

Using wide biological pores to cap and contain the COVID-19 spike protein

G. Sampath

Abstract. Geometric analysis shows that the spike (S) protein in the COVID-19 virus (SARS-Cov-2) can fully or partially enter into the channel of a wide biological pore like perforin (PFN) or streptolysin (SLO) when the latter is anchored in a bilayer lipid membrane. The PFN channel is a β barrel formed from multiple monomers, for example a ~ 14 nm diameter channel is formed from 22 monomers. Coincidentally the wide canopy of S (which has three identical chains) has an enclosing diameter of ~ 14 nm. While inside the channel peripheral residues in the canopy may bind with residues on the pore side of the barrel. If there are no adverse cross-reactions this would effectively prevent S from interacting with a target cell. Calculations with data obtained from PDB and other sources show that there are ~ 12 peripheral residue triples in S within a circle of diameter ~ 14 nm that can potentially bind with 22 exposed residues in each barrel monomer. The revised Miyazawa-Jernighan matrix is used to calculate the binding energy of canopy-PFN barrel residue pairs. The results show a large number of binding pairs over distances of up to 38 Å into the pore. This geometric view of capture and containment points to the possibility of using biological pores to neutralize SARS-Cov-2 in its many variant forms. Some necessary conditions that must be satisfied for such neutralization to occur are noted.

1. Introduction

The first step in the infection of a host cell by the COVID-19 virus is usually the attachment of the spike (S) protein jutting out of the virus to the ACE2 receptor protein on the membrane of the target cell [1,2]. Many antiviral drugs are designed to neutralize S by preventing such attachment. Broadly there are three approaches to counter the virus [3]: 1) neutralize the receptor-binding domain (RBD) in S by targeting residues in it [4]; 2) block pockets in other domains of S [5]; 3) directly deactivate the ACE2 receptor [6]. The inhibiting agent may be one of the following depending on the approach used: antibody proteins [7], peptides [8,9], nanobodies [10], small molecules [11]. Here the possibility of capping and containing S inside a biological pore [12] is explored by looking at the geometries of S and the pore and considering how S can enter the pore and be retained inside the pore channel. Such an approach can in principle deactivate the RBD as well as any other pockets that may be present in the other domains of S. As it is based primarily on physical containment it is likely to be indifferent to local mutations or homological differences among variants, evolutionary or otherwise,. The sequences of S and the pore barrel are examined for contact residues in S and binding sites on the inner wall of the channel. Data are presented for the SARS-Cov2 virus.

2. Geometry of the COVID-19 spike protein

The SARS-Cov-2 (COVID-19) virus is a spherical particle with a protein membrane whose interior contains a protein capsid enclosing a single RNA string that codes for the virus. The membrane surface has about 25-40 spike-like protrusions, each of which is a protein with three identical sequence chains A, B, and C forming a mushroom shape with a canopy and a stem. There are two domains in the spike: S1, the canopy consisting of a receptor binding domain (RBD) and a non-terminal domain (NTD); and S2, below the canopy, which directs fusion of the virus to a target cell prior to invading it. The RBD may be closed or open outward to provide a binding site to the ACE2 receptor that is located on the surface of a target cell [1].

Here the focus is on the geometry of the SARS-Cov-2 spike protein. Figure 1 shows plan and side views of 6VXX. The two graphics were obtained from RasMOL views and are annotated to include other pertinent information, see below.

The geometrical properties of the spike can be studied in terms of its extent, which can be calculated in a number of ways, including such well-known methods as minimum enclosing ellipsoid, minimum enclosing right cylinder, and minimum enclosing box or parallelepiped, all of which are global in character [13]. Here attention is focused on the diameter of the canopy so that it can be related to biological pores, which are roughly symmetric cylinders. To this end structure data for the spike protein (PDB id: 6VXX) was downloaded from the RCSB website and atomic coordinates extracted; the results can be found in Supplementary File 1. The centroid of the three $C\alpha$ atoms of each residue triple in the three chains was calculated. From it the smallest enclosing diameter circle and the distances of the centroids from the centroids of the topmost and bottommost residues were computed. The results for the 19 largest diameter triples are shown in Table 1. Of these the eight residue triples with diameter within ~ 15 Å of 14 nm are shaded.

