Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Short Communication

**In silico** exploration of enzymes involved in sialic acid biosynthesis and their possible role in SARS-CoV-2 infection

V.C. Divya a, Balasubramanian Saravanakarthikeyan b, *

**Article info**

Article history:
Received 9 June 2021
Received in revised form
15 August 2021
Accepted 16 August 2021
Available online xxx

**Abstract**

Salivary glands are considered important targets of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. Recent evidence suggests that along with angiotensin converting enzyme 2, certain cell surface sialic acids (Sia) may function as receptors for binding SARS-CoV-2 spike protein. Over 50 forms of Sia have been identified in nature, with N-acetylneuraminic acid (Neu5Ac) being the most abundant. We explored the Human Protein Atlas repository to analyze important enzymes in Neu5Ac biosynthesis and propose a hypothesis that further highlights the significance of salivary glands in coronavirus disease 19 (COVID-19). This work may facilitate research into targeted drug therapies for COVID-19.

© 2021 Japanese Association for Oral Biology. Published by Elsevier B.V. All rights reserved.

Salivary glands have been suggested as important targets of infection by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), causing coronavirus disease 19 (COVID-19). Angiotensin converting enzyme 2 (ACE2), the primary receptor for SARS-CoV-2 entry into the host cell, along with TMPRSS2 (a type II transmembrane serine protease) and furin, which are responsible for priming and activating the spike protein of SARS-CoV-2, respectively, are highly expressed in the salivary glands [1–3]. Recently, researchers have provided **in silico** evidence of a dual strategy, wherein in addition to ACE2, certain sialic acids (Sia) present on the cell surface may also function as potential receptors for binding the spike protein of SARS-CoV-2 [4–7].

Sia are monosaccharides found on the outermost ends of sugar chains of glycoproteins or glycolipids (glycoconjugates), which are present on the cell surface of vertebrates, higher invertebrates, and few bacteria and play significant roles in several physiological and pathological processes [8,9]. More than 50 forms of Sia have been identified in nature, of which the most abundant is N-acetylneuraminic acid (Neu5Ac). Sia and other host sugar molecules are often used as receptors by a wide range of viruses, including coronaviruses [10,11]. Recently, it has also been demonstrated that Neu5Ac exhibits affinity for the SARS-CoV-2 spike protein [12].

We explored the Human Protein Atlas (HPA) repository [13] to investigate the expression of genes related to Sia metabolism in human tissues, with an emphasis on three main enzymes involved in the biosynthesis of Neu5Ac, namely, glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase (GNE) enzyme, Neu5Ac 9-phosphate synthase (NANS) and Neu5Ac-9-phosphate phosphatase (NANP) [9]. The HPA repository is a freely available interactive resource that maps the human tissue proteome to analyze tissue profiles of specific protein classes in order to achieve its spatial localization down to the single-cell level [13–15], as depicted in Fig. 1. All the information in HPA is provided without any restrictions to allow researchers to get a holistic map of the human body.

**Abbreviations**: severe acute respiratory syndrome coronavirus 2, (SARS-CoV-2); coronavirus disease 19, (COVID-19); angiotensin converting enzyme 2, (ACE2); N-acetylneuraminic acid, (Neu5Ac); glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase, (GNE); Neu5Ac 9-phosphate synthase, (NANS); Neu5Ac-9-phosphate phosphatase, (NANP); Human Protein Atlas, (HPA); Genotype-Tissue Expression, (GTEX); Functional Annotation of Mammalian Genomes 5, (FANTOM5); cap analysis of gene expression, (CAGE).

**E-mail addresses**: skmdc2006@gmail.com, saravanb@srmist.edu.in

B. Saravanakarthikeyan.

* Corresponding author.

https://doi.org/10.1016/j.job.2021.08.004

1349-0079/© 2021 Japanese Association for Oral Biology. Published by Elsevier B.V. All rights reserved.

