Journal of Applied Biological Sciences
E-ISSN: 2146-0108 16(1): 89-101, 2022
DOI: https://doi.org/10.71336/jabs.950
Research Article

GLYCOINFORMATICS APPROACH FOR IDENTIFYING TARGET POSITIONS TO INHIBIT INITIAL BINDING OF SARS-COV-2 S1 PROTEIN TO THE HOST CELL

Muhammet Uslupehlivan, Ecem Şener Uslupehlivan*

Ege University, Faculty of Science, Department of Biology, Molecular Biology Section, İzmir, Turkey.

**Corresponding Author:*
E-mail: ecem.sener@ege.edu.tr

(Received 22nd May 2021; accepted 11th August 2021)

ABSTRACT. COVID-19 outbreak is still threatening the public health. Therefore, in the middle of the pandemic, all kind of knowledge on SARS-CoV-2 may help us to find the solution. Determining the 3D structures of the proteins involved in host-pathogen interactions are of great importance in the fight against infection. Besides, post-translational modifications of the protein on 3D structure should be revealed in order to understand the protein function since these modifications are responsible for the host-pathogen interaction. Based on these, we predicted O-glycosylation and phosphorylation positions using full amino acid sequence of S1 protein. Candidate positions were further analyzed with enzyme binding activity, solvent accessibility, surface area parameters and the positions determined with high accuracy rate were used to design 3D O-glycoprotein structure of the S1 protein using carbohydrate force field. In addition, the interaction between the C-type lectin CD209L and α-mannose residues was examined and carbohydrate recognition positions were predicted. We suggest these positions as a potential target for the inhibition of the initial binding of SARS-CoV-2 S1 protein to the host cell.

Keywords: *CD209L–α-mannose interaction, 3D glycoprotein structure, O-glycosylation, phosphorylation, molecular docking*

INTRODUCTION

The coronavirus spike protein (S) plays a key role in the early steps of viral infection. S protein comprises of two subunits S1 and S2 which are responsible for the binding of the host cell receptor and fusion of the cellular membrane, respectively [1,2,3]. S protein also contains furin cleavage site at the boundary between the S1-S2 subunits which mediates the membrane fusion and virus infectivity [1,2]. It has been suggested that different domains within a single S protein could bind multiple alternative receptors. Although ACE2 (angiotensin-converting enzyme 2) is known to be the SARS-CoV receptor [2,4], CD209L, a C-type lectin that binds to high-mannose glycans on glycoproteins, has also been found to be as an alternative receptor for SARS-CoV [5]. C type lectins having characteristic C-type lectin-like domains (CTLDs) show specificity for mannose and galactose type carbohydrates thanks to conserved residues in the CTLD. Since C-type lectins have ability to recognize self and non-self ligands, they have numerous roles in the physiological functions including development, signalling, inflammation, homeostasis, cell death, and cancer. Their crucial role in immunity is to participate in the antigen presentation to T cells [6,7,8]. CD209L lectin has been found to

bind to SARS-CoV, human coronavirus 229E, Ebola, Hepatitis C, HIV, Influenza virus glycoproteins and may mediate the endocytosis of pathogens [9,10,11,12,13,14]. However, recent studies have mainly focused on the S protein-ACE2 interaction, there is a paucity of information on the SARS-CoV-2 S1 protein-CD209L lectin interaction. Thus, identification of the carbohydrate-lectin interaction sites on the 3D structure for understanding of the initial binding of virus to a host cell is crucial. Besides, as post-translational modifications regulate the host-pathogen interaction [15], identifying SARS-CoV-2 S1 protein modifications may help us to inhibit initial binding of the virus. SARS-CoV-2 S protein has been found to be a glycoprotein with N-linked, high mannose oligosaccharide chains. It contains 22 N-glycosylation sequons. Once transported from ER to the Golgi, O-glycan structures are also added to the S protein [16,17]. 3D crystal structure of SARS-CoV-2 S protein with N-glycosylation positions was identified but, furin cleavage and O-glycosylation sites are missing on the 3D structure [18]. As glycan structures typically exist in solution or on proteins, it is a big challenge to characterize the 3D structure of glycoproteins experimentally. However, computational structural biology allows us to generate the 3D protein and glycoprotein modelling with high accuracy rate using all amino acid sequence [19,20,21]. Herein, we predicted O-glycosylation and phosphorylation positions using full amino acid sequence of S1 protein. Candidate positions were further analyzed with enzyme binding activity, solvent accessibility, surface area parameters and the positions determined with high accuracy rate were used to design 3D glycoprotein structure of the S1 protein using carbohydrate force field. In addition, the interaction between the C-type lectin and α-mannose residues was examined and carbohydrate recognition positions were predicted.

MATERIALS AND METHODS

Prediction of O-Glycosylation and Phosphorylation Positions

Amino acid sequence of SARS-CoV-2 S1 protein was taken from NCBI with QHD43416 ID. The potential O-glycosylation (O-GalNAc) and O-β-GlcNAc sites of S1 protein were analysed via NetOGlyc 4.0 [22] and YinOYang 1.2 Server [23], respectively. The threshold was chosen as 0.5 for both to predict high potential sites. The potential glycosylation positions which are found on NetOGlyc and YinOYang were analysed with the ISOGlyP server for the enzyme binding activity [24]. ISOGlyP was used to calculate all the potential positions with ppGalNAc transferase isoforms which showed high enzyme binding activity and calculates the enhancement value product (EVP) values as an indication of glycosylation rates [25]. The potential phosphorylation positions of S1 protein was analysed with NetPhos 3.1 Server [26]. NetSurfP v1.1 was used to assess the surface and solvent accessibility of predicted Ser and Thr positions with glycosylation and phosphorylation positions [27]. The relative solvent accessibility of potential glycosylation positions was analysed using Poly-View 2D-SABLE protein structure prediction server [28].