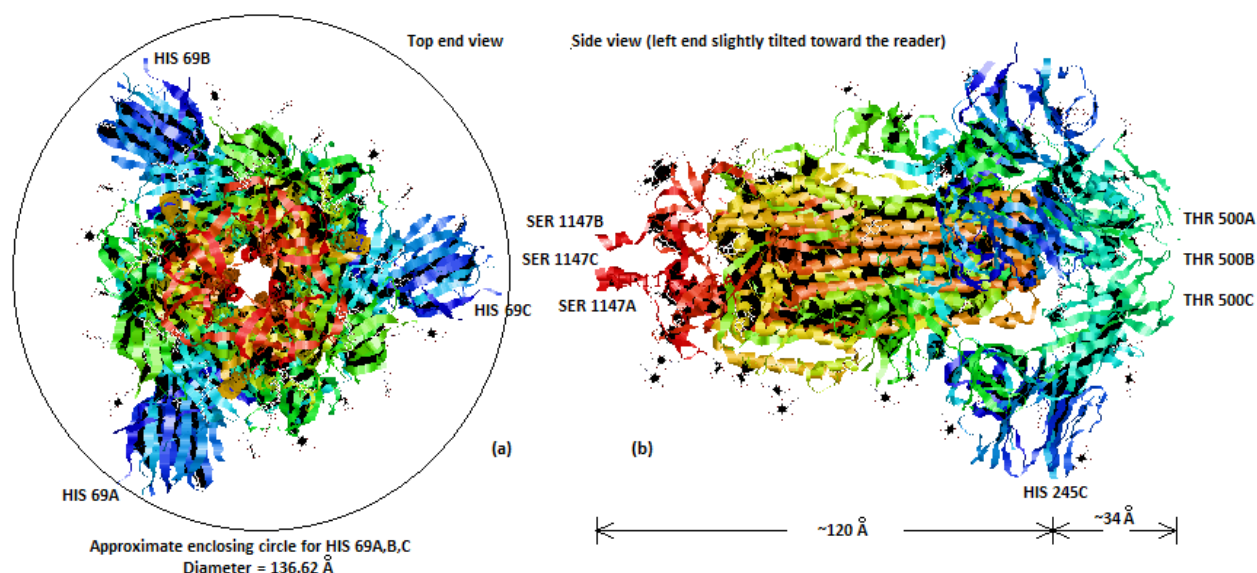


Figure 1. (a) Plan view of spike protein structure looking down from canopy side; smallest enclosing circle for all residues in spike protein is shown. (b) Side view of spike protein; distances of widest diameter residue (HIS 245 A, B, C) to tip of spike (residue SER 1147 A, B, C) and to top of canopy (residue 500 THR A, B, C) are shown.

Table 1. 19 triples (residues in A, B, C) with the largest enclosing diameters; the eight largest are shaded.

Residue	Residue number in sequence (A, B, and C chains)	Diameter of covering circle for triple (Å)	Distance from spike tip (residue triple sequence number 1147, SER) (Å)	Distance from top of canopy (residue triple sequence number 500, THR) (Å)
HIS	69	136.62	113.56	-42.24
HIS	245	135.35	120.5	-35.29
VAL	143	130.35	123.07	-32.72
ILE	68	130.18	112	-43.79
LEU	244	128.15	120.6	-35.19
GLY	142	126.62	124.93	-30.87
SER	98	125.87	112.09	-43.7
LYS	97	124.24	108.43	-47.37
ALA	67	122.91	113.01	-42.78
VAL	213	122.21	101.65	-54.15
ALA	243	121.72	120.12	-35.67
PHE	186	121.64	104.04	-51.75
ARG	214	120.38	100.68	-55.12
LYS	187	120.09	103.83	-51.96
ILE	100	119.79	115.58	-40.21
ALA	263	119.45	110.66	-45.13
ASN	211	119.38	98.6	-57.19
ALA	123	119.26	120.33	-35.47
ASN	99	119.19	113.96	-41.84

Figure 2 shows the projections of the 19 triples on to the XY plane. Note the clustering of the projection points, the result of 3-way symmetry in the structure of the protein (see Figure 1a).

