We are IntechOpen,
the world's leading publisher of Open Access books
Built by scientists, for scientists

5,600
Open access books available

138,000
International authors and editors

175M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Post-Translational Modifications of Proteins Exacerbate Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV2)

Alok Raghav, Renu Tomar and Jamal Ahmad

Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV2) is severely affecting the worldwide population. It belongs to the coronavirus family which exhibit protein constituted enveloped single-stranded RNA. These viral proteins undergo post-translational modifications (PTMs) that reorganized covalent bonds and modify the polypeptides and in turn modulate the protein functions. Being viral machinery, it uses host cells system to replicate itself and make their copes, their proteins are also subject to PTMs. Glycosylation, palmitoylation of the spike and envelope proteins, phosphorylation, of the nucleocapsid protein are among the major PTMs responsible for the pathogenesis of the viral infection phase. The current knowledge of CoV proteins PTMs is limited and need to be exploring for to understand the viral pathogenesis mechanism and PTMs effect of infection phase.

Keywords: SARS-CoV-2, Post-translational modifications, Protein, glycosylation palmitoylation, phosphorylation

1. Introduction

SARS-CoV-2 is a culprit of the COVID-19 pandemic, caused 168, 599, 045 infections in people followed by 3,507,477 deaths worldwide as of 28th May 2021 according to the World Health Organization (WHO). Despite protective immunity and vaccination against SARS-CoV-2, it is spreading and affecting the worldwide population causing an increase in severity of illness due to the absence of the pre-existing immunity against SARS-CoV-2. This virus belongs to the family of coronavirus (CoV) with *Nidovirales* order exhibiting protein-containing enveloped positive-strand RNA that causes disease in both humans and animals. SARS-CoV-2 exhibits 30 kb genome size constituted in single-stranded RNA. Its genome comprises 6–11 open reading frames (ORFs) associated with 5′ and 3′ flanking untranslated regions (UTRs).

Coronaviruses are morphologically found spherical accounting average diameter of 80–120 nm with trimeric S-glycoprotein, sometimes with homodimeric HE protein [1, 2]. M-glycoprotein is the most abundant protein of virion which provides structural support to the virion. Moreover, E protein is also an essential protein needed for virion assemble and release [3, 4]. The nucleocapsid of the virus also comprised of the N protein. In the pathophysiology of virus replication, the S protein (180-200 kDa) plays a major role in binding and recognition with the

host cell via a cognate receptor(s). This trimeric S protein comprises two S1, S2, HR1 and HR2 subunits. This S protein comprised of two domain on its N-terminal extracellular transmembrane with another short intracellular C-chain domain [5]. The total length of the S protein of SARS-CoV-2 is 1273 amino acids (aa) arranged in a single peptide (1–13 aa) situated at N-terminus, S1 subunit (14–685 aa) and the S2 (686–1273 aa). The last two-sector is responsible for receptor binding and membrane fusion respectively. Furthermore, the S1 subunit comprises of N-terminal domain (14–305 aa) along with the receptor-binding domain (319–541 aa). Similarly, S2 protein constitutes the fusion peptide (788–806 aa), heptapeptide sequence (HR1) (912–984 aa), HR2 (1163–1213 aa), transmembrane domain (1213–1237 aa) along with cytoplasm domain ((1237–1273 aa) [6].

Post-translational modifications (PTMs) refers to the covalent bonds modifications of the proteins post-release from the ribosomes. PTMs adds new functional groups, like phosphate and carbohydrates along with other biological molecules of desired interest. PTMs is a naturally occurring process responsible for the regulation of protein folding, stability, enzymatic activity, protein to protein, and cell-to protein interaction. The most common and routinely occurring PTMs includes glycosylation, phosphorylation and lipidation (addition of such as palmitoylation and myristoylation) via proteolytic cleavage, formation of disulfide bonds. Proteins can also be modified through other covalent modifications like ubiquitination, sumoylation, glycation and neddylation.

PTMs are naturally occurring process catalyzed by enzymes. For instance, N-linked glycosylation needs series of enzyme reaction that generates precursor dolichol-linked oligosaccharide, oligosaccharyltransferase that initiate the transfer of the glycan to a specific consensus sequence (N-X-S/T, where X is any amino acid except proline) along with glycosidases and glycosyltransferases that are needed for the processing of N-linked glycan. In another series of example, protein ubiquitination needed three different enzymes including ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2) and ubiquitin ligases (E3) acting sequentially. It is a known fact that a virus utilizes host machinery to replicate its copies, therefore several viral proteins are susceptible to PTMs. Many viral proteins including structural, non-structural and accessory proteins are modified by PTMs which are affecting viral replication and pathogenesis.

