

Proteins Nsp12 and 13 of SARS-CoV-2 have mitochondrial recognition signal: A connection with cellular mitochondrial dysfunction and disease manifestation

Feroza Begum^{1,2}, Upasana Ray^{*1,2}

¹CSIR-Indian Institute of Chemical Biology, 4, Raja S.C., Mullick Road, Jadavpur, Kolkata-700032, West Bengal, India.

²Academy of Scientific and Innovative Research (AcSIR), Ghaziabad- 201002, India

Correspondence: ray.upasana@gmail.com

upasana.ray@iicb.res.in

Keywords: SARS-CoV-2, ARDS, non-structural proteins, mitochondria

Abstract

Mitochondria are classically termed as powerhouse of a mammalian cell. Most of the cellular chemical energy in the form of adenosine tri phosphate (ATP) is generated by mitochondria and dysregulation of mitochondrial functions thus can be potentially fatal of cellular homeostasis and health. Acute respiratory distress has been earlier linked to mitochondrial dysfunction. SARS-CoV-2 infection severity leads to acute respiratory distress syndrome (ARDS) and can be fatal. We tried to investigate possible connection between SARS-CoV-2, ARDS and mitochondria. Here, we report identification of SARS-CoV-2 non-structural proteins (particularly Nsp12 and 13) that have recognition sequence with respect to mitochondrial entry. We also report that these proteins can potentially shuttle between cytoplasm and mitochondria based on the localization signals and help in downstream maintenance of the virus. Their properties to use ATP for enzymatic activities may cause ATP scavenging allowing viral RNA functions in lieu of host cell health.

Introduction

Viruses have been long known to hijack host cell machinery to carry out successful completion of their life cycles within the host cells. Viral proteins contribute to modulation of functions of various host cell factors and signaling pathways to accomplish better acclimatization in the host environment. Viral proteins have also been often reported to translocation to multiple host cell organelles leading to modulation of organellar functions that gets translated into pathogenic outcomes.

In this article we have explored the subcellular localization of SARS-CoV2 proteins with a specific goal to identify those proteins that can potentially be linked to mitochondria. Mitochondria being the powerhouse of host cells, determine cellular homeostasis and other functions that ultimately

determine the cell health. Since, SARS-CoV-2 has been linked to respiratory disorders, pathogenic effects of SARS-CoV-2 proteins on mitochondrial functions might influence extent of pathogenesis of this virus.

Methods

For this work, sequences corresponding to 16 non-structural proteins, 4 structural proteins, 6 ORF proteins and 1 hypothetical protein (ORF9b) encoded by the genome of the severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1 (complete genome NCBI Reference Sequence: NC_045512.2) were used.

Multiple freely available *in silico* tools were used to study the localization pattern of the proteins. WoLF-PSORT (<https://wolfpsort.hgc.jp/>) was used to study the localization pattern of the viral proteins. As our study was focused on exploring the mitochondrial localization of the SARS-CoV2 proteins, tools like MitoFates [1], MitoPort [2], DeepMito [3], TargetP [4] and Predotar [5] were used to further validate the result.

Results and Discussion

We analyzed the non-structural proteins of SARS-CoV-2 for their localization to identify those which can potentially gain mitochondrial entry. We used multiple *in silico* tools to narrow down our selections.

As per MitoFates prediction tool, multiple proteins were shown to have TOM20 recognition motifs, amphipathic alpha helix and bear probable heterodimer mitochondrial processing peptidase (MPP) cleavage site in the sequence. MPP cleavage occurs in the mitochondrial matrix and is required to cleave the N terminal mitochondrial signal sequence before a protein can enter inside. There are reports on other corona viruses where experiments have shown that the viral proteins showing mitochondrial localization possess TOM20 recognition motifs in the N terminal region [6]. Hence, these factors are important to consider while assessing the localization of any viral protein in mitochondria. TOM20 recognition motifs, amphipathic helix and MPP cleavage sites were identified in 19 proteins (Nsp2, Nsp3, Nsp5, Nsp6, Nsp7, Nsp8, Nsp10, Nsp12, Nsp13, Nsp14, Nsp16, surface, envelope, membrane, orf3a, orf6, orf7a, orf8 and orf9b). But the probability of detecting mitochondrial presequence (target sequence) in the above mentioned proteins were very low.

