The Spike Protein S1 Subunit of SARS−CoV−2 Contains an LxxIxE-like Motif that is Known to Recruit the Host PP2A-B56 Phosphatase

Halim Maaroufi

Institut de biologie intégrative et des systèmes (IBIS). Université Laval. Quebec. Canada

Halim.maaroufi@ibis.ulaval.ca

ABSTRACT The novel betacoronavirus (SARS-CoV-2) is highly contagious and can cause serious acute respiratory illness syndromes, often fatal, called covid-19. It is an urgent priority to better understand SARS-CoV-2 infection mechanisms that will help in the development of prophylactic vaccines and therapeutics that are very important to people health and socioeconomic stability around the world. The surface coronavirus spike (S) glycoprotein is considered as a key factor in host specificity because it mediates infection by receptor-recognition and membrane fusion. Here the analysis of CoV-2 S protein revealed in S1subunit a B56-binding LxxIxE-like motif that could recruit the host protein phosphatase 2A (PP2A). This motif is absent in SARS-CoV and MERS-CoV. PP2A is a major family of serine/threonine phosphatases in eukaryotic cells. Phosphatases and kinases are big players in the regulation of pro-inflammatory responses during pathogenic infections. Moreover, studies have shown that viruses use multiple strategies to target PP2A in order to manipulate host's antiviral responses. The latest studies have indicated that SARS-CoV-2 is involved in sustained inflammation in the host. Therefore, by controlling acute inflammation; it is possible to eliminate its dangerous effects on the host. Among efforts to fight covid-19, the interaction between LxxIxE-like motif and PP2A-B56-binding pocket could be a target for the development of a bioactive peptide and ligand inhibitors for therapeutic purposes.

KEYWORDS Coronavirus; SARS-CoV-2; spike S glycoprotein; PP2A-B56 phosphatase; LxxIxE-like motif; inflammation; therapeutic peptides.

INTRODUCTION

In March 11th 2020, the World Health Organization (WHO) announced that covid-19 situation is a pandemic because of the speed and scale of transmission. Coronaviruses (CoVs) are a large family of enveloped single positive-stranded RNA viruses that can infect both mammalian and avian species because their rapid mutation and recombination facilitate their adaptation to new hosts (Graham and Baric, 2010; Li, 2013). They can cause severe, often fatal acute respiratory disease syndromes named covid-19. CoVs are classified into *Alpha-*, *Beta-*, *Gamma-*, and *Deltacoronavirus* genetic genera. The novel betacoronavirus (betaCoVs) SARS-CoV-2 is relatively close to other betaCoVs: severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), bat coronavirus HKU4, mouse hepatitis coronavirus (MHV), bovine coronavirus (BCoV), and human OC43 coronavirus (HCoV-OC43). SARS-CoV emerged in China (2002–2003) and spread to other countries (more than 8,000 infection cases and a fatality rate of ~10%) (Peiris et al., 2003). In 2012, MERS-CoV was detected in the Middle East. It spread to multiple countries, infecting more than 1,700 people with a fatality rate of ~36%.

The surface-located SARS-CoV-2 spike glycoprotein S (S) is a 1273 amino acid residues. It is a homotrimeric, multidomain, and integral membrane protein that give coronaviruses the appearance of having crowns (*Corona* in Latin) (Li, 2016). It is a key piece of viral host recognition (receptor-recognition) and organ tropism and induces strongly the host immune reaction (Li, 2015). It is subdivided to S1 subunit that binds to a receptor on the host cell surface and S2 subunit that permits viral and host membranes fusion. S1 subunit is divided into two domains, an N-terminal domain (NTD) and a C-terminal receptor-binding domain (RBD) that can function as viral receptors-binding (Li, 2012). In addition, S1 subunit is normally more variable in sequence among different CoVs than is the S2 subunit (Masters, 2006).