2. N-linked glycosylation

N-linked glycosylation of the S protein of coronavirus is reported first time in the 1980s [7]. Murine strain MHV S protein acquire several mannose residues in the rough endoplasmic reticulum (ER). It was demonstrated that the S protein of infectious bronchitis virus (IBV), transmissible gastroenteritis coronavirus (TGEV) and bovine coronavirus is modified by the process of N-liked glycosylation [8]. S protein was also supposed to acquire mannose oligosaccharides and undergoes trimerization post entry to the endoplasmic reticulum (ER) before entering to Golgi complex [9]. In previously published literature the glycosylation sites in S protein of MHV predicted are 20 and 21 in numbers [10].

N-linked glycosylation plays important role in maintaining the conformation of coronavirus S protein, can affect the binding with the receptor of host cells and antigenicity of S protein. In a previously published study of IBV, it was found that mutation on the sites of N-linked glycosylation sites significantly causes shifting in the antigenicity of IBV [11]. In another study, it was found that TGEV infected cells when incubated with tunicamycin (inhibitor of N-linked glycosylation) showed a reduction in the antigenicity of both S and M protein [12]. These research findings

suggest that N-linked glycosylation is playing a vital role in maintaining the structure and conformation of the S protein and any mutations occurring at these glycosylation sites contribute to the reduction of antigenicity. It may be proposed here that mutations in the N-linked glycosylation in the S protein of SARS-CoV-2 can be beneficial in controlling the infections and reduction in their antigenicity. In another study conducted on IBV, it was demonstrated that N-D or N-Q mutations at the N-linked glycosylation sites N212 or N276 inhibit the function of S protein and in turn, hampers the cell-to-cell fusion and recognition [13].

Moreover, N-linked glycosylation of Dipeptidyl-peptidase 4 (DPP4), a similar receptor of Middle East respiratory syndrome coronavirus (MERS-CoV) significantly affect the binding of MERS-CoV S protein. E protein of SARS-CoV is another important constituent of the virus, it has been demonstrated that it contains two N-linked glycosylation sites N48 and N66, while IBV E protein contains one site at N5. In one of the studies published in past associated with SARS-CoV, it was demonstrated that transfected E protein with N-terminal Myc-tag showed glycosylation co-translationally [14]. However, two transmembrane domains are required to interact with SARS-CoV M protein and the hydrophilic region [14].

In previously published literature it was demonstrated that M protein of α-coronavirus transmissible gastroenteritis virus (TGEV) and Porcine Epidemic Diarrhea Virus (PDEV), gamma coronavirus IBV along with turkey enteric coronavirus are susceptible to N-linked glycosylation which can be inhibited by the tunicamycin [15–17]. The SARS-CoV-2 S glycoprotein possesses 22 N-linked glycosylation sites as confirmed from the recently published literature [18–20]. SARS-CoV-2 S glycoprotein showed a conserved S2 subunit for N-linked glycosylation with a low tendency for O-linked glycosylation. The N-linked glycosylation in SARS-CoV-2 is featured by binding of GlcNAc with the Asp amino acid residue in the Asp-X-Ser/Thr consensus sequence in which the residue X is amino acid except for proline.

It is known that N-linked glycosylation is compulsory for understanding the location, structure and infectivity of the viruses and also their interaction with the host cells. These glycoproteins plays important role in immune responses but still need more research [21, 22]. The S glycoprotein represents unique pathogen-associated molecular patterns (PAMPs) that further recognized by the host pattern Recognition Receptors (PPRs). These PPRs includes Toll-like receptors 3, 4, 7, 8 and 9 along with C-type lectins and collectins [23, 24]. SARS-CoV is recognized by the toll-like receptors 3 and 4 via MyD88 and TRIF, furthermore, the same process may be proposed for the pathogenesis and infectivity in SARS-CoV2 [22, 25].

3. O-linked glycosylation

O-linked glycosylation involved in providing structural and functional stability to protein and believed to play important role in the maintenance of viral entity and biological activities associated with these viral proteins [26]. It has been demonstrated from a previously published study that Ser673, Thr678 and Ser686 are the conserved sites of O-linked glycosylation in human SARS-CoV-2 and other coronaviruses especially in S protein [26]. Moreover, O-glycosylation sites were predicted using the tool Net-O-Gly server 4.0 and found three sites for O-linked glycosylation at Ser673, Thr678 and Ser686 [26]. In another study, it was found that O-glycosylation at Thr 323 and Ser 325 and Thr 323 of the S1 glycoprotein are the possible and predicted sites of O-linked glycosylation in SARS-CoV-2 viral proteins [27].

