGTITVEELKKLLEQWNLVIGFLFLTWICLLQFAYANRNRFLYIIKLIFLWLLWP
VTLACFVLAAVYRINWITGGIAIAMACLVGLMWLSYFIASFRLFARTRSMWSFNPETNI
LLNVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTLS
YYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIALLVQ

Fig. 1: Amino acid sequence of the SARS-CoV-2 membrane protein: in the blue region on the surface of the membrane that forms the envelope, in the red region on the inside of the envelope and in the underlined region, indicating interaction with the membrane. The blue highlighted region was used for epitope synthesis.

The first step is to discriminate the epitopes that are potentially immunoprotective from epitopes that are not. The methods used to predict continuous antibody epitopes from the protein sequence were: Sequential B-Cell Epitope Predictor (BepiPred-2.0), and Antibody Epitope Prediction (IEDB), both of which showed that the start and end regions can be used to generate epitopes, as shown in Figure 2Figure 2 .

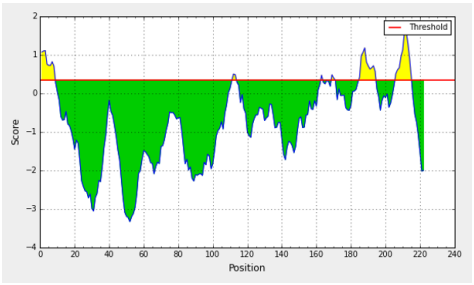

Fig. 2: Predicted B-cell epitopes using BepiPred 2.0 epitope predictor. Yellow areas above the thres hold (red line) are proposed as part of B cell epitopes in membrane proteins.

BepiPred-2.0 showed that the first 20 and the last 20 amino acids are the best regions to generate epitopes. The first 20 amino acids were chosen because they are on the surface of the viral envelope. The evaluation in CTLPred is a direct method for predicting CTL epitopes of T cells, where epitopes with 9 amino acids starting at amino acid 5 (score 0.99), 7 (score 0.79) and 8 (score 0.80) are excellent for generating epitopes. The analysis in AlgPred, based on amino acid composition, showed that the amino acid sequence has allergenic potential (score 0.5).

The DNA sequence used for synthesis was 60 bp (ATG GCA GAT TCC AAC GGT ACT ATT ACC GTT GAA GAG CTT AAA AAG CTC CTT GAA CAA TGG), generating a sequence of 20 amino acids (Met Ala Asp Ser Asn Gly Thr Ile Thr Val Glu Glu Leu Lys Lys Leu Leu Glu Gln Trp), with the first 20 amino acids in Table 1. For production of the heterologous protein in *Escherichia coli,* two expression systems were used in pET-28a (Novagen), with His tag for purification and pFLAG-ATS (Sigma) for purification and detection of the FLAG fusion protein. The synthesis was performed by Twist Bioscience. One of the first steps in epitope validation is to track immunoreactivity through serological analysis. This requires expression of the recombinant protein in a suitable host. *Escherichia coli* expression systems are the preferred choice for the production of recombinant proteins. The potential advantages of epitope-based vaccines also include increased safety, the opportunity to rationally design epitopes to increase potency and amplitude, and the ability to focus immune responses on conserved epitopes. Therefore, *in vivo* it is expected that this protein will interact with antigen-presenting cells, eliciting a potentially protective immune response (Shey et al., 2019)[6], in which isolated epitopes, capable of stimulating a specific immune response, can be identified and used to obtain advanced vaccine formulations, replacing those consisting of complete formulations of pathogens. The next step currently planned is to express this peptide in a bacterial system and carry out the various immunological tests necessary to validate the results obtained here through immuno-computer analysis.

2. Source of Funding

None.

3. Conflict of interest

None.

References

1. Soria-Guerra RE, Nieto-Gomez R, Govea-Alonso DO, Rosales-Mendoza S. An overview of bioinformatics tools for epitope prediction: Implications on vaccine development. *J Biomed Inf*. 2015;53:405–14.
2. Ujike M, Taguchi F. Incorporation of Spike and Membrane Glycoproteins into Coronavirus Virions. *Viruses*. 2015;7(4):1700–25.
3. Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. *Virol J*. 2019;16(1):69.
4. Rahman MS, Hoque MN, Islam MR, Akter S, Rubayet-Ul-Alam A, Siddique MA. Epitope-based chimeric peptide vaccine design against S, M and E proteins of SARS-CoV-2 etiologic agent of global pandemic COVID-19: an in silico approach. *bioRxiv*. 2020;.
5. Holmes KV, Williams RK. Background Paper Functions of Coronavirus Glycoproteins. *Adv Exp Med Biol*. 1990;276.
6. Shey RA, Ghogomu SM, Esoh KK, Nebangwa ND, Shintouo CM, Nongley NF, et al. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Scientific Rep*. 2019;9(1):4409.

Author biography

Rudi Emerson De Lima Procopio Teacher

Cite this article: Procopio REDL. *In-silico* design of epitopes against SARS-CoV-2 based on membrane protein. *IP Int J Med Microbiol Trop Dis* 2020;6(3):191-192.