Glycoprotein Building on 3D Structure using Carbohydrate Force Field

3D structure models involving full amino acid sequence of S1 protein was taken with QHD43416 ID from I-TASSER (Iterative Threading ASSEmbly Refinement) [29]. The structure model has been generated by the C-I-TASSER pipeline, which utilizes deep convolutional neural-network based contact-map predictions to guide the I-TASSER

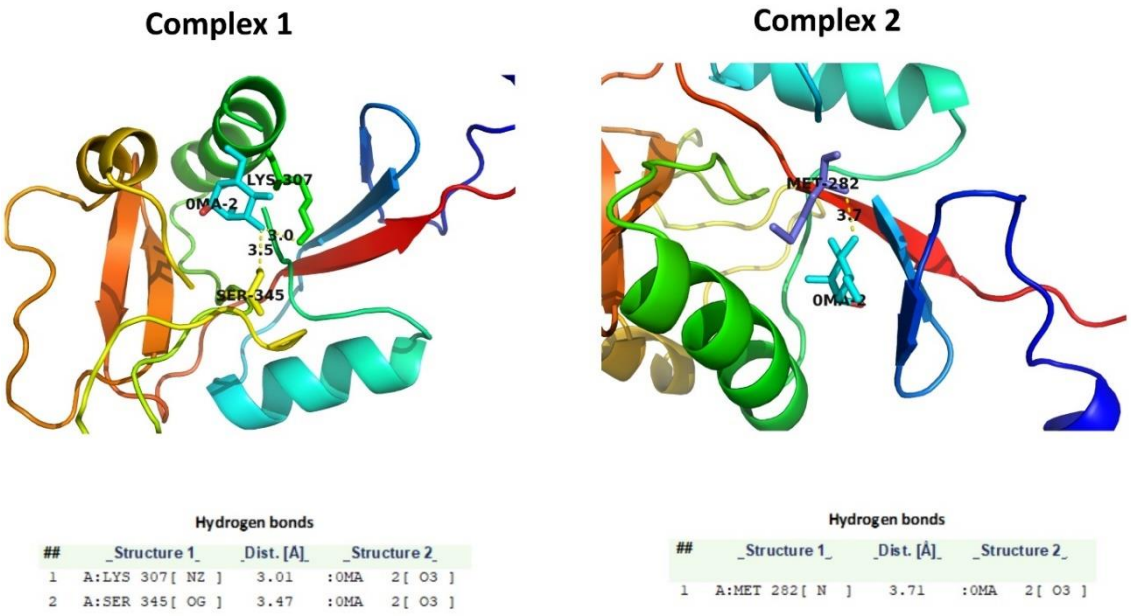
fragment assembly simulations. GLYCAM-Web Server (AMBER carbohydrate force field) was used to screen highly reliable glycosylation sites on the protein 3D structure. GLYCAM-Web is dedicated to simplifying the prediction of three-dimensional structures of carbohydrates and macromolecular structures involving carbohydrates. Glycoprotein Builder runs SASA (Solvent Accessible Surface Area) prediction and finds the most appropriate glycosylation sites for adding glycan units on the 3D structure using molecular dynamics simulation [30,31]. Glycan units skeleton was chosen as Core 1 type which is the most common O-glycan motifs on proteins for O-glycan units [32]. Additionally, all potential phosphorylation sites were analyzed for SASA parameters on the 3D protein structure using GLYCAM and the positions with high accuracy rates were chosen as phosphorylation positions.

Lectin-Carbohydrate Docking

In order to find out the CD209L lectin–mannose interaction, ZDOCK docking server was used [33]. CD209L lectin 3D structure (1.4 Å resolution) was taken from PDB (ID: 1XPH) [34] and α-D-mannose structure was taken from Glyco3D database [35]. PDBePISA server was used for the identification of macromolecular interfaces between the receptor-ligand [36]. All structures were visualized with PyMOL.

RESULTS AND DISCUSSION

Determining the 3D structures of the proteins involved in host-pathogen interactions are of great importance in the fight against infection. Besides, post-translational modifications of the protein on 3D structure should be revealed to understand the protein function since these modifications are responsible for the host-pathogen interaction. Therefore, researchers primarily have focused on the S protein structure of SARS-CoV-2 virus and revealed the 3D structure with N-glycosylation positions. But, the detailed structure of the O-glycosylation positions is not included in the current model [18]. Also, it is known that the initial binding of other coronaviruses to host cell occurs via the alternative receptor C-type lectin CD209L besides ACE2 protein [5]. Taken together, herein, we first studied the terminal mannose in the N-glycan structure of the S1 protein and CD209L interaction. When 3D structure of S1 protein analyzed, N-glycosylation positions has been found to be located mainly on the N-Terminal Domain (NTD) [18]. Since the Receptor Binding Domain (RBD) is known to be involved in ACE2 binding, we suggest that N-glycan positions on NTD may associated with the binding of C-type lectin. C-type lectins interact with α-linked mannose residues on mannosylated N-glycan structures [32]. There has been a structural study on C-type lectin-glycan interaction but terminal sugar is not α-mannose here [37]. Thus, in this study, the interaction between the C-type lectin and α-D-mannose was analyzed and interface residues were identified. ZDOCK docking and PDBePISA interface interaction analysis showed in two models with high accuracy rate that mannose sugar interacts with C-type lectin at Met282, Lys307, and Ser345 positions with hydrogen bonding (Fig. 1). We suggest these positions as the α-D-mannose recognizing sites having function on the lectin-sugar binding.