Protein phosphatase 2A (PP2A) is a major family of serine/threonine phosphatases in eukaryotic cells and regulates diverse biological processes through dephosphorylation of numerous signaling molecules. PPA2 and phosphatase 1 (PP1), regulates over 90% of all ser/thr dephosphorylation events in eukaryotic cells (Eichhorn et al., 2009). PP2A is a heterotrimeric holoenzyme composed of a stable heterodimer of the scaffold A-subunit (PP2A-A) and catalytic C-subunit (PP2A-C) and a variable mutually exclusive regulatory subunit from four families (B (B55), B′ (B56), B″ and B‴)

which provide substrate specificity. The human B56 family consists of at least five different members (α, β, γ, δ and ε). Phosphatases and kinases are big players in the regulation of pro-inflammatory responses during microbial infections. Moreover, studies have revealed that viruses use multiple strategies to target PP2A in the aim to manipulate host antiviral responses (Guergnon et al., 2011). Here, face to urgent priority to fight the novel SARS-CoV-2 due to its grave consequences in the human health and socioeconomic stability around the world, S protein was analyzed because its importance in mediating infection. This analysis revealed in S1subunit a B56-binding LxxIxE-like motif that could recruit the host PP2A. The interaction S1 subunit-host PP2A-B56 could be a target for the development of a bioactive peptide and ligand inhibitors for therapeutic purposes.

RESULTS AND DISCUSSION

Two LxxIxE-like motifs in S1 and S2 subunits of Spike S

Sequence analysis of SARS-CoV-2 spike protein by the eukaryotic linear motif (ELM) resource (http://elm.eu.org/) revealed short linear motifs (SLiMs) known as LxxIxE-like motif ,^{293}**L**DP**L**S**E**^{298} in S1 subunit and ^{1197}**L**ID**L**Q**E**^{1202} in S2 subunit (Fig. 1). SLiMs are few amino acid residues (3-15) in proteins that facilitate protein sequence modifications and protein-protein interactions (Davey et al., 2012; Van Roey et al., 2014). Viruses are known to mutate quickly and thus create mimic motifs, on very short time scales, that could hijack biological processes in the host cell such as cell signaling networks (Davey et al., 2015; Via et al., 2015; Davey et al., 2011). Interestingly, ^{293}LDPLSET^{299} is only present in SARS-CoV-2 (Fig. 1A). It is absent in the other coronaviruses S protein analysed in this study. In order to interact with protein(s), ^{293}LDPLSE^{298} must be present at the surface of S1 subunit. Indeed, it is exposed in the surface of S1 subunit in the end of NTD (Fig. 3B). Additionally, this motif could be an antigenic epitope to generate antibodies and/or can help the design of vaccine components and immuno-diagnostic reagents. A second motif ^{1197}LIDLQEL^{1203} is present in S2 subunit. It is conserved in S2 subunit of SARS-CoV-2, SARS-CoV, SARS-like of bat from China and Kenya (Fig. 1B). These last betacoronaviruses are phylogenetically close (Fig. 2). Unfortunately, the region containing ^{1197}LIDLQEL^{1203} peptide has not been resolved in all known structures of spike S protein of coronaviruses to know if it is exposed in the surface of S2 subunit.

Interactions of ^{293}LDPLSET^{299}and ^{1197}LIDLQEL^{1203} with Subunit B56-PP2A