Please cite this article as: V.C. Divya and B. Saravanakarthikeyan, **In silico** exploration of enzymes involved in sialic acid biosynthesis and their possible role in SARS-CoV-2 infection, Journal of Oral Biosciences, https://doi.org/10.1016/j.job.2021.08.004
Interestingly, we observed that the GNE gene was also highly enhanced in the salivary glands compared to the lungs (5.5 NX) [13,15,16]. This is also in accordance with the findings in the individual datasets. The mRNA expression in the HPA data set revealed an 80.4 pTPM in the salivary glands and a 57.3 pTPM in the lungs. GNE protein expression was neither present in the alveolar cells (Table 1). These findings corroborate the observation that SARS-CoV-2 can engage in Sia found in human respiratory cells [13].

According to the HPA repository, the normalized mRNA expression levels of NANS were also reported to be higher in the salivary glands (80.0 NX) than in the lungs (17.1 NX), as revealed by HPA, GTEx, and FANTOM5 datasets [13,15,16]. This is also in accordance with the findings in the individual datasets. The mRNA expression in the HPA data set was 1.2 pTPM in the salivary glands and 4.4 pTPM in the lungs. The GNE data set revealed a 2.4 pTPM in the salivary glands and 2.7 pTPM in the lungs. The NANS data set revealed 1.7 and 7.1 scaled tags per million in the salivary glands and lungs, respectively. However, no data are available regarding NANS protein expression (Table 1).

The central dogma of molecular biology closely connects DNA, RNA, and protein molecules [13,15,16]. The nucleotide sequence determines the sequence of its mRNA product, and the mRNA sequence determines the amino acid sequence of the resulting polypeptide. The relationship between the concentration of a transcript and that of a protein derived from a particular locus is not trivial. Systematic studies at the genomic level that quantify transcripts and proteins revealed the significance of several processes beyond transcript concentration that influence the protein expression level. These processes include translation rates, translation rate modulation, protein half-life, protein synthesis delay, and protein transport. Thus, a direct comparison between protein and mRNA abundances from the same location or from the same cell type may not be appropriate. This could explain the variations between the protein expression and mRNA levels of several enzymes observed in the data sets in the present study. For instance, GNE protein expression was minimally detected in the glandular cells of the salivary glands despite its high mRNA expression. In addition, the protein in silico findings in the HPA are concerned with immunohistochemistry, which may not be the best tool for quantitative evaluation as it simply provides information pertaining to localization [2]. This may also partially explain why the data repositories yielded contradictory information that could be reconciled by missing data in one being present in the other.

Hence, we hypothesize that the high expression of enzymes such as GNE, NANS, and NAPN in the salivary glands, which play a pivotal role in the synthesis of Sia, needs to be considered. This correlation further implicates the significance of salivary glands as the precursor of Sia. Considering the consensus among several processes including translation, protein half-life, and protein synthesis delay, and protein transport. Thus, a direct comparison between protein and mRNA abundances from the same location or from the same cell type may not be appropriate. This could explain the variations between the protein expression and mRNA levels of several enzymes observed in the data sets in the present study. For instance, GNE protein expression was minimally detected in the glandular cells of the salivary glands despite its high mRNA expression. In addition, the protein in silico findings in the HPA are concerned with immunohistochemistry, which may not be the best tool for quantitative evaluation as it simply provides information pertaining to localization [2]. This may also partially explain why the data repositories yielded contradictory information that could be reconciled by missing data in one being present in the other.

Hence, we hypothesize that the high expression of enzymes such as GNE, NANS, and NAPN in the salivary glands, which play a pivotal role in the synthesis of Sia, needs to be considered. This correlation further implicates the significance of salivary glands as the precursor of Sia. Considering the consensus among several processes including translation, protein half-life, and protein synthesis delay, and protein transport. Thus, a direct comparison between protein and mRNA abundances from the same location or from the same cell type may not be appropriate. This could explain the variations between the protein expression and mRNA levels of several enzymes observed in the data sets in the present study. For instance, GNE protein expression was minimally detected in the glandular cells of the salivary glands despite its high mRNA expression. In addition, the protein in silico findings in the HPA are concerned with immunohistochemistry, which may not be the best tool for quantitative evaluation as it simply provides information pertaining to localization [2]. This may also partially explain why the data repositories yielded contradictory information that could be reconciled by missing data in one being present in the other.