The O-linked glycosylation at the Thr323 is confirmed by the presence of proline amino acids at position 322, making the possibility that the presence of proline amino acid is higher adjacent to the O-linked glycosylation sites [28]. Cryo-electron

microscopic images of the SAS-CoV2 indicate that the binding of S protein to the human angiotensin I-converting enzyme 2 (hACE2) receptor involves an association between receptor-binding domain (RBD) and the hACE2 peptidase domain [29, 30]. The RBD of the S protein in the S1 subunit endures hinge-like dynamic movement of accelerating the detention of RBD with hACE2, exhibiting a 10–20 fold increase in affinity for the hACE2 receptors [31, 32].

In another published study it was found that application of tunicamycin showed normal glycosylation of the M protein despite inhibiting the N-linked glycosylation of S protein [33]. In one of the study, the structures of the associated glycans to the M protein during the O-linked glycosylation showed that it added into two-step processes; GalNAc first added before the addition of galactose and the sialic acid. After the possession of the GalNAc, galactose and sialic acid sequentially, the M protein was further undergone modification in the trans-Golgi apparatus [34].

Recently it was reported that there are low levels of O-linked glycosylation in the S protein of SARS-CoV2 [35]. These glycans regulate the recognition of the antibodies and impinge on priming by the host proteases enzyme system. Mucin-type O-linked glycosylation is featured with the presence of GalNAc associated with the hydroxyl group of serine and threonine amino acid residues. Mucins contain a significant number of O-GalNAc glycans [36]. The presence of the O-linked glycans involved in the O-linked glycosylation of viral proteins suggests a vital role in biological activity. In the SARS-CoV-2 S1 protein, the O-linked glycosylation as GalNAc and O-GlcNAc appears to be involved in the structural and functional stability of the protein. The current scenario involves the use of a vaccine that utilizes the S protein glycosylation as a target.

4. Palmitoylation

Palmitoylation refers to the attachment of the palmitic fatty acid to the cysteine (S-palmitoylation) but less frequently to the serine and threonine amino acid residues (O-palmitoylation). In coronaviruses studies, the S protein undergone palmitoylation in infected cells and in the presence of tunicamycin it does not undergoes palmitoylation [37]. In another study conducted on MHV S protein reduces infectivity of MHV when treated with palmitoyl acyltransferase inhibitors 2-bromopalmitate [38]. The cytoplasmic part of the SARS-CoV S protein consists of four cysteine-rich clusters among them 2 clusters modified upon palmitoylation. However, cell surface expression of SARS-CoV S protein was unaffected due to this palmitoylation. In one of the previously published study, it was found that treatment of nitric oxide significantly leads to a reduction in the palmitoylation of the S protein of SARS-CoV [39].

In one of the study, it was found that there is three cysteines at position 40, 43 and 44 are found to undergo palmitoylation in the E protein of the SARS-CoV [40]. In another study, homologous cysteine of the E protein of MV-A59 at position C40, C44 and C47 were mutated to the alanine residues as resultant infectivity decreased [41]. It is therefore concluded that palmitoylation of the MHV E protein contributes to the stability and biological activity of the mature virions. Contrary, palmitoylation of the SARS CoV E protein is not mandatory for its interception with N protein.

5. Phosphorylation

In SARS-CoV2 the most abundant genomic protein encoded is the N protein with a significantly higher level of translation at the early stage of the infection. In

all form of the coronaviruses, the N-protein is almost the same and conserved containing two globular domains, the N- terminal domain (NTD) and the C-terminal domain (CTD). Around these domains, intrinsically disorganized regions are present. N protein is dimeric with multiple RNA binding sites including one major RNA-binding groove, which is created by the two CTD piling on each other on NTD [42]. In previously published literature it has been found that in the disorganized region, there is an abundance of serine-arginine residues that is essential for the essential function and regulation of the N-protein [42]. Cytoplasmic kinases mediate the phosphorylation of the N-protein in the early infection phase. N protein, a helical nucleocapsid, constituent of SARS-CoV2 virus arranged in beads on a string pattern which shows binding with the RNA. It has two domains namely the N-terminal domain and C-terminal domain. It has been found that both domains contribute to the binding of the viral genome. In one of the recently published study, it was demonstrated in truncation analysis that an L/Q-rich region placed within the intrinsically disordered region of the SARS-CoV-2 N protein plays a vital role in RNA-mediated phase separation, which is located adjacent to the phosphorylated SR-rich region (constituting residues 176–206) [43]. In the same study, it was concluded that N protein central intrinsically disordered region shown to be involved in protein–protein interactions mediated via putative hydrophobic α-helix spanning residues (213–225 residues) [43].