To compare the interactions between these peptides and B56 regulatory subunit of PP2A, molecular docking was performed with the software AutoDock vina (Trott and Olson, 2010). Fig. 4 shows that peptides are localized in the same region as pS-RepoMan peptide (PDBid: 5SW9) and important amino acids of LxxIxE-like motif are superposed with those of pS-RepoMan peptide (Fig. 4C). Interestingly, ^{293}LDPLSET^{299} contains a serine and threonine that could be phosphorylated generating a negative charge that will interact with positive patch in subunit B56 of PP2A, enhancing binding affinity (Fig. 4A) (Nygren and Scott, 2015). According to Autodock software, binding affinity of ^{293}LDPLSET^{299}is -4.8 Kcal/mol and this of ^{1197}LIDLQEL^{1203}is -3.5 Kcal/mol. The difference of binding affinity may explained by the phosphorylation of serine and threonine. It is known that the binding affinity of SLiMs is relatively weak (low µmolar range) (Gouw et al., 2018). This knowledge of molecular interactions of ^{293}LDPLSET^{299} and B56-PP2A will pave the way to design a peptide able to mimic the surface of B56-PP2A and strongly bind to ^{293}LDPLSET^{299} surface precluding PP2A's recruitment (Zaidman and Wolfson, 2016).

Protein phosphatase 2A and single RNA viruses

It has been shown in single RNA viruses, Ebola virus (EBOV) and Dengue fever virus (DENV) that they recruit the host PP2A through its regulatory subunit B56-binding LxxIxE motif to activate transcription and replication (Kruse et al., 2018; Oliveira et al., 2018). In addition, it has been shown in mice infected with rhinovirus 1B (the most common viral infectious agent in humans) an exacerbation of lung inflammation. Administrating Salmeterol (beta-agonist) treatment exerts anti-inflammatory effects by increasing PP2A activity. It is probable that beta-agonists have the potential to target distinct proinflammatory pathways unresponsive to corticosteroids in patients with rhinovirus-induced exacerbations. (Hatchwell et al., 2014). It is interesting to learn about Salmeterol drug and the possibility of using it in covid-19's patients with sustained and dangerous inflammatory reaction.

MATERIALS AND METHODS

Sequence analysis

To search probable short linear motifs (SLiMs), SARS-CoV-2 spike protein sequence was scanned with the eukaryotic linear motif (ELM) resource (http://elm.eu.org/).

3D modeling and molecular docking

For docking, the coordinates of the ^{293}LDPLSET^{299} peptide were extracted from spike S protein of CoV-2 structure (PDBid: 6VSB_A). Unfortunately, the region containing ^{1197}LIDLQEL^{1203} peptide has not been resolved in all known structures of spike S protein. So, Pep-Fold (Thevenet et al., 2012) software was used to model *de novo* this peptide. The model quality of the peptide was assessed by analysis of a Ramachandran plot through PROCHECK (Vaguine et al., 1999). The docking of the two peptides into regulatory subunit B56 of PPA2 (PDBid: 5SWF_A) was performed with the software AutoDock vina (Trott and Olson, 2010). The 3D complex containing regulatory subunit B56 of PPA2 and peptides was refined by using FlexPepDock (London et al., 2011), which allows full flexibility to the peptide and side-chain flexibility to the receptor. The electrostatic potential surface of the regulatory subunit B56 of PPA2 was realized with PyMOL software (http://pymol.org/).

Phylogeny

To establish the phylogenetic relationships between spike S protein of SARS-CoV-2 and representative betacoronaviruses, amino acid residues sequences were aligned with Clustal omega (Sievers et al., 2011) and a phylogenetic tree was constructed with MrBayes (Huelsenbeck and Ronquist, 2001) using: Likelihood model (Number of substitution types: 6(GTR); Substitution model: Poisson; Rates variation across sites: Invariable + gamma); Markov Chain Monte Carlo parameters (Number of generations: 100 000; Sample a tree every: 1000 generations) and Discard first 500 trees sampled (burnin).

ACKNOWLEDGMENTS

I would like to thank the IBIS bioinformatics group for their assistance.

CONFLICT OF INTERESTED

The author declares that he has no conflicts of interest.

REFERENCES

Davey, N.E., Cyert, M.S., Moses, A.M., 2015. Short linear motifs - ex nihilo evolution of protein regulation. Cell Commun. Signal. 13, 43. https://doi.org/10.1186/s12964-015-0120-z.