Localization pattern was further studied using WoLF-PSORT tool where proteins like Nsp1, Nsp2, Nsp3, Nsp4, Nsp8, Nsp9, Nsp14, Nsp15, Nsp16 and orf9b were detected to be majorly cytoplasmic with the highest predicted number of nearest neighbors to the query which localize to cytoplasm of 17.5, 17, 13, 26, 22, 15, 20.5, 24, 18.5 and 22.5 respectively. Nsp5 (3C like Proteinase) and orf6 were detected to be cytoplasmic as the highest predicted number of nearest neighbors to the query which localize to cytoplasm were 11.5 and 11 respectively but both the proteins also showed the second highest score of 11 and 10 respectively for mitochondrial localization. Proteins like Nsp12

and Nsp13 were detected to be mitochondrial as the highest predicted number of nearest neighbors to the query which localize to mitochondria was 15 and 23.5 respectively. Also, Nsp13 was detected to be a shuttling protein between cytoplasm and mitochondria as the predicted number of nearest neighbors to the query which localize to both mitochondria and cytoplasm is 13.833. Other proteins like Nsp6, Nsp7, Nsp10, Envelope, orf7A, orf8 and orf10 were detected to be extracellular as their highest predicted number of nearest neighbors to the query which localize to Extracellular region were 13, 29, 28, 17, 27, 20 and 14 respectively. Nsp6 showed an equal score of 13 for both extracellular and plasma membrane localization. Nucleocapsid was shown to be nuclear with the highest score of 29 and structural proteins like surface glycoprotein and membrane showed plasma membrane localization with the highest score of 30 and 29 respectively. Also, ORF proteins like ORF3a and 7b showed plasma membrane localization with the highest score of 32 and 22 respectively. Hence, 2 proteins (Nsp12 and Nsp13) were predicted to show mitochondrial localization, out of which Nsp13 is shown to commute between cytoplasm and mitochondria. Nsp5 and ORF6 have a very close score for being cytoplasmic or mitochondria, therefore both can have a property to travel between the two sub cellular locations.

To further analyze the sequences, MITOPROT II tool was used where Nsp6, Nsp7, Nsp9, Nsp12, Nsp13, nucleocapsid, ORF3a and ORF10 were predicted to possess target sequence cleavage site and their respective cleaved sequence was also mentioned. Out of the, few proteins (Nsp6, Nsp12, Nsp13) showed a higher probability of export to mitochondria (0.8425, .5067, 0.7848 respectively) and other proteins like Nsp7, Nsp9, nucleocapsid, ORF 3a and ORF10 showed a lower probability of export to mitochondria (0.0338, 0.2221, 0.0433, 0.2027, 0.0151 respectively). If the cut off value is considered to be 0.5 then Nsp6, Nsp12 and Nsp13 can be considered to contain mitochondrial localization signal. No cleavage site was predicted in Nsp5 protein and showed a very low probability score of 0.0973 for mitochondrial export. Results obtained from the two tools (MitoFates and MITOPROT II) regarding few proteins like Nsp6, Nsp12 and Nsp13 indicates their possible mitochondrial localization. According to MitoFates tool, NSP6 has been shown to contain MPP cleavage site at 94th a.a. position (96th a.a. position w.r.t. MITOPROT II) and TOM20 recognition motif at 91-95 a.a. position. Nsp12 has been shown to possess MPP cleavage site at 34th a.a. position (21st a.a. position detected by MITOPROT II) with a TOM20 recognition motif at three different positions (8th-12th, 31st-35th, 45th-49th amino acids positions). For Nsp13, MPP cleavage site was detected at 23rd a.a. position (18th a.a. position detected as per MITOPROT II) and Tom20 recognition motifs were detected at two different positions (26th-30th, 31st-35tha.a positions). Taking together, these 3 proteins likely harbor mitochondrial localization signals as predicted by these two tools.