6. Prospects

The functional role of PTMs in SARS-CoV2 associated proteins has not been fully explored and many trials are needed for proposing the role of all types of glycosylation like N-linked and O-linked along with palmitoylation and phosphorylation in the initial phase of the infection. However multiple modification sites on the proteins of the SARS-CoV2 virus provides opportunities to explore more about the replication and pathogenesis of the virus into the host cells. Moreover, newer techniques for the detection of the PTMs are also needed to detect the modifications at multiple sites in dynamically changing virus structure. It is also needed in the current scenario to better understand the molecular mechanism of these PTMs. Also, the PTMs of the coronavirus proteins might be attractive targets for the therapeutic regime. PTMs of the coronavirus proteins might also provide a prospective target for the development of the vaccines.

Author details

Alok Raghav[1*], Renu Tomar[2] and Jamal Ahmad[3]

1 Multidisciplinary Research Unit, GSVM Medical College, Kanpur, India

2 Department of Public Health, Poornima University, Jaipur, India

3 Rajiv Gandhi Centre for Diabetes and Endocrinology, Aligarh Muslim University, Aligarh, India

*Address all correspondence to: alokalig@gmail.com

IntechOpen

References

[1] Masters PS. The molecular biology of coronaviruses. Advances in Virus Research. 2006;**66**:193-292

[2] de Groot RJ. Structure, function and evolution of the hemagglutinin-esterase proteins of corona-and toroviruses. Glycoconjugate Journal. 2006;**23**(1-2): 59-72

[3] Liu DX, Inglis SC. Association of the infectious bronchitis virus 3c protein with the virion envelope. Virology. 1991;**185**(2):911-917

[4] To J, Surya W, Fung TS, et al. Channel-inactivating mutations and their revertant mutants in the envelope protein of infectious bronchitis virus. Journal of Virology. 2017;**91**(5): e02158-e02116

[5] Huang Y, Yang C. Xu, Xf. et al. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. ActaPharmacol Sin. 2020;**41**: 1141-1149

[6] Xia S, Zhu Y, Liu M, Lan Q, Xu W, Wu Y, et al. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. Cellular & Molecular Immunology. 2020;**17**:765-767

[7] Niemann H, Boschek B, Evans D, Rosing M, Tamura T, Klenk HD. Post-translational glycosylation of coronavirus glycoprotein E1: inhibition by monensin. The EMBO Journal. 1982;**1**(12):1499-1504

[8] Delmas B, Laude H. Assembly of coronavirus spike protein into trimers and its role in epitope expression. Journal of Virology. 1990;**64**(11): 5367-5375

[9] Nal B, Chan C, Kien F, et al. Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. The Journal of General Virology. 2005;**86**(Pt 5):1423-1434

[10] Schmidt I, Skinner M, Siddell S. Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. The Journal of General Virology. 2010;**68**(Pt 1):47-56

[11] Smati R, Silim A, Guertin C, et al. Molecular characterization of three new avian infectious bronchitis virus (IBV) strains isolated in Quebec. Virus Genes. 2002;**25**(1):85-93

[12] Delmas B, Laude H. Carbohydrate-induced conformational changes strongly modulate the antigenicity of coronavirus TGEV glycoproteins S and M. Virus Research. 1991;**20**(2): 107-120

[13] Zheng J, Yamada Y, Fung TS, Huang M, Chia R, Liu DX. Identification of N-linked glycosylation sites in the spike protein and their functional impact on the replication and infectivity of coronavirus infectious bronchitis virus in cell culture. Virology. 2018;**513**:65-74

[14] McBride CE, Machamer CE. Palmitoylation of SARS-CoV S protein is necessary for partitioning into detergent-resistant membranes and cell-cell fusion but not interaction with M protein. Virology. 2010;**405**(1): 139-148

[15] Hogue BG, Nayak DP. Expression of the porcine transmissible gastroenteritis coronavirus M protein. Advances in Experimental Medicine and Biology. 1990;**276**:121-126

[16] Utiger A, Tobler K, Bridgen A, Ackermann M. Identification of the membrane protein of porcine epidemic

diarrhea virus. Virus Genes. 1995;**10**(2): 137-148

[17] Dea S, Verbeek AJ, Tijssen P. Antigenic and genomic relationships among turkey and bovine enteric coronaviruses. Journal of Virology. 1990;**64**(6):3112-3118