Hydrogen bonds

##	_Structure 1_	_Dist. [A]_	_Structure 2_
1	A:LYS 307[NZ]	3.01	:0MA 2[O3]
2	A:SER 345[OG]	3.47	:0MA 2[O3]

Hydrogen bonds

##	_Structure 1_	_Dist. [Å]_	_Structure 2_
1	A:MET 282[N]	3.71	:0MA 2[O3]

Fig. 1. *Receptor-ligand interactions of CD209L–α-D-mannose. ZDOCK and PDBePISA results showed that two complexes have receptor-ligand interaction by hydrogen bonds at Lys307, Ser345, and Met282 positions. The mannose ligand is shown in cyan, receptor positions Lys307 in green, Ser345 in yellow and Met282 in blue colors. All residues are shown in the stick representation.*

Secondly, we identified O-GalNAc and O-GlcNAc modifications of SARS-CoV-2 S1 protein. O-glycans affect virus transmission, immune escape of the virus, and virulence and pathogenicity of viruses [38,39]. For instance with O-glycan shielding, HIV antigenic sites are blocked and virus can escape from the neutralizing antibodies [38]. Further, O-glycans have been found to have a crucial role in pathogenicity of the avian influenza virus strain HPAI [40]. O-glycans are known to be involved in protein stability and function [41]. The presence of O-glycans in some viral proteins suggesting that glycans may play a role on the biological activity of viral proteins. In the comparative study on human SARS-CoV-2 and the S proteins of other coronaviruses have shown that Ser673, Thr678, and Ser686 are conserved O-glycosylated positions and have suggested that SARS-CoV-2 S1 protein may show O-glycosylation at these positions [42]. In this study, S1 protein was found to be O-GalNAcylated at Thr632, Thr678 and O-β-GlcNAcylated at Thr323, Thr638, Ser686 positions (Table 1 and Fig. 2). When compared our results with the previous study, SARS-CoV-2 S1 protein found not to be O-glycosylated at Thr673, but O-GalNAcylated at Thr678 and O-β-GlcNAcylated at Ser686. Besides, we found additional O-glycosylation positions at Thr323, Thr632, and Thr638 on the S1 protein. On the 3D glycoprotein structure, Thr323 was found to be located at RBD and Thr678 and Ser686 located near Furin Recognizing Site (FRS) (Fig. 2 and 3). Since O-glycans are involved in protein-protein interactions [41], we suggest that O-glycans at Thr323 may play a role on binding to ACE2 and O-glycans at Thr678, Thr686 may responsible for furin protease enzyme binding. It has been known that O-glycans are responsible for the protein stability and creating mucin like domain as glycan shields involved in immunoevasion [43]. Based on this, we suggest that the O-glycosylation

positions found may be involved in the stability of the S1 protein and immunoevasion of virus to the host cell.

Phosphorylation of viral proteins are catalyzed by host cell enzymes like glycosylation modifications [15,44]. In this study, phosphorylation modification of SARS-CoV-2 S1 protein was also examined and 36 positions of Ser/Thr residues were found to be phosphorylated on almost all domains of the protein (Table 2). Davidson et al. have shown that S1 protein is phosphorylated at 7 positions [45]. Amongst them, Ser459, Ser637, Ser640 positions were found to be correlated with our results. Phosphorylation modification regulates the protein activity and function in cooperation with glycosylation [46,47]. When we examined the phosphorylation and glycosylation locations of S1 protein on the 3D structure; the phosphorylation positions at Ser161 and Ser162 were found to be located close to glycosylation position Asn165. Likewise, Thr109 with Asn234; Ser680 with Thr678 and FRS; Thr604 with Asn603; Thr630 with Thr632; Thr637 and Thr640 with Thr638; Ser659 with Asn657 were found to be located close to each other on the 3D structure (Fig. 3). Therefore, we suggest these sites as critical sites for S1 protein activity.