Hence, we hypothesize that the high expression of enzymes such as GNE, NANS, and NAPN in the salivary glands, which play a pivotal role in the synthesis of Sia, needs to be considered. This correlation further implicates the significance of salivary glands as the precursor of Sia. Considering the consensus among several processes including translation, protein half-life, and protein synthesis delay, and protein transport. Thus, a direct comparison between protein and mRNA abundances from the same location or from the same cell type may not be appropriate. This could explain the variations between the protein expression and mRNA levels of several enzymes observed in the data sets in the present study. For instance, GNE protein expression was minimally detected in the glandular cells of the salivary glands despite its high mRNA expression. In addition, the protein in silico findings in the HPA are concerned with immunohistochemistry, which may not be the best tool for quantitative evaluation as it simply provides information pertaining to localization [2]. This may also partially explain why the data repositories yielded contradictory information that could be reconciled by missing data in one being present in the other.
an important target in COVID-19 infection. Hypothetically, this observation could contribute to obtaining greater insight into the SARS-CoV-2 disease process, which will aid in developing future interventions and research.

Ethical approval

Ethical approval is not required.

CRediT authorship contribution statement

V.C. Divya: Literature search, writing, reviewing, and editing. B. Saravanakarthikeyan: Literature search, table, writing, and reviewing.

Conflicts of interest

The authors declare that there are no competing interests.

Acknowledgments

The data used for the analyses described in this manuscript were obtained from the HPA repository and GTEx Portal on 21/07/2021 and/or dbGaP accession number phs000424.vN.pN on 21/07/2021. There are no sources of funding for the present study.

Table 1

| mRNA and protein expression of ACE2, GNE, NANS and NANP in salivary glands and lungs. |
|--------------------------------------|--------------------------------------|
|                                      | Salivary Glands                      | Lungs                          |
| Angiotensin converting enzyme 2 (ACE2) |                                      |                                |
| RNA                                  |                                      |                                |
| Consensus, NX                        | 1.1                                  | 0.8                            |
| HPA, pTPM                            | 0.5                                  | 1.7                            |
| GTEx, pTPM                           | 1.8                                  | 1.1                            |
| FANTOM5, scaled tags per million      | 0.3                                  | 2.8                            |
| Protein                              | Glandular cells: not detected        | Macrophages: not detected       |
|                                      | Alveolar cells: not detected         |                                |
| GNE (Glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase) |                                      |                                |
| RNA                                  |                                      |                                |
| Consensus, NX                        | 45.3                                 | 5.5                            |
| HPA, pTPM                            | 33.9                                 | 13.7                           |
| GTEx, pTPM                           | 17.8                                 | 9.3                            |
| FANTOM5, scaled tags per million      | 47.3                                 | 28.2                           |
| Protein                              | Glandular cells: low detection       | Macrophages: high detection     |
|                                      | Alveolar cells: not detected         |                                |
| Neu5Ac 9-phosphate synthase (NANS)   |                                      |                                |
| RNA                                  |                                      |                                |
| Consensus, NX                        | 80.0                                 | 17.1                           |
| HPA, pTPM                            | 80.4                                 | 57.3                           |
| GTEx, pTPM                           | 101.7                                | 68.8                           |
| FANTOM5, scaled tags per million      | 56.2                                 | 62.4                           |
| Protein                              | Glandular cells: high detection      | Macrophages: high detection     |
|                                      | Alveolar cells: not detected         |                                |
| Neu5Ac-9-phosphate phosphatase (NANP) |                                      |                                |
| RNA                                  |                                      |                                |
| Consensus, NX                        | 7.8                                  | 6.7                            |
| HPA, pTPM                            | 1.2                                  | 4.4                            |
| GTEx, pTPM                           | 2.4                                  | 2.7                            |
| FANTOM5, scaled tags per million      | 1.7                                  | 7.1                            |
| Protein                              | No data                              | No data                        |