3. Geometry of some biological pores

Pores play a major role in biological function, which may be benign or malign. Proteinaceous pores in the membrane of biological cells allow transport in one direction or the other of particles of various sizes, from Ca^{+} ions to large proteins. Such transport is usually benign and beneficial to normal cellular functions. Pores may also be formed by viruses to attack healthy cells. There are many pore-forming proteins in nature that are toxins and

function by puncturing the cell membrane to cause cell death [14]. On the other hand such an intrinsically malign property has been turned to advantage *in vitro* (and *in silico*) by using these pores as analytical tools in a wide range of applications, including sequencing/identification of biomolecules (DNA, protein, oligomers, peptides) [15,16], analysis of species [17], protein conformation and folding [18], and protein dynamics [19].

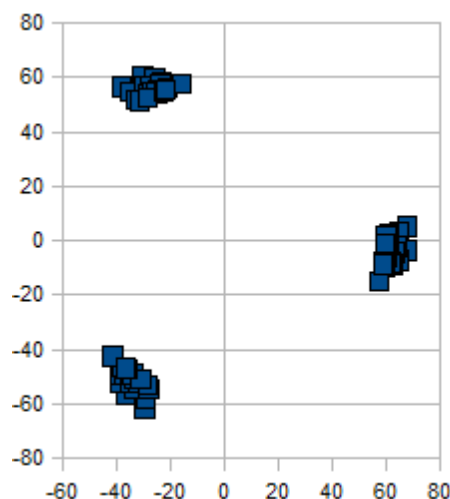


Figure 2. Projections onto XY plane of the 19 largest diameter residue triples in canopy of S

When the naturally occurring protein *alpha hemolysin* is added to an electrolytic cell in which a lipid bilayer separates two chambers *cis* and *trans*, it causes a pore to form through the membrane [20]. If an analyte like DNA or protein is added to the *cis* chamber and a potential difference is applied across *cis* and *trans* the analyte translocates through the pore to *trans* thereby causing changes in the pore current. These changes can be measured and used to identify bases in the DNA sequence, residues in the protein sequence, or the DNA or protein as a whole. Such pores usually have diameters of ~1-3 nm. Biological pores [21,22], however, have a wide range of diameters, going from 1 nm all the way up to the micron range. In the present context wide pores with diameters in the 10-50 nm range are of immediate interest.

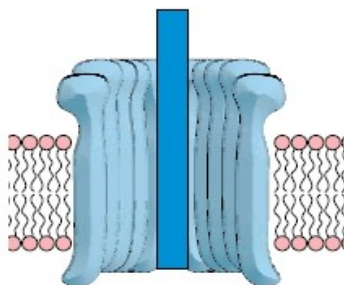


Figure 3. Cutaway view of PFN pore in a bilayer lipid membrane. (Extracted from Figure 1 in [26].)

Over the last few years, wide diameter pores have studied in depth [23]. Two wide pores of considerable interest are perforin (PFN) and streptolysin (SLO). PFN is a toxin that has been studied in detail in [24-26] and is a multifold pore. That is, it is formed from multiple monomers; 22 monomers can form a 14 nm diameter pore. Likewise SLO is a toxic protein that can form pores with inner diameter of up to 25 nm [27]. Both PFN and SLO can form pores that are full rings or arcs in a giant β barrel (a cylindrical structure made up of linked β sheets). The relationship between the pore diameter and the number of monomers has been studied in [28].

The present report focuses on the potential use of PFN as a device to cap and contain the spike protein in SARS-Cov-2. Figure 3 shows a cutaway view of a PFN pore anchored in a bilayer lipid membrane (graphic extracted from Figure 1 in [26]). The structure of the monomer is shown in Figure 4a below. The sequence of residues in the β sheet

at the bottom of the structure that face the inside of the pore is shown in Figure 4b, the figure is an adaptation from [26] .