Davey, N.E., Travé, G., Gibson, T.J., 2011. How viruses hijack cell regulation. Trends Biochem. Sci. 36, 159–169. https://doi.org/10.1016/j.tibs.2010.10.002

Davey, N.E., Van Roey, K., Weatheritt, R.J., Toedt, G., Uyar, B., Altenberg, B., Budd, A., Diella, F., Dinkel, H., Gibson, T.J., 2012. Attributes of short linear motifs. Mol. Biosyst. 8, 268–281. https://doi.org/10.1039/c1mb05231d

Eichhorn, P.J.A., Creyghton, M.P., Bernards, R., 2009. Protein phosphatase 2A regulatory subunits and cancer. Biochim. Biophys. Acta 1795, 1–15. https://doi.org/10.1016/j.bbcan.2008.05.005

Gouw, M., Michael, S., Sámano-Sánchez, H., Kumar, M., Zeke, A., Lang, B., Bely, B., Chemes, L.B., Davey, N.E., Deng, Z., Diella, F., Gürth, C.-M., Huber, A.-K., Kleinsorg, S., Schlegel, L.S., Palopoli, N., Roey, K.V., Altenberg, B., Reményi, A., Dinkel, H., Gibson, T.J., 2018. The eukaryotic linear motif resource - 2018 update. Nucleic Acids Res. 46, D428–D434. https://doi.org/10.1093/nar/gkx1077

Graham, R.L., Baric, R.S., 2010. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. J. Virol. 84, 3134–3146. https://doi.org/10.1128/JVI.01394-09

Guergnon, J., Godet, A.N., Galioot, A., Falanga, P.B., Colle, J.-H., Cayla, X., Garcia, A., 2011. PP2A targeting by viral proteins: a widespread biological strategy from DNA/RNA tumor viruses to HIV-1. Biochim. Biophys. Acta 1812, 1498–1507. https://doi.org/10.1016/j.bbadis.2011.07.001

Hatchwell, L., Girkin, J., Dun, M.D., Morten, M., Verrills, N., Toop, H.D., Morris, J.C., Johnston, S.L., Foster, P.S., Collison, A., Mattes, J., 2014. Salmeterol attenuates chemotactic responses in rhinovirus-induced exacerbation of allergic airways disease by modulating protein phosphatase 2A. J. Allergy Clin. Immunol. 133, 1720–1727. https://doi.org/10.1016/j.jaci.2013.11.014

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755. https://doi.org/10.1093/bioinformatics/17.8.754

Kruse, T., Biedenkopf, N., Hertz, E.P.T., Dietzel, E., Stalmann, G., López-Méndez, B., Davey, N.E., Nilsson, J., Becker, S., 2018. The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. Mol. Cell 69, 136–145.e6. https://doi.org/10.1016/j.molcel.2017.11.034

Li, F., 2016. Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu Rev Virol 3, 237–261. https://doi.org/10.1146/annurev-virology-110615-042301

Li, F., 2015. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. J. Virol. 89, 1954–1964. https://doi.org/10.1128/JVI.02615-14

Li, F., 2013. Receptor recognition and cross-species infections of SARS coronavirus. Antiviral Res. 100, 246–254. https://doi.org/10.1016/j.antiviral.2013.08.014

Li, F., 2012. Evidence for a common evolutionary origin of coronavirus spike protein receptor-binding subunits. J. Virol. 86, 2856–2858. https://doi.org/10.1128/JVI.06882-11

London, N., Raveh, B., Cohen, E., Fathi, G., Schueler-Furman, O., 2011. Rosetta FlexPepDock web server--high resolution modeling of peptide-protein interactions. Nucleic Acids Res. 39, W249–53. https://doi.org/10.1093/nar/gkr431

Masters, P.S., 2006. The molecular biology of coronaviruses. Adv. Virus Res. 66, 193–292. https://doi.org/10.1016/S0065-3527(06)66005-3