HPA, human protein atlas; GTEx, genotype tissue expression; FANTOM5, functional annotation of mammalian genomes 5; NX, normalized expression; pTPM, 1 million transcripts per kilobase million. Protein data are interpreted as follows (according to The Human Protein Atlas): “Protein expression score is manually annotated on immunohistochemical figures based on the intensity of staining (negative, weak, moderate or strong) as well as fraction of stained cells (<25%, 25%–75% or >75%): negative-no detection; weak <25%- no detection; weak combined with 25%–75% or >75%- low detection”.

References

[1] Xu J, Li Y, Gan F, Du Y, Yao Y. Salivary glands: potential reservoirs for COVID-19 asymptomatic infection. J Dent Res 2020;99:989.
[2] Pascolo L, Zupin L, Melato M, Tricarico PM, Crovella S. TMPRSS2 and ACE2 Co-expression in SARS-CoV-2 salivary glands infection. J Dent Res 2020;99:1120 e1.
[3] Zupin L, Pascolo L, Crovella S. Is FURIN gene expression in salivary glands related to SARS-CoV-2 infectivity through saliva? J Clin Pathol 2021;74:209 e11.
[4] Morniroli D, Giannì ML, Consales A, Pietrasanta C, Mosca F. Human sialome and coronavirus disease-2019 (COVID-19) pandemic: an understated correlation? Front Immunol 2020;11:1480.
[5] Milanetti E, Miotto M, Di-Rienzo L, Monti M, Gosti G, Ruocco G. In-Silico evidence for two receptors based strategy of SARS-CoV-2. bioRxiv 2020.https://doi.org/10.1101/2020.03.24.006197.
[6] Kuchipudi SV, Nelli RK, Gontu A, Satyakumar R, Surendran NM, Subbiah M. Sialic acid receptors: the key to solving the enigma of zoonotic virus spillover. Viruses 2021;3:262.
[7] Sun XL. The role of cell surface sialic acids for SARS-CoV-2 infection. Glycobiology 2021 Apr 28: cwab032.https://doi.org/10.1093/glycob/cwab032. Epub ahead of print. PMID: 33909065.
[8] Varki A. Sialic acids in human health and disease. Trends Mol Med 2008;14:351 e60.
[9] Li Y, Chen X. Sialic acid metabolism and sialyl transferases: natural functions and applications. Appl Microbiol Biotechnol 2012;94:887–905.
[10] Wasik BR, Barnard KN, Parrish CR. Effects of sialic acid modifications on virus binding and infection. Trends Microbiol 2016 Dec;24(12):991–1001. https://doi.org/10.1016/j.tim.2016.07.005. Epub 2016 Aug 1. PMID: 27491885; PMCID: PMC5123965.
[11] Distinct roles for sialoside and protein receptors in coronavirus infection. mBio 2020;11(1). https://doi.org/10.1128/mBio.02764-19. e02764-19.
[12] Baker AN, Richards SJ, Guy CS, Tr Congdon, Hasan M, Zwetsloot AJ, et al. The SARS-COV-2 spike protein binds sialic acids and enables rapid detection in a lateral flow point of care diagnostic device. ACS Cent Sci 2020;6:2046–52.
[13] Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. Science 2015;347:1260419.

[14] Human protein atlas (HPA) available from: https://www.proteinatlas.org/about/assays+annotation [Last accessed online 21 July 2021].

[15] Genotype-tissue expression (GTEx) available from: https://www.gtexportal.org/home/documentationPage (Last accessed online 21 July 2021).

[16] Liu Y, Beyer A, Aebersold R. On the dependency of cellular protein levels on mRNA abundance. Cell 2016;165:535–50.