DeepMito tool was further used to identify the mitochondrial localization of the proteins and none of the proteins were predicted to be mitochondrial in nature except Nsp9. The program is designed only for mitochondrial proteins to check their sub mitochondrial localization but it can be used to check if an unknown protein belongs to the mitochondrial group. As per this software, Nsp9 was detected to be localized in the mitochondrial inner membrane with a score of 0.12. NSP9 was also

found to have an MPP cleavage site at 11th a.a. position by MitoFates tool and a cleavage site at 13th a.a. position as per MITOPROT II.

As per Predotar detection tool, none of the proteins were predicted to be located in Mitochondria with a very low probability score. The score for Nsp5, Nsp6, Nsp9, Nsp12, Nsp13 and ORF 6 are 0.04, 0.09, 0, 0.01, 0.1, 0 respectively. Studies done by Target P2.0 tool further indicates low likelihood of such proteins to be located in mitochondria with a score of 0.0006, 0.0001, 0.0584, 0.0368, 0.1878, 0.0308 respectively. Nsp13 is likely a shuttle protein between cytoplasm and mitochondria as it shows a high score for localization in other regions (possibly cytoplasm) of the cell (0.89 and 0.8003 for Predotar and Target P respectively) and the second highest score of mitochondrial confinement (0.1 and 0.1878 for Predotar and Target P respectively).