Nygren, P.J., Scott, J.D., 2015. Therapeutic strategies for anchored kinases and phosphatases: exploiting short linear motifs and intrinsic disorder. Front. Pharmacol. 6, 158. https://doi.org/10.3389/fphar.2015.00158

Oliveira, M., Lert-Itthiporn, W., Cavadas, B., Fernandes, V., Chuansumrit, A., Anunciação, O., Casademont, I., Koeth, F., Penova, M., Tangnararatchakit, K., Khor, C.C., Paul, R., Malasit, P., Matsuda, F., Simon-Lorière, E., Suriyaphol, P., Pereira, L., Sakuntabhai, A., 2018. Joint ancestry and association test indicate two distinct pathogenic pathways involved in classical dengue fever and dengue shock syndrome. PLoS Negl. Trop. Dis. 12, e0006202. https://doi.org/10.1371/journal.pntd.0006202

Peiris, J.S.M., Lai, S.T., Poon, L.L.M., Guan, Y., Yam, L.Y.C., Lim, W., Nicholls, J., Yee, W.K.S., Yan, W.W., Cheung, M.T., Cheng, V.C.C., Chan, K.H., Tsang, D.N.C., Yung, R.W.H., Ng, T.K., Yuen, K.Y., SARS study group, 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 361, 1319–1325. https://doi.org/10.1016/s0140-6736(03)13077-2

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539. https://doi.org/10.1038/msb.2011.75

Thévenet, P., Shen, Y., Maupetit, J., Guyon, F., Derreumaux, P., Tufféry, P., 2012. PEP-FOLD:

an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res. 40, W288–93. https://doi.org/10.1093/nar/gks419

Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461. https://doi.org/10.1002/jcc.21334

Vaguine, A.A., Richelle, J., Wodak, S.J., 1999. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. Acta Crystallogr. D Biol. Crystallogr. 55, 191–205. https://doi.org/10.1107/S0907444998006684

Van Roey, K., Uyar, B., Weatheritt, R.J., Dinkel, H., Seiler, M., Budd, A., Gibson, T.J., Davey, N.E., 2014. Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. Chem. Rev. 114, 6733–6778. https://doi.org/10.1021/cr400585q

Via, A., Uyar, B., Brun, C., Zanzoni, A., 2015. How pathogens use linear motifs to perturb host cell networks. Trends Biochem. Sci. 40, 36–48. https://doi.org/10.1016/j.tibs.2014.11.001

Wang, X., Bajaj, R., Bollen, M., Peti, W., Page, R., 2016. Expanding the PP2A Interactome by Defining a B56-Specific SLiM. Structure 24, 2174–2181. https://doi.org/10.1016/j.str.2016.09.010

Zaidman, D., Wolfson, H.J., 2016. PinaColada: peptide–inhibitor ant colony ad-hoc design algorithm. Bioinformatics 32, 2289–2296. https://doi.org/10.1093/bioinformatics/btw133

Figures

Figure 1. Multiple alignment of the spike glycoprotein of betacoronaviruses using Clustal omega (Sievers et al., 2011). LxxIxE-like motifs are indicated by green stars. Numbers at the start of each sequence corresponding to the GenBank and UniProt accession number. Green stars indicated LxxIxE-like motif. The figure was prepared with ESPript (http://espript.ibcp.fr).