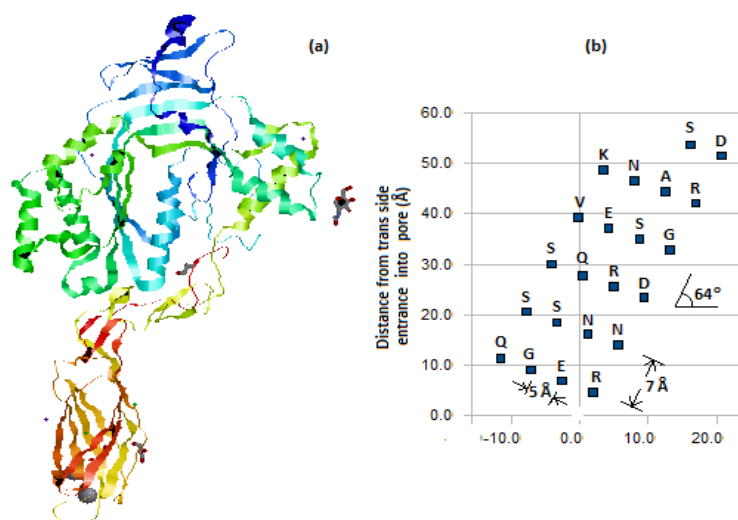


Figure 4. (a) Structure of single monomer of perforin (PFN). (b) Residues facing inside of pore channel

Table 2 shows the 22 residues in the β sheet at the bottom of the PFN monomer sorted by z value. This ordering on z is used to define a barrel array for use in calculating the total energy of the spike-barrel ensemble.

Table 2. Shows x and z coordinates of residues in Figure 4. Last column shows integer value of z.

Sorted on z	x	z	Truncated z
R	1.9	4.7	4
E	-2.6	6.9	6
G	-7.1	9.1	9
Q	-11.6	11.3	11
N	5.7	14.0	14
N	1.2	16.2	16
S	-3.3	18.4	18
S	-7.8	20.6	20
D	9.4	23.4	23
R	4.9	25.6	25
Q	0.5	27.8	27
S	-4.0	30.0	30
G	13.2	32.8	32
S	8.7	35.0	35
E	4.2	37.1	37
V	-0.3	39.3	39
R	17.0	42.1	42
A	12.5	44.3	44
N	8.0	46.5	46
K	3.5	48.7	48
D	20.8	51.5	51
S	16.3	53.7	53

Figure 5 shows the PFN β barrel schematically as a cut and unrolled cylinder laid out flat.

$\text{trunc}(0W/360)$ is 1. The contents of the two binary arrays are shown in Figure 7 with ■ or ■ used in place of 1.