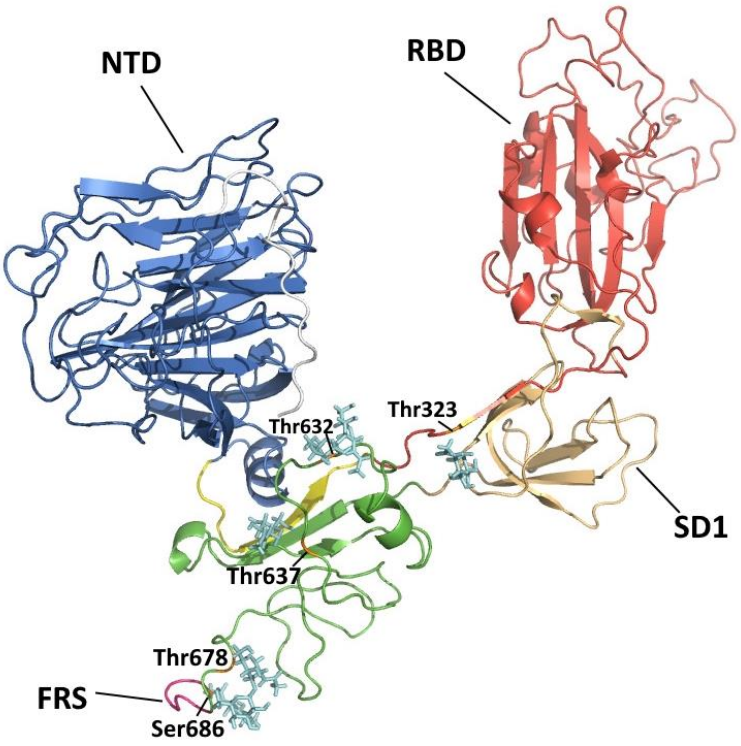


Fig. 2. *O-glycosylated 3D glycoprotein structure of SARS-CoV-2 S1 protein. Glycans are shown in cyan color. O-GalNAcylation positions: Thr632, Thr678; O-β-GlcNAcylation positions: Thr323, Thr638, Ser686. NTD: N-Terminal Domain, RBD: Receptor Binding Domain, FRS: Furin Recognizing Site. (The domain information was cited from Wong et al., 2004 and Wrapp et al., 2020.)*

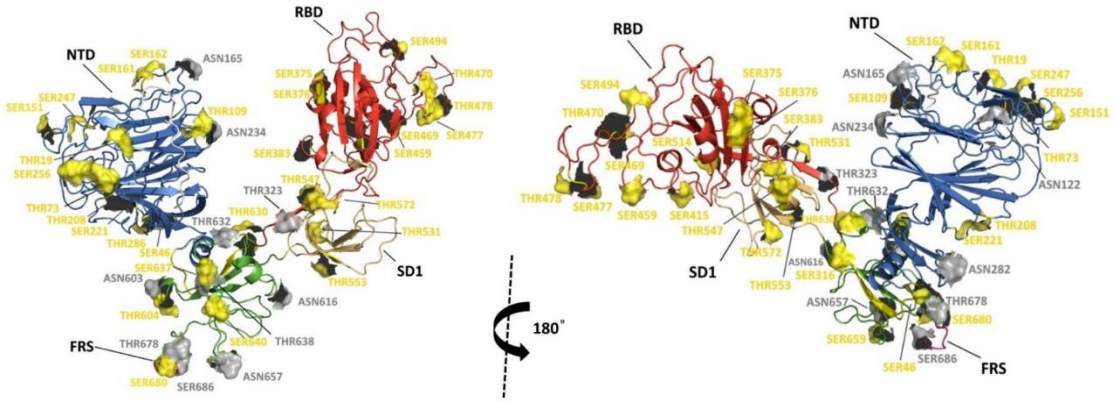


Fig. 3. *N-, O-glycosylation and Ser/Thr phosphorylation positions of SARS-CoV-2 S1 protein on the 3D structure. The glycosylation and phosphorylation positions are shown in grey and yellow, respectively. NTD: N-Terminal Domain, RBD: Receptor Binding Domain, FRS: Furin Recognizing Site (The domain information and N-glycosylation positions was cited from Wong et al., 2004 and Wrapp et al., 2020.).*

Table 1. *The predicted O-glycosylation positions of SARS-CoV-2 S1 protein.*

Glycosylation Position	NetOGlyc/YinOYang Score	ISOGlyP ppGalNAc Tranferase Isoforms Binding Value											Glycosylation Type
		T1	**T2**	**T3**	**T4**	**T5**	**T10**	**T11**	**T12**	**T13**	**T14**	**T16**	
	T/S Ratio	15:1	6:61	15:6	2:8	8:01	3:0	8:15	2:5	14:5	6:48	6:35	
T323†	0.5301	2.17	**6.96***	0.42	0.73	0.35	0.47	1.38	0.19	1.66	1.41	4.28	O-β-GlcNAc
T632ζ	0.5927	1.24	**2.87***	0.93	1.11	2.43	1.80	0.72	0.41	0.90	1.37	1.84	O-GalNAc
T638	0.5329	1.39	1.17	0.69	0.23	0.51	0.44	0.99	0.51	**1.48***	0.50	1.01	O-β-GlcNAc
T678	0.5458	1.34	0.44	1.03	1.94	2.20	0.23	1.51	**6.11***	1.28	0.34	0.60	O-GalNAc
S686ς	*0.6770*	*0.09*	*0.15*	*0.05*	*0.05*	*0.42*	*0.08*	*0.71*	*0.05*	***1.28****	*0.11*	*0.22*	*O-β-GlcNAc*

* indicates the highest enzyme binding score between the transferase isoforms. † indicates that position localized on Receptor Binding Domain and near the SD1 Variant Residue Glu324. ζ indicates that position is also phosphorylated. ς indicates that position is the first amino acid of S2 protein and localized near the FRS: Furin Recognizing Site (Protease Cleavage Site).