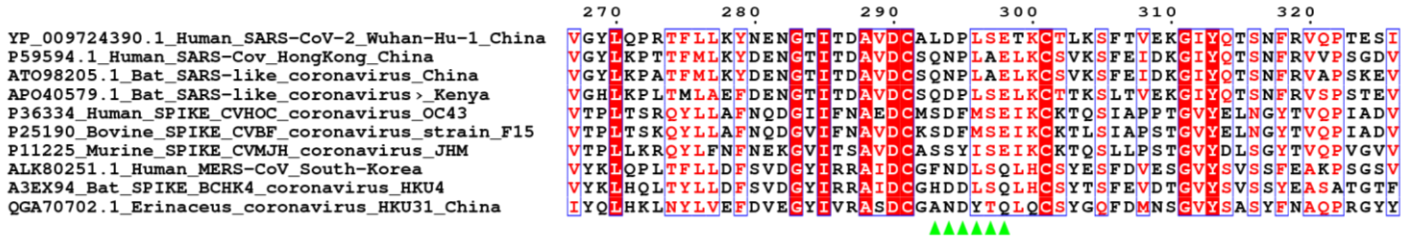
Figure 2. Unrooted phylogenetic tree of spike protein of representative betacoronaviruses. The tree was constructed using Mr Bayes method (Huelsenbeck and Ronquist, 2001) based on the multiple sequence alignment by Clustal omega (Sievers et al., 2011). Numbers at the start of each sequence corresponding to the GenBank and UniProt accession number. Red rectangle assembles betacoronaviruses with the same ^{1197}LIDLQE^{1202}. Green star indicated the only betacoronavirus with ^{293}LDPLSE^{298}.

Figure 3. (A) Diagram representation of the S1 subunit of spike protein of SARS-CoV-2 colored by domain. N-terminal domain (NTD, cyan), receptor-binding domain (RBD, green), subdomains 1 and 2 (SD1-2, orange) and the localization of ^{293}LDPLSE^{298} in the end of NTD.
(B) Surface structure representation of the S1subunit of spike protein (PDBid: 6VSB_A). ^{293}LDPLSE^{298} peptide is localized in the surface (red).

Figure 4. Electrostatic potential surface representation of the region of the regulatory subunit B56 of PP2A (PDBid: 5SWF_A) with docked peptides. (A) ^{293}**L**DP**L**p**S**E**p**T^{299} (green), (B) ^{1197}**L**ID**L**Q**E**L^{1203} (cyan) and (C) ^{293}**L**DP**L**p**S**E**p**T^{299} superposed to pS-RepoMan (^{581}RDIASKKPL**L**pSP**I**P**E**LPEVPE^{601}) peptide (orange, PDBid: 5SW9_B). The surfaces are colored by electrostatic potential with negative charge shown in red and positive charge in blue. Images were generated using PyMol (www.pymol.org).