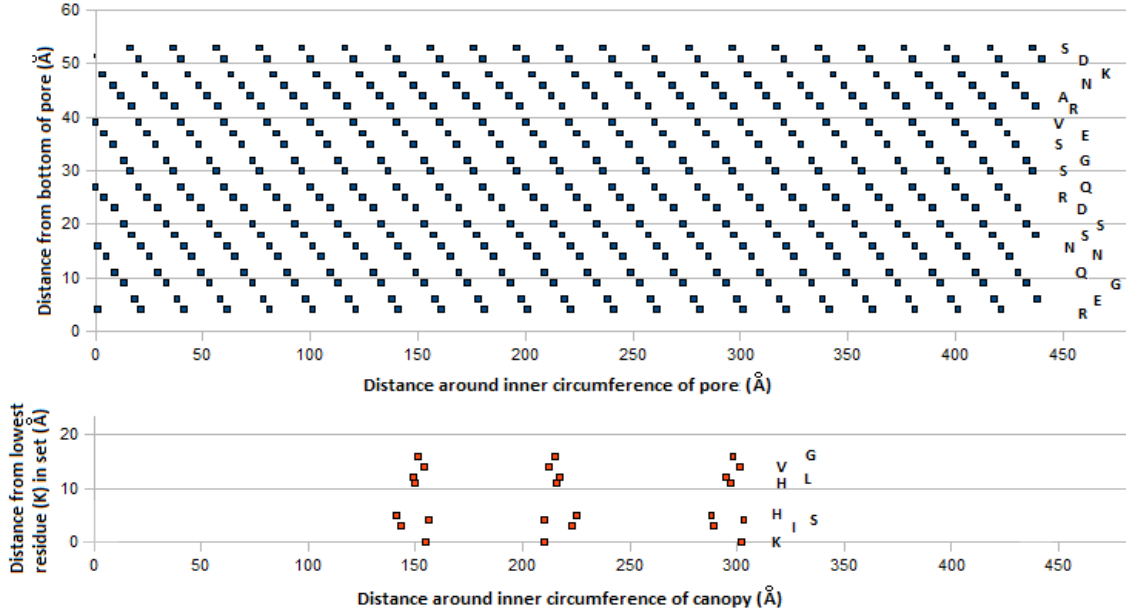


Figure 7. Contents of binary barrel and canopy matrices. ■ = ■ = 1; 0s not shown. Letters on the right are 1-letter names of residues on corresponding row.

Binding energy of pore-canopy ensemble

1) A simple energy model considers a canopy residue to be bound to (or coincident with) a pore residue if the corresponding entries in the barrel and canopy arrays are both 1. To determine binding pairs, the two arrays are AND-ed. Entries in the AND-ed result that are 1 corresponding to binding pairs (R_B , R_S), where B and S refer to the barrel and the spike. This is repeated for every circular position of the canopy by doing a circular rotation of the canopy array 1 column to the right; there are thus W instances of the canopy array. The canopy array is then moved 1 row up on the barrel array and the procedure repeated L-H+1 times.

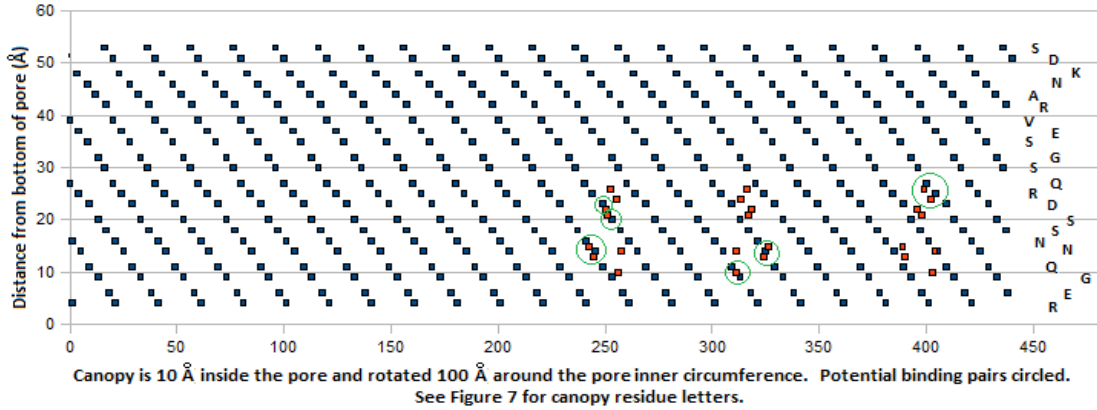


Figure 8. Canopy array (■) superimposed on barrel array (■). Spike has moved 10 Å into pore and rotated 80 degrees (with respect to Figure 3).

Figure 8 shows an example of the spike-barrel ensemble in which the spike has traveled $z_m = 10$ Å into the pore and rotated 100 Å around the inner circumference of the pore (≈ 80 degrees with reference to the angular position of the spike in Figure 3). Corresponding to this movement the canopy array has moved up 10 rows and shifted right by 100 columns relative to the barrel array.

2) A more involved model considers binding to occur between a canopy residue at $(z+z_{in},x)$, where z_{in} = distance of spike up the pore, and a barrel residue at any of the following locations in the barrel array: $(z+z_{in},x)$, $(z+z_{in}+1,x)$, $(z+z_{in}-1,x)$, $(z+z_{in},x+1)$, $(z+z_{in},x-1)$. The total energy is then the sum of the energies in all applicable cases.

Figure 9 shows the results of the two models.