Table 2. *The predicted phosphorylation positions of SARS-CoV-2 S1 protein.*

	NetPhos		**NetSurfP**			**SABLE-POLYVIEW**	**Protein Localization**
Amino acid	Score	Kinase	Buried or Exposed Surface	RSA	ASA	RSA*	
T19	0.503	cdc2	E	0.320	44.426	2	NTD
S46	0.957 0.783 0.721	unsp PKC PKA	B	0.098	11.486	2	NTD
T73	0.642	PKC	E	0.359	49.738	2	NTD
T109	0.799 0.613 0.561	PKC unsp cdc2	B	0.196	27.227	1	NTD
S151	0.985 0.579	unsp CKII	E	0.390	45.731	3	NTD
S161	0.753	PKA	B	0.226	26.475	3	NTD
S162	0.557	PKA	B	0.298	34.984	4	NTD
T208	0.543	p38MAPK	E	0.329	45.618	3	NTD
S221	0.601	cdc2	E	0.312	36.555	3	NTD
S247	0.536	cdc2	E	0.351	41.196	2	NTD
S256	0.509	CKI	E	0.517	60.628	4	NTD
T286	0.549	CKII	B	0.170	23.565	2	NTD
S316	0.531	cdc2	E	0.443	51.931	2	Near RBD
S359	0.945 0.802 0.514	unsp PKA cdc2	B	0.295	34.527	1	RBD
S375	0.707	PKC	E	0.330	38.676	3	RBD
T376	0.846	PKC	E	0.379	52.526	3	RBD
S383	0.501	p38MAPK	E	0.341	39.942	2	RBD
T415	0.873	unsp	B	0.272	37.726	2	RBD
S459	0.895 0.504	unsp RSK	E	0.445	52.119	3	RBD (Also VR)
S469	0.968	unsp	E	0.479	56.139	2	RBD
T470	0.515	CKII	B	0.230	31.929	2	RBD
S477	0.805 0.510	PKC cdc2	B	0.244	28.620	2	RBD
T478	0.733	unsp	B	0.195	26.977	2	RBD (Also VR)
S494	0.538	PKC	E	0.280	32.757	4	RBD (Also VR)
S514	0.587	unsp	B	0.027	3.164	1	RBD
T531	0.625	PKC	B	0.314	43.566	2	SD1
T547	0.502	CKI	E	0.395	54.731	2	SD1
T553	0.521	CKII	E	0.299	41.444	4	SD1

Table 2. *(continued).*

	NetPhos		**NetSurfP**			**SABLE-POLYVIEW**	**Protein Localization**
Amino acid	Score	Kinase	Buried or Exposed Surface	RSA	ASA	RSA*	
T572	0.874 0.621 0.502	unsp cdk5 GSK3	E	0.276	38.240	2	SD1
T604	0.723	unsp	E	0.549	76.146	4	SD2 VR
T630	0.564 0.529	p38MAPK cdk5	E	0.368	51.000	2	Near GP T632
T632	0.859	PKC	B	0.172	23.926	1	Also O-Glycosylated
S637	0.967 0.579	unsp PKA	E	0.473	55.459	4	Near GP T638
S640	0.501	cdc2	E	0.484	56.690	3	Near GP T638
S659	0.957 0.511	unsp CKI	E	0.318	37.281	2	
S680	0.635	unsp	B	0.287	33.613	3	Near FRS and GP T678

All positions were analyzed with the kinase binding, surface, and solvent accessibility parameters. The kinase specificity and protein localizations were also given on the table. B: Buried, E: Exposed, RSA: Relative Surface Accessibility, ASA: Absolute Surface Accessibility, RSA*: Score reliability increases from 1 to 9. NTD: N-terminal Domain, RBD: Receptor Binding Domain (ACE2), VR: Variant Residue, GP: O-Glycosylation Position, FRS: Furin Recognizing Site (Protease Cleavage Site).

Many bacteria and viruses bind to mammalian host cells via specific glycan ligands. Thus, glycan mimetrics can be used to block the initial binding of the pathogens [21]. The CBAs (carbohydrate-binding agents), like large group of natural proteins, peptides and even synthetic agents that can interact with glycans [48]. Since viral surfaces of many enveloped viruses like HIV, DENV and hepatitis C are glycosylated, CBAs could interact with the glycosylated envelope of the virus and block the entry of viruses into host cells. Many lectins have previously been identified showing antiviral activity against hepatitis C and HIV [49,50,51]. Likewise, the same strategy may work here and lectins may be used to block the entry of the SARS-CoV-2 into the host cells.

CONCLUSION

COVID-19 outbreak is still threatening the public health. At this point, inhibiting the initial binding of virus to the host cell is crucial in order to find treatment. Since post-translational modifications regulate the host-pathogen interaction, identifying SARS-CoV-2 S1 protein modifications may help us to inhibit initial binding of the virus. Therefore, we focused on the CD209L lectin-α-D-mannose interaction of S1 protein and suggested Met282, Lys307, and Ser345 positions as targets for the initial binding. Also,

we found that phosphorylation and glycosylation positions where located at close sites may be critical for S1 protein activity.

Finally, glycan-based therapeutics are in great interest of vaccine and drug development [52,53,54]. For instance, envelope protein gp120 is expressed on the surface of HIV-1, and the cell type specific glycosylation of gp120 influences its binding to the receptors. Thus, adding new glycan epitopes to gp120 for the vaccine development has increased the ability of antibodies to recognize HIV-1 [21,55]. In the same way, protein glycosylation may be taken into account to develop a more effective SARS-CoV-2 vaccine. For this, the positions we suggested in this study may be used as a therapeutic target for vaccine or antiviral drug development against COVID-19.