294

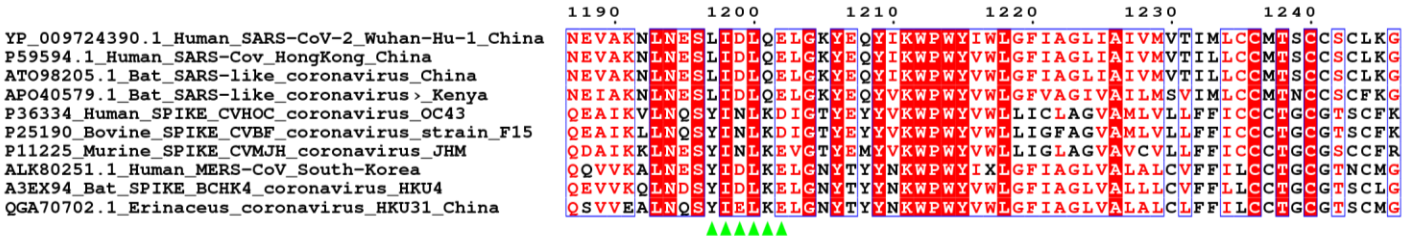
A



295

296

B

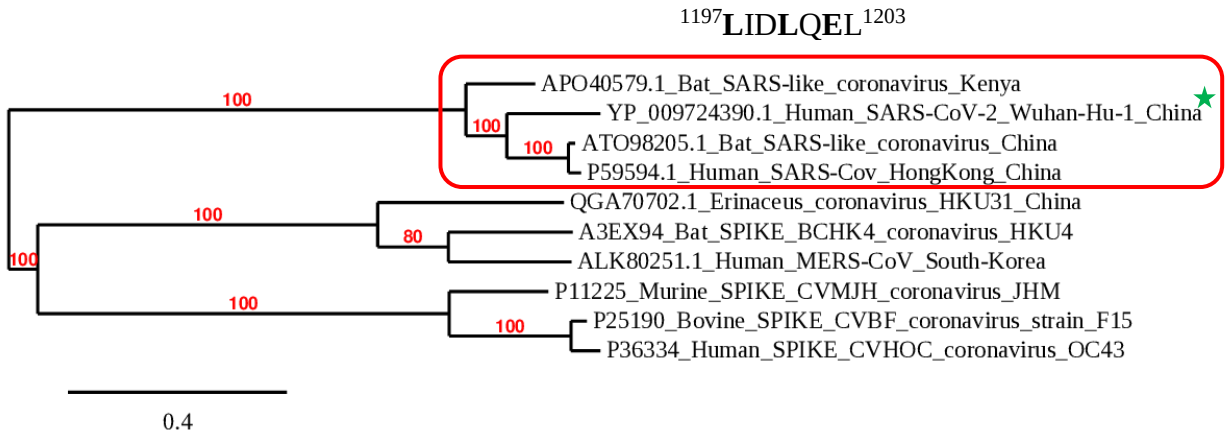


297

298 **Fig. 1**

299