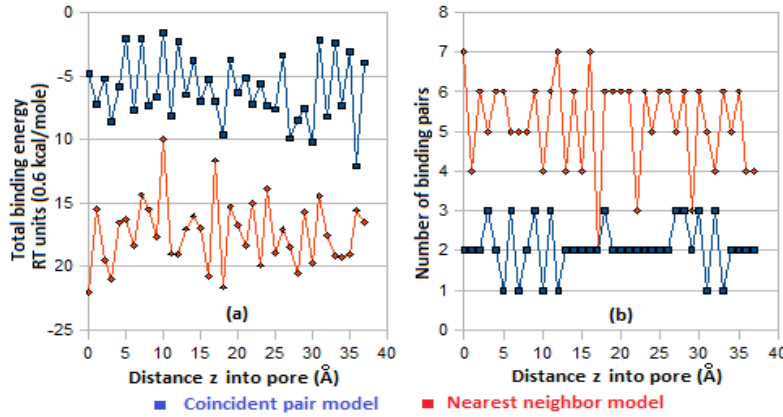


Figure 9. (a) Binding energy of pore-canopy ensemble as a function of the distance z_{in} traveled into the pore by the spike; (b) Number of binding pairs vs distance z_{in} .

5. Discussion

Translation of the method proposed above into an anti-viral protocol needs to consider a wide range of issues. Some of them are: a) developing an appropriate method of delivery; b) determining the right dosage, c) ensuring that the integrity of the bilayer lipid membrane is not diminished after delivery. The last is especially important because most biological pores, including PFN and SLO, are toxins in their natural form [14]; the cure should not be worse than the contagion. Some practical issues are noted next.

1) Improved binding between S and the pore channel can be achieved by artificially introducing residues into PFN through selective mutagenesis at a desired set of positions [35].

2) Notice that with the entire canopy of the spike capped by the pore, other pockets in the spike that can potentially bind with the target cell are effectively closed off.

3) The maximum number of coincidences between a residue on a barrel row and a residue on a canopy row (see Figure 8) is determined by how the number of monomers (22) and the number of canopy residues on any row (3) relate to each other. This number is minimal because 22 and 3 are relatively prime. It can be increased by using 21 or 24 monomers, but this also affects the diameter of the pore and hence its relationship to the diameter of a canopy residue triple.

4) Since capping is largely a physical process, whether the RBD (or any other domain) is conserved or not may not be an issue.

5) One of the problems with attempts to neutralize the spike is that the latter is covered with a glycan shield that thwarts attempts by antibodies to detect it [36]. Such evasion by the virus may not be an issue if the spike is capped by a pore as long as there are no other cross-reactions.

6) An interesting property of the SARS-Cov-2 virus is that the spherical membrane appears to be quite robust and resists mechanical attempts to change its shape [37].

7) When the pore (anchored in the bilayer lipid membrane) enters the extracellular environment, like any other foreign body it can be expected to trigger an immune response, possibly in the form of antibodies. As long as these antibodies do not block the pore or otherwise adversely impact the pore the latter's ability to cap and contain the spike will likely remain unaffected.

References

- [1] David Goodsell, "Molecule of the Month: SARS-CoV-2 Spike", June 2020, RCSB.org. doi: 10.2210/rcsb_pdb/mom_2020_6
- [2] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, and X. Wang, "Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor", *Nature*, 581, 215-229, 2020. doi:

10.1038/s41586-020-2180-5

- [3] S. Pomplun, "Targeting the SARS-CoV-2-spike protein: from antibodies to miniproteins and peptides", *RSC Med. Chem.*, 12, 197-202, 2021.
- [4] G. Zhang, S. Pomplun, A. R. Loftis, X. Tan, A. Loas, and B. L. Pentelute, "Investigation of ACE2 N-terminal fragments binding to SARS-CoV-2 Spike RBD", *bioRxiv preprint*, June 17, 2020. doi: 10.1101/2020.03.19.999318
- [5] S. Yazdani, N. De Maio, Y. Ding, V. Shahani, N. Goldman, and M. Schapira, "Genetic variability of the SARS-CoV-2 pocketome", *J. Proteome Res.* 2021. doi: 10.1021/acs.jproteome.1c00206
- [6] G. Wang, M.-L. Yang, Z.-L. Duan, F.-L. Liu, L. Jin, C.-B. Long, M. Zhang, X.-P. Tang, L. Xu, Y.-C. Li, P. M. Kamau, L. Yang, H.-Q. Liu, J.-W. Xu, J.-K. Chen, Y.-T. Zheng, X.-Z. Peng, and R. Lai, "Dalbavancin binds ACE2 to block its interaction with SARS-CoV-2 spike protein and is effective in inhibiting SARS-CoV-2 infection in animal models", *Cell Research*, 31, 17-24, 2021. doi: 10.1038/s41422-020-00450-0
- [7] C. O. Barnes, C. A. Jette, M. E. Abernathy, K.-M. A. Dam, S. R. Esswein, H. B. Gristick, A. G. Malyutin, N. G. Sharaf, K. E. Huey-Tubman, Y. E. Lee, D. F. Robbani, M. C. Nussenzweig, A. P. West Jr., and P. J. Bjorkman, "SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies", *Nature*, 588, 682-701, 2020. doi: 10.1038/s41586-020-2852-1
- [8] M. Tonk, D. Ružek, and A. Vilcinskas, "Compelling evidence for the activity of antiviral peptides against SARS-CoV-2", *Viruses*, 13, 912, 2021. doi: 10.3390/v13050912
- [9] W. Lin, J. Rafeya, V. Roschewitz, D. Smith, A. Keller, and Y. Zhang, "Peptide scanning of SARS-CoV and SARS-CoV-2 spike protein subunit 1 reveals potential additional receptor binding sites", *bioRxiv preprint*, August 16, 2021. doi: 10.1101/2021.08.16.456470
- [10] T. F. Custódio, H. Das, D. J. Sheward, L. Hanke, S. Pazicky, J. Pieprzyk, M. Sorgenfrei, M. Schroer, A. Gruzinov, C. Jeffries, M. Graewert, D. Svergun, N. Dobrev, K. Remans, M. A. Seeger, G. M. McInerney, B. Murrell, B. M. Hällberg, and C. Löw, "Selection, biophysical and structural analysis of synthetic nanobodies that effectively neutralize SARS-CoV-2", *bioRxiv preprint*, June 23, 2020. doi: 10.1101/2020.06.23.165415
- [11] D. Bojadzic, O. Alcazar, J. Chen, S.-T. Chuang, J. M. C. Capcha, L. A. Shehadeh, and P. Buchwald, "Small-molecule inhibitors of the coronavirus spike: ACE2 protein-protein interaction as blockers of viral attachment and entry for SARS-CoV-2", *ACS Infect. Dis.* 2021, 7, 1519-1534. doi: 10.1021/acsinfecdis.1c00070
- [12] H. Bayley and L. Jayasinghe, "Functional engineered channels and pores (Review)", *Molecular Membrane Biology*, 21, 209-220, 2004.
- [13] P. Schneider and D. H. Eberly, *Geometric Tools for Computer Graphics*. Elsevier, 2002.
- [14] H. Bayley, "Piercing insights: Pore-forming proteins", *Nature* 459, 651-652, 2009.
- [15] D. Deamer, M. Akeson, and D. Branton, "Three decades of nanopore sequencing". *Nat. Biotechnol.* 34, 518–524, 2016.
- [16] N. Varongchayakul, J. Song, A. Meller, and M. W. Grinstaff, "Single-molecule protein sensing in a nanopore: a tutorial", *Chem Soc Rev.* 47, 8512–8524, 2018. doi:10.1039/c8cs00106e
- [17] N. Liu, Z. Yang, X. Ou, B. Wei, J. Zhang, Y. Jia1, and F. Xia, "Nanopore-based analysis of biochemical species", *Microchim. Acta*, 183, 2016, 955–963. doi: 10.1007/s00604-015-1560-2
- [18] G. Oukhaled, J. Mathe, A.-L. Biance, L. Bacri, J.-M. Betton, D. Lairez, J. Pelta, and L. Auvray, "Unfolding of proteins and long transient conformations detected by single nanopore recording", *Phys. Rev. Lett.* 98, 158101, 2007.
- [19] S. Schmid and C. Dekker, "Nanopores – a versatile tool to study protein dynamics", *Report*, TU Delft, 2020.
- [20] H. Bayley, "Nanopore sequencing: from imagination to reality", *Clin Chem.* 61, 25–31, 2015. doi: 10.1373/clinchem.2014.223016
- [21] K. Huang, "Engineering biological nanopores for proteomics study", *Dissertation*, University of Groningen, 2019. doi: 10.33612/diss.102598418
- [22] S. Wang, "Development of New Biological Nanopores and Their Application for Biosensing and Disease Detection", *Dissertation*, U Kentucky, 2016. doi: 10.13023/ETD.2016.428
- [23] H. Watanabe, A. Gubbiotti, M. Chinappi, N. Takai, K. Tanaka, K. Tsumoto, and R. Kawano, "Analysis of pore formation and protein translocation using large biological nanopores", *Anal. Chem.* 2017. doi: 10.1021/acs.analchem.7b01550
- [24] R. H. P. Law, N. Lukoyanova, I. Voskoboinik, T. T. Caradoc-Davies, K. Baran, M. A. Dunstone, M. E. D'Angelo, E. V. Orlova, F. 'li Coulibaly, S. Verschoor, K. A. Browne, A. Ciccone, M. J. Kuiper, P. I. Bird, J. A. Trapani, H. R. Saibil, and J. C. Whisstock, "The structural basis for membrane binding and pore formation by lymphocyte perforin", *Nature*, 2010. doi: 10.1038/nature09518
- [25] I. Osinska, K. Popko, and U. Demkow, "Perforin: an important player in immune response", *Central Eur. J. Immunol.* 39, 109-115, 2014. doi: 10.5114/ceji.2014.42135