[18] Wang Q, Zhang Y, Wu L, et al. Structural and functional basis of SARS-CoV-2 entry using human ACE2. Cell. 2020;**181**:894-904.e9

[19] Wu D, Wu T, Liu Q, et al. The SARS-CoV-2 outbreak: what we know. International Journal of Infectious Diseases. 2020;**94**:44-48

[20] Yuan M, Wu NC, Zhu X, et al. A highly conserved cryptic epitope in the receptor-binding domains of SARS-CoV-2 and SARS-CoV. Science. 2020;**368**:630-633

[21] Hompson AJ, de Vries RP, Paulson JC. Virus recognition of glycans receptors. CurrOpinVirol. 2019;**34**: 117-129

[22] van Kooyk Y, Rabinovich GA. Protein-glycan interactions in the control of innate and adaptive immune responses. Nature Immunology. 2008;**9**:593-601

[23] Kawai T, Akira S. The role of pattern-recognition receptor in innate immunity: update on Toll-like receptor. Nature Immunology. 2010;**11**:373-384

[24] Alexopoulou L, Holt AC, Medzhitov R, et al. Recognition of double-stranded RNA and activation of NF-kappaB by toll-like receptor 3. Nature. 2001;**413**:732-738

[25] Li G, Fan Y, Lai Y, et al. Coronavirus infections and immune responses. Journal of Medical Virology. 2020;**92**:424-432

[26] Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nature Medicine. 2020;**26**:450-452

[27] Uslupehlivan M, Sener E. Glycoinformatics approach for identifying target positions to inhibit initial binding of SARS-CoV-2 S1 protein to the host cell. bioRxiv. 2020. DOI: 10.1101/2020.03.25.007898

[28] ThankaChristlet TH, Veluraja K. Database analysis of O-glycosylation sites in proteins. Biophysical Journal. 2001;**80**:952-960

[29] Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell. 2020;**180**:1-12

[30] Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020;**181**:1-10

[31] Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020;**367**: 1260-1263

[32] Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science. 2020;**367**:1444-1448

[33] Holmes KV, Doller EW, Sturman LS. Tunicamycin resistant glycosylation of coronavirus glycoprotein: demonstration of a novel type of viral glycoprotein. Virology. 1981;**115**(2): 334-344

[34] Locker JK, Griffiths G, Horzinek MC, Rottier PJ. O-glycosylation of the coronavirus M protein. Differential localization of sialyltransferases in N- and O-linked

glycosylation. The Journal of Biological Chemistry. 1992;**267**(20):14094-14101

[35] Zhao P, Praissman JL, Grant OC, et al. Virus-receptor interactions of glycosylated SARS-CoV-2 spike and human ACE2 receptor. Cell Host & Microbe. 2020;**24**:S1931-3128(20) 30457-1

[36] Bagdonaite I, Wandall HH. Global aspect of viral glycosylation. Glycobiology. 2018;**28**:443-467

[37] van Berlo MF, van den Brink WJ, Horzinek MC, van der Zeijst BA. Fatty acid acylation of viral proteins in murine hepatitis virus-infected cells. Brief Report. Archives of Virology. 1987;**95**(1-2):123-128

[38] Thorp EB, Boscarino JA, Logan HL, Goletz JT, Gallagher TM. Palmitoylations on murine coronavirus spike proteins are essential for virion assembly and infectivity. Journal of Virology. 2006;**80**(3):1280-1289

[39] Akerstrom S, Gunalan V, Keng CT, Tan Y-J, Mirazimi A. Dual effect of nitric oxide on SARS-CoV replication: viral RNA production ¨and palmitoylation of the S protein are affected. Virology. 2009;**395**(1):1-9

[40] Liao Y, Yuan Q, Torres J, Tam JP, Liu DX. Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein. Virology. 2006;**349**(2):264-275

[41] Boscarino JA, Logan HL, Lacny JJ, Gallagher TM. Envelope protein palmitoylations are crucial for murine coronavirus assembly. Journal of Virology. 2008;**82**(6):2989-2999

[42] Carlson CR et al. Phosphorylation Modulates Liquid-Liquid Phase Separation of The Sars-Cov-2 N Protein.

bioRxiv preprint2020. DOI: 10.1101/ 2020.06.28.176248

[43] Lu S, Ye Q, Singh D, et al. The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. Nature Communications. 2021;**12**:502