300



301

302 **Fig. 2**

303

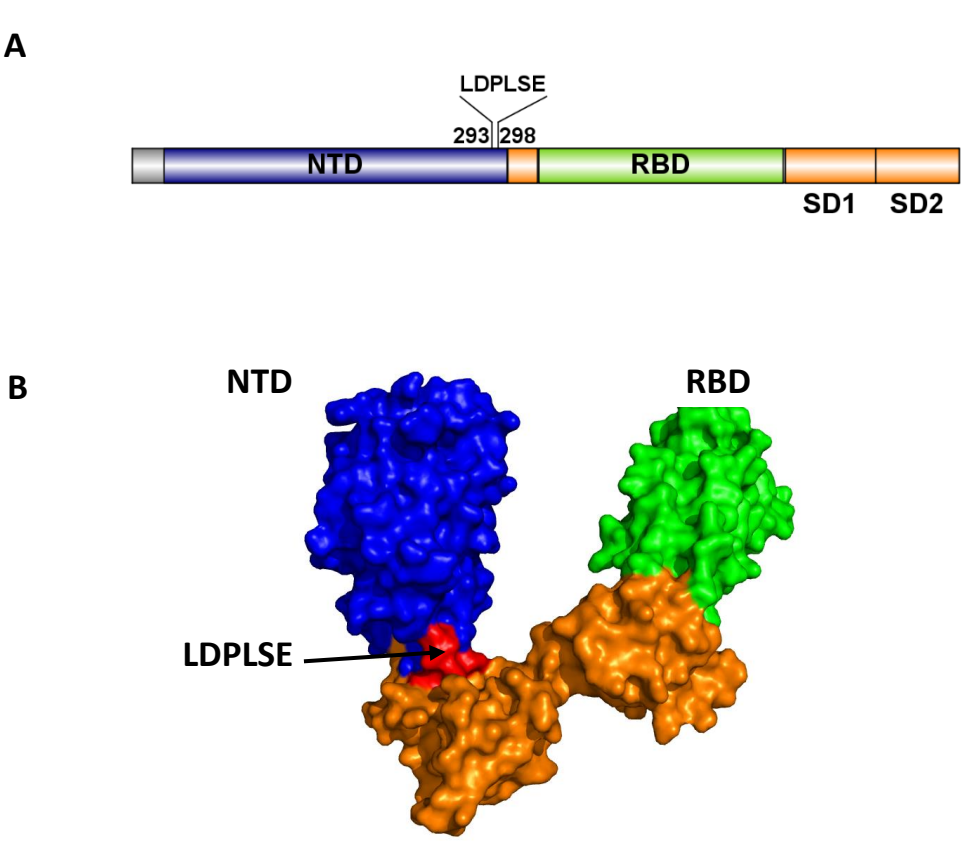


Fig. 3

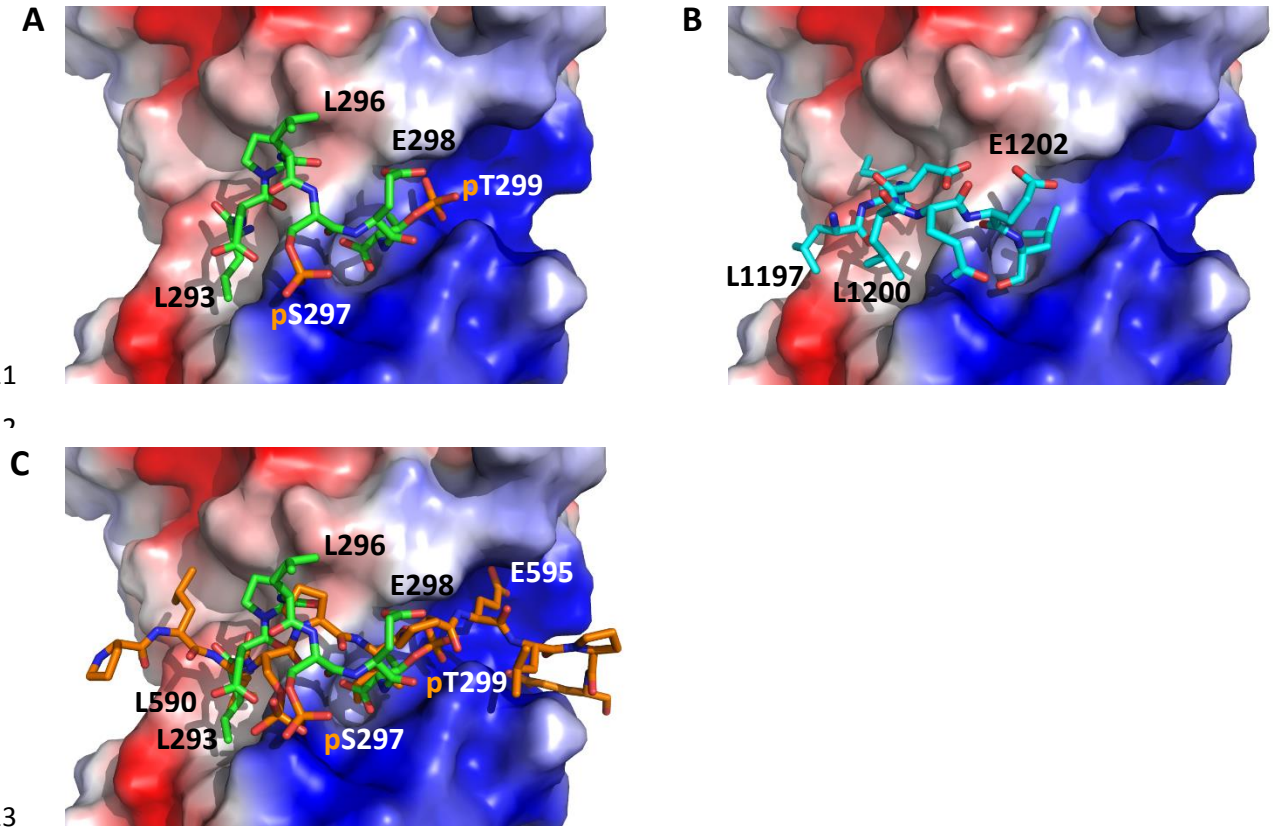


Fig. 4