- [26] M. E. Ivanova, N. Lukoyanova, S. Malhotra, M. Topf, J. A. Trapani, I. Voskoboinik, and H. R. Saibil, "The pore conformation of lymphocyte perforin", *bioRxiv preprint*, July 3, 2021. doi: 10.1101/2021.07.03.450947
- [27] H. Yasuga, R. Kawano, M. Takinoue, Y. Tsuji, T. Osaki, K. Kamiya, N. Miki, and S. Takeuchi, "Logic gate operation by DNA translocation through biological nanopores", *PLoS ONE* 11:e0149667, 2016. doi:10.1371/journal.pone.0149667
- [28] C. F. Reboul, K. Mahmood, J. C. Whisstock, and M. A. Dunstone, "Predicting giant transmembrane β -barrel architecture". *Bioinformatics*, 2012. doi:10.1093/bioinformatics/bts152.
- [29] S. Miyazawa and R. L. Jernigan, "Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation", *Macromolecules*, 18, 534–552, 1985.
- [30] S. Miyazawa and R. L. Jernigan, "Residue–residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading", *J. Mol. Biol.* 256, 1996, 623–644.
- [31] J. C. Biro, "Amino acid size, charge, hydropathy indices and matrices for protein structure analysis", *Theor. Biol. and Medical Modelling*, 3, 15, 2006. doi:10.1186/1742-4682-3-15
- [32] A. Vangone, A. M. Bonvin, "Contacts-based prediction of binding affinity in protein–protein complexes", *eLife*. 2015;4:e07454
- [33] T. Siebenmorgen and M. Zacharias, "Computational prediction of protein–protein binding affinities", *Comput Mol Sci.*, 10:e1448, 2020. doi: 10.1002/wcms.1448
- [34] M. Heo, S. Maslov, and E. Shakhnovich, "Topology of protein interaction network shapes protein abundances and strengths of their functional and nonspecific interactions", *PNAS*, 108, 4258–4263, 2011. doi: 10.1073/pnas.1009392108
- [35] A. Crnkovic, M. Srnko, and G. Anderluh, "Biological nanopores: engineering on demand", *Life* 11, 27, 2021. doi: 10.3390/life11010027
- [36] O. C. Grant, D. Montgomery, K. Ito, R. J. Woods, "Analysis of the SARS-CoV-2 spike protein glycan shield: implications for immune recognition", *bioRxiv preprint*, May 1, 2020. doi: 10.1101/2020.04.07.030445
- [37] B. Kiss, Z. Kis, B. Pa'lyi, and M. S. Z. Kellermayer, "Topography, spike dynamics, and nanomechanics of individual native SARS-CoV-2 virions", *Nano Lett.* 21, 2675-2680, 2021.

Email: sampath_2068@yahoo.com