Conflict of Interest. The authors declared that there is no conflict of interest.

Authorship Contributions. Concept: M.U., Design: M.U., E.Ş.U., Data Collection or Processing: M.U., Analysis or Interpretation: M.U., E.Ş.U., Literature Search: M.U., E.Ş.U., Writing: E.Ş.U.

Financial Disclosure. This research received no grant from any funding agency/sector.

REFERENCES

[1] Belouzard, S., Chu, V. C., Whittaker G. R. (2009): Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proceedings of the National Academy of Sciences 106(14): 5871-5876. https://doi.org/10.1073/pnas.0809524106.

[2] Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., Veesler, D. (2020): Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 181(2): 281-292. 10.1016/j.cell.2020.02.058.

[3] Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N. H., Nitsche, A., Müller, M. A., Drosten, C., Pöhlmann, S. (2020): SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 181(2): 271-280. https://doi.org/10.1016/j.cell.2020.02.052.

[4] Wong, S. K., Li, W., Moore, M. J., Choe, H., Farzan, M. (2004): A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. Journal of Biological Chemistry 279(5): 3197-3201. 10.1074/jbc.C300520200.

[5] Jeffers, S. A., Tusell, S. M., Gillim-Ross, L., Hemmila, E. M., Achenbach, J. E., Babcock, G. J., Thomas Jr., W. D., Thackray, L. B., Yung, M. D., Mason, R. J., Ambrosino, D. M., Wentworth, D. E., DeMartini, J. C., Holmes, K. V. (2004): CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. Proceedings of the National Academy of Sciences 101(44): 15748-15753. https://doi.org/10.1073/pnas.0403812101.

[6] Dambuza, I. M., Brown, G. D. (2015): C-type lectins in immunity: recent developments. Current Opinion in Immunology 32:21-27. https://doi.org/10.1016/j.coi.2014.12.002.

[7] Drickamer, K., Taylor, M. E. (2015): Recent insights into structures and functions of C-type lectins in the immune system. Current Opinion in Structural Biology 34:26-34. https://doi.org/10.1016/j.sbi.2015.06.003.

[8] Brown, G.D., Willment, J. A., Whitehead, L. (2018): C-type lectins in immunity and homeostasis. Nature Reviews Immunology 18:374-389. https://doi.org/10.1038/s41577-018-0004-8.

[9] Alvarez, C. P., Lasala, F., Carrillo, J., Muñiz, O., Corbí, A. L., Delgado, R. (2002): C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. Journal of Virology 76(13): 6841-6844. doi: 10.1128/JVI.76.13.6841-6844.2002.

[10] Lin, G., Simmons, G., Pöhlmann, S., Baribaud, F., Ni, H., Leslie, G. J., Haggarty, B. S., Bates, P., Weissman, D., Hoxie, J. A., Doms, R. W. (2003): Differential N-linked glycosylation of human immunodeficiency virus and Ebola virus envelope glycoproteins modulates interactions with DC-SIGN and DC-SIGNR. Journal of Virology 77(2): 1337-1346. doi: 10.1128/JVI.77.2.1337-1346.2003.

[11] Cormier, E. G., Durso, R. J., Tsamis, F., Boussemart, L., Manix, C., Olson, W. C., Gardner, J. P., Dragic, T. (2004): L-SIGN (CD209L) and DC-SIGN (CD209) mediate transinfection of liver cells by hepatitis C virus. Proceedings of the National Academy of Sciences 101(39): 14067-14072. https://doi.org/10.1073/pnas.0405695101.

[12] Marzi, A., Gramberg, T., Simmons, G., Möller, P., Rennekamp, A. J., Krumbiegel, M., Geier, M., Eisemann, J., Turza, N., Saunier, B., Steinkasserer, A., Becker, S., Bates, P., Hofmann, H., Pöhlmann, S. (2004): DC-SIGN and DC-SIGNR interact with the glycoprotein of Marburg virus and the S protein of severe acute respiratory syndrome coronavirus. Journal of Virology 78(21): 12090-12095. https://doi.org/10.1128/JVI.78.21.12090-12095.2004.

[13] Jeffers, S. A., Hemmila, E. M., Holmes, K. V. (2006): Human coronavirus 229E can use CD209L (L-SIGN) to enter cells. In: Perlman, S., Holmes, K. V. (ed.) The Nidoviruses. Advances in Experimental Medicine and Biology, Springer, Boston: MA, USA.

[14] Londrigan, S. L., Turville, S. G., Tate, M. D., Deng, Y. M., Brooks, A. G., Reading, P. C. (2011): N-linked glycosylation facilitates sialic acid-independent attachment and entry of influenza A viruses into cells expressing DC-SIGN or L-SIGN. Journal of Virology 85(6): 2990-3000. doi: 10.1128/JVI.01705-10.

[15] Sugrue, R. J. (2007): Glycovirology Protocols. Methods in Molecular Biology, Vol. 379, Humana Press, Totowa, New Jersey.

[16] Duan, L., Zheng, Q., Zhang, H., Niu, Y., Lou, Y., Wang, H. (2020): The SARS-CoV-2 Spike Glycoprotein Biosynthesis, Structure, Function, and Antigenicity: Implications for the Design of Spike-Based Vaccine Immunogens. Frontiers in Immunology 11:576622. https://doi.org/10.3389/fimmu.2020.576622.