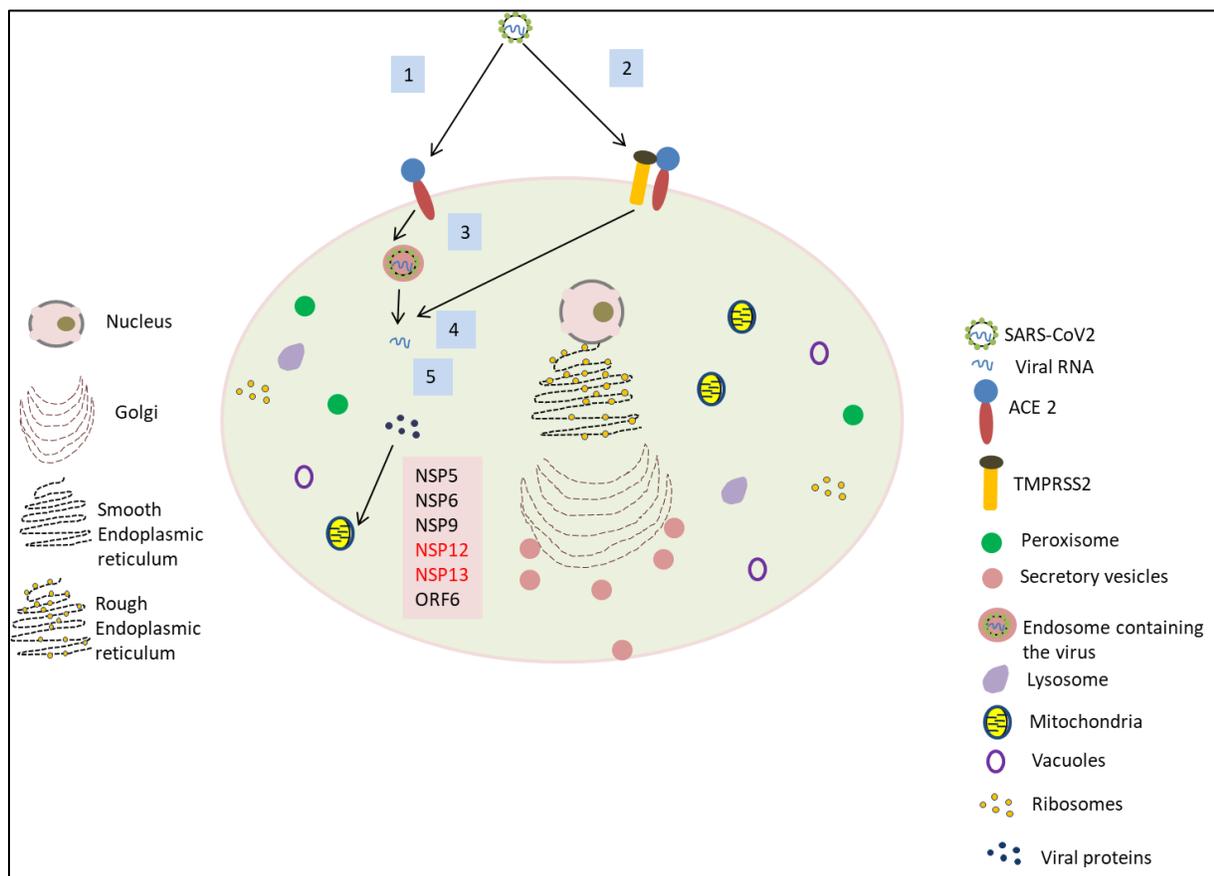


Figure1. Diagrammatic representation of the different proteins of SARS-CoV2 that may enter the mitochondria in the cell. After interaction with the receptor (ACE2) (1) the virus may enter the cell either indirectly via receptor mediated endocytosis or (2) directly via cell membrane fusion once its fusion peptide is exposed by TMPRSS2 cleavage. (3) If the virus enters via pathway 1 then it forms an endosome and (4) the viral genome is further released in the cytoplasm (5) The viral genome eventually gets translated generating viral proteins. The viral proteins further gets located to different subcellular localisations. The proteins mentioned in the figure likely enter the

mitochondria and the ones marked in red have a higher chance of mitochondrial localisation than others.

The study indicates that NSP13 has a high potential to translocate from cytoplasm to mitochondria as detected by 5 out of 6 prediction tools used. Nsp13 is the helicase enzyme and is one of the most essential protein for the virus for RNA replication as it helps in RNA unwinding during the replication. Another protein Nsp12, the RNA dependent RNA polymerase has also been detected to possess mitochondrial localization signal (detected by 3 out of 6 prediction tools) and should be studied in future to explore its possible role. Both the proteins are vital for the replication of the viral genome. A hypothetical protein, Orf9b a product of an alternate reading frame of ORF9 was shown to be localizing in mitochondria in case of SARS-CoV. The BLAST result between SARS-CoV Orf9b and SARS-CoV-2 Orf9b showed them to be 72.45% identical. The SARS-CoV ORF9b has been shown to regulate host system and hence it would be important to study the SARS-CoV2 Orf9b protein as well to check for its similar functions and possible associations with mitochondrial dysfunction.

The helicase and RdRp use ATP for their enzymatic activities. We also hypothesize that by collective scavenging of cellular ATP pool and modulating mitochondrial functions they can potentially lead to imbalance in cellular respiration and homeostasis thereby contributing to acute respiratory illness.

References

1. Fukasawa Y, Tsuji J, Fu SC, Tomii K, Horton P, Imai K. MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Mol Cell Proteomics*. 2015;14(4):1113-1126. doi:10.1074/mcp.M114.043083
2. Claros MG. MitoProt, a Macintosh application for studying mitochondrial proteins. *Comput Appl Biosci*. 1995;11(4):441-447. doi:10.1093/bioinformatics/11.4.441
3. Savojardo C, Bruciaferri N, Tartari G, Martelli PL, Casadio R. DeepMito: accurate prediction of protein sub-mitochondrial localization using convolutional neural networks. *Bioinformatics*. 2020;36(1):56-64. doi:10.1093/bioinformatics/btz512
4. Almagro Armenteros JJ, Salvatore M, Emanuelsson O, et al. Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance*. 2019;2(5):e201900429. Published 2019 Sep 30. doi:10.26508/lsa.201900429
5. Small I, Peeters N, Legeai F, Lurin C. Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics*. 2004;4(6):1581-1590. doi:10.1002/pmic.200300776
6. Acar DD, Stroobants VJE, Favoreel H, Saelens X, Nauwynck HJ. Identification of peptide domains involved in the subcellular localization of the feline coronavirus 3b protein. *J Gen Virol*. 2019;100(10):1417-1430. doi:10.1099/jgv.0.001321