[17] Sanda, M., Morrison, L., Goldman, R. (2021): N- and O-Glycosylation of the SARS-CoV-2 Spike Protein. Analytical Chemistry 93:2003-2009. https://doi.org/10.1021/acs.analchem.0c03173.

[18] Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., Graham, B. S., McLellan, J. S. (2020): Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367(6483): 1260-1263. DOI: 10.1126/science.abb2507.

[19] Uemura, S., Kurose, T., Suzuki, T., Yoshida, S., Ito, M., Saito, M., Horiuchi, M., Inagaki, F., Igarashi, Y., Inokuchi, J. (2006): Substitution of the N-glycan function in glycosyltransferases by specific amino acids: ST3Gal-V as a model enzyme. Glycobiology 16(3): 258-270. https://doi.org/10.1093/glycob/cwj060.

[20] Schultz, M. J., Swindall, A. F., Bellis, S. L. (2012): Regulation of the metastatic cell phenotype by sialylated glycans. Cancer and Metastasis Reviews 31(3-4): 501-518. https://doi.org/10.1007/s10555-012-9359-7.

[21] Varki, A., Cummings, R. D, Esko, J. D., Stanley, P., Hart, G.W. et al. (ed.) Essentials of Glycobiology (2015): Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, USA.

[22] Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. G., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Bennett, E. P., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., Clausen, H. (2013): Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO Journal 32(10): 1478-1488. https://doi.org/10.1038/emboj.2013.79.

[23] Gupta, R., Brunak, S. (2001): Prediction of glycosylation across the human proteome and the correlation to protein function. Pacific Symposium on Biocomputing 7, 310-322.

[24] Gerken, T. A., Raman, J., Fritz, T. A., Jamison, O. (2006): Identification of common and unique peptide substrate preferences for the UDP-GalNAc: polypeptide α-N-acetylgalactosaminyltransferases T1 and T2 derived from oriented random peptide substrates. Journal of Biological Chemistry 281(43): 32403-32416. https://doi.org/10.1074/jbc.M605149200.

[25] Gerken, T. A., Jamison, O., Perrine, C. L., Collette, J. C., Moinova, H., Ravi, L., Markowitz, S. D., Shen, W., Patel, H., Tabak, L. A. (2011): Emerging paradigms for the initiation of mucin-type protein O-glycosylation by the polypeptide GalNAc transferase family of glycosyltransferases. Journal of Biological Chemistry 286(16): 14493-14507. doi: 10.1074/jbc.M111.218701.

[26] Blom, N., Gammeltoft, S., Brunak, S. (1999): Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of Molecular Biology 294(5): 1351-1362. https://doi.org/10.1006/jmbi.1999.3310.

[27] Petersen, B., Petersen, T. N., Andersen, P., Nielsen, M., Lundegaard, C. (2009): A generic method for assignment of reliability scores applied to solvent accessibility predictions. BMC Structural Biology 9(1): 51. https://doi.org/10.1186/1472-6807-9-51.

[28] Adamczak, R., Porollo, A., Meller, J. (2004): Accurate prediction of solvent accessibility using neural networks–based regression. Proteins: Structure, Function, and Bioinformatics 56(4): 753-767. https://doi.org/10.1002/prot.20176.

[29] Zhang, Y. (2008): I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9(1): 40. https://doi.org/10.1186/1471-2105-9-40.

[30] Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., Woods, R. J. (2008): GLYCAM06: a generalizable biomolecular force field. Carbohydrates. Journal of Computational Chemistry 29(4): 622-655. https://doi.org/10.1002/jcc.20820.

[31] Nivedha, A. K., Makeneni, S., Foley, B. L., Tessier, M. B., Woods, R. J. (2014): Importance of ligand conformational energies in carbohydrate docking: Sorting the wheat from the chaff. Journal of Computational Chemistry 35(7): 526-539. https://doi.org/10.1002/jcc.23517.

[32] Taylor, M. E., Drickamer, K. (2006): Introduction to Glycobiology, 3rd edition. Oxford University Press, London, UK.

[33] Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T., Weng, Z. (2014): ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. Bioinformatics 30(12): 1771-1773. https://doi.org/10.1093/bioinformatics/btu097.

[34] Snyder, G. A., Colonna, M., Sun, P. D. (2005): The structure of DC-SIGNR with a portion of its repeat domain lends insights to modeling of the receptor tetramer. J. of Mol. Bio. 347(5): 979-989. doi: 10.1016/j.jmb.2005.01.063.

[35] Pérez, S., Sarkar, A., Rivet, A., Breton, C., Imberty, A. (2015): Glyco3D: a portal for structural glycosciences. In: Lütteke, T., Frank, M. (ed.) Glycoinformatics, Humana Press New York, USA doi: 10.1007/978-1-4939-2343-4_18.

[36] Krissinel, E., Henrick, K. (2007): Inference of macromolecular assemblies from crystalline state. Journal of Molecular Biology 372(3): 774-79. https://doi.org/10.1016/j.jmb.2007.05.022.

[37] Feinberg, H., Mitchell, D. A., Drickamer, K., Weis, W. I. (2001): Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. Science 294(5549): 2163-2166. doi: 10.1126/science.1066371.

[38] Li, Y., Liu, D., Wang, Y., Su, W., Liu, G., Dong, W. (2021): The Importance of Glycans of Viral and Host Proteins in Enveloped Virus Infection. Frontiers in Immunology 12:638573. https://doi.org/10.3389/fimmu.2021.638573.

[39] Iwashkiw, J. A., Seper, A., Weber, B. S., Scott, N. E., Vinogradov, E., Stratilo, C., Reiz, B., Cordwell, S. J., Whittal, R., Schild, S., Feldman, M. F. (2012): Identification of a General O-linked Protein Glycosylation System in Acinetobacter baumannii and Its Role

in Virulence and Biofilm Formation. Plos Pathogens 8(6):e1002758. https://doi.org/10.1371/journal.ppat.1002758.

[40] Mayr, J., Lau, K., Lai, J. C. C., Gagarinov, I. A., Shi, Y., McAtamney, S., Chan, R. W. Y., Nicholls, J., Itzstein, M., Haselhorst, T. (2018): Unravelling the Role of O-glycans in Influenza A Virus Infection. Scientific Reports 8:16382. https://doi.org/10.1038/s41598-018-34175-3.

[41] Cumming, D. A. (2003): Pathways and functions of mammalian protein glycosylation. New Comprehensive Biochemistry 38: 433-455. https://doi.org/10.1016/S0167-7306(03)38026-3.

[42] Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., Garry, R. F. (2020): The proximal origin of SARS-CoV-2. Nature Medicine 26: 450-452. https://doi.org/10.1038/s41591-020-0820-9.

[43] Bagdonaite, I., Wandall, H. H. (2018). Global aspects of viral glycosylation. Glycobiology 28(7): 443-467. doi: 10.1093/glycob/cwy021.

[44] Bretana, N. A., Lu, C. T., Chiang, C. Y., Su, M. G., Huang, K. Y., Lee, T. Y., Weng, S. L. (2012): Identifying protein phosphorylation sites with kinase substrate specificity on human viruses. PloS One 2012 7(7): e40694. https://doi.org/10.1371/journal.pone.0040694.

[45] Davidson, A. D., Kavangh Williamson, M., Lewis, S., Shoemark, D., Carroll, M. W., Heesom, K. J., Zambon, M., Ellis, J., Lewis, P: A., Hiscox, J. A., Matthews, D. A. (2020): Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. Genome Medicine 12:68. https://doi.org/10.1186/s13073-020-00763-0.

[46] Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., Brunak, S. (2004): Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4(6): 1633-1649. https://doi.org/10.1002/pmic.200300771.

[47] Hart, G. W., Kreppel, L. K., Comer, F. I., Arnold, C. S., Snow, D. M. (1996): O-GlcNAcylation of key nuclear and cytoskeletal proteins: reciprocity with O-phosphorylation and putative roles in protein multimerization. Glycobiology 6(7): 711-716. https://doi.org/10.1093/glycob/6.7.711.

[48] François, K. O., Balzarini, J. (2010): Potential of carbohydrate-binding agents as therapeutics against enveloped viruses. Medicinal Research Reviews 32(2): 349-387. https://doi.org/10.1002/med.20216.

[49] Balzarini, J., Herrewege, Y. V., Vermeire, K., Vanham, G., Schols, D. (2007): Carbohydrate-binding agents efficiently prevent dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin (DC-SIGN)-directed HIV-1 transmission to T lymphocytes. Molecular Pharmacology 71(1): 3-11. doi: 10.1124/mol.106.030155.

[50] Bertaux, C., Daelemans, D., Meertens, L., Cormier, E. G., Reinus, J. F., Peumans, W. J., Van Damme, E. J. M., Igarashi, Y., Oki, T., Schols, D., Dragic, T., Balzarini, J. (2007): Entry of hepatitis C virus and human immunodeficiency virus is selectively inhibited by carbohydrate-binding agents but not by polyanions. Virology 366(1): 40-50. https://doi.org/10.1016/j.virol.2007.04.008.

[51] Alen, M., Schols, D., Broad antiviral activity of carbohydrate-binding agents against dengue virus infection in Fa-Chang, C. (ed.) (2012): Carbohydrates-Comprehensive Studies on Glycobiology and Glycotechnology IntechOpen doi: 10.5772/50631.

[52] Osborn, H. M. I., Evans, P. G., Gemmell, N., Osborne, S. D. (2004): Carbohydrate-based therapeutics. Journal of Pharmacy and Pharmacology 56(6): 691-702. https://doi.org/10.1211/0022357023619.

[53] Cipolla, L., Araujo, A. C., Bini, D., Gabrielli, L., Russo, L., Shaikh, N. (2010): Discovery and design of carbohydrate-based therapeutics. Expert Opinion on Drug Discovery 5(8): 721-737. https://doi.org/10.1517/17460441.2010.497811.

[54] Fernandez-Tejada, A., Canada, F. J., Jimenez-Barbero, J. (2015): Recent developments in synthetic carbohydrate-based diagnostics, vaccines, and therapeutics. Chemistry 21(30): 10616-28. https://doi.org/10.1002/chem.201500831.

[55] Raska, M., Czernekova, L., Moldoveanu, Z., Zachova, K., Elliott, M. C., Novak, Z., Hall, S., Hoelscher, M., Maboko, L., Brown, R., Smith, P. D., Mestecky, J., Novak, J. (2014): Differential glycosylation of envelope gp120 is associated with differential recognition of HIV-1 by virus-specific antibodies and cell infection. AIDS Research and Therapy 11(23): 1-16. https://doi.org/10.1186/1742-6405-11-23.