ALFERRED MOLECULAR PATHWAYS OBSERVED IN NASO-OROPHARYNGEAL SAMPLES OF SARS-CoV-2 PATIENTS

Emel Akgun¹,², Mete Bora Tuzuner², Betul Sahin², Meltem Kílercik¹,², Canan Kulah², Hacer Nur Cakroglu²,
Mustafa Serteser¹,², Ibrahim Unsal², Ahmet Tarik Baykal¹,²*¹

¹,² Department of Medical Biochemistry, Faculty of Medicine, Acibadem University, Istanbul, Turkey.
² Acibadem Labmed Clinical Laboratories, Istanbul, Turkey.

*To whom correspondence may be addressed: Assoc. Prof. Ahmet Tarik Baykal, PhD, Acibadem University,
Department of Medical Biochemistry, School of Medicine, Icerenkoy Mah. No:32, Kayisdagi Cad., 34752
Ataşehir, Istanbul, Turkey, Tel: +902165004828, Fax: +902165443940, E-mail: ahmet.baykal@acibadem.edu.tr

Abstract
Background: COVID-19 or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) appeared throughout the World and currently affected more than 3.6 million people and caused the death of around 252,000 people. The novel strain of the coronavirus disease is transmittable at a devastating rate with a high rate of severe hospitalization even more so for the elderly population. Currently around 50,000 patients are in a seriously critical situation. Although 1.2 million patients recovered from the disease there are still more than 2.1 Million active cases. Naso-oro-pharyngeal swab samples as the first step towards detecting suspected infection of SARS-CoV-2 provides a non-invasive method for PCR testing at a high confidence rate. Furthermore, proteomics analysis of PCR positive and negative naso-oropharyngeal samples provides information on the molecular level which highlights disease pathology.

Method: Samples from 15 PCR positive cases and 15 PCR negative cases were analyzed with nanoLC-MS/MS to identify the differentially expressed proteins.

Results: Proteomic analyses identified 207 proteins across the sample set and 17 of them were statistically significant. Protein-protein interaction analyses emphasized pathways like Neutrophil degranulation, Innate Immune System, Antimicrobial Peptides.

Conclusion: Neutrophil Elastase (ELANE), Azurocidin (AZU1), Myeloperoxidase (MPO), Myeloblastin (PRTN3), Cathepsin G (CTSG) and Transcobalamin-1 (TCN1) were found to be significantly altered in naso-oropharyngeal samples of SARS-CoV-2 patients. The identified proteins are linked to alteration in the innate immune system specifically via neutrophil degranulation and NETosis.

Keywords: SARS-CoV-2, Neutrophil Degranulation, Proteinase-3, Myeloperoxidase, ANCA, NETosis
1. Background

Severe acute respiratory distress syndrome-associated coronavirus-2 (SARS-CoV-2) also known as COVID-19 first appeared in Wuhan, Hubei, China in December 2019, and now it is widespread around the World. The replication of the virus is through the human airway epithelial cells where it targets the receptors of human Angiotensin-Converting Enzyme 2 (ACE-2). The high mortality observed in COVID-19 is associated with severe acute respiratory distress and systemic coagulopathy. Many COVID-19 patients show a postponed onset of respiratory issues but then develop into more severe situations. We wanted to investigate the molecular changes in the COVID-19 patients’ naso-oropharyngeal swab samples via comparison to the proteome of PCR negative cases. Our goal is to find pathways associated with the site of infection through proteomics analysis. 17 statistically significant protein alterations lead us specifically to neutrophil degranulation pathways.

During airway infections, Neutrophils are the first wave of defense that also defines the disease outcome. Neutrophils have multiple functions in viral infections such as inactivation of the virus, achieved by phagocytosis, ROS production, proteolytic enzyme release, NET activation (NETosis, Neutrophil Extracellular Traps). Neutrophils also interact with immune cells and secreting cytokines participate in eliciting an antiviral response.

Neutrophil granulocytes express several enzymes linked to controlling host infection. As the cytotoxic molecules are released from the granules they have an impact on the inflammatory response. Such adverse molecules cause severe damage to the host where it exhibits itself as perivascular infiltrates around the capillaries in the lungs as observed in SARS-CoV-2 patients. The identified up-regulated proteins Neutrophil Elastase, Azurocidin, Myeloperoxidase, Myeloblastin, Cathepsin G, and Transcobalamin-1 (TCN1, AZU1, CTSG,
PRTN3, MPO, ELANE) in naso-oropharyngeal swab samples are discussed to highlight the molecular mechanism changes in the site of infection.

2. Study design

2.1. Study Population and Sample Collection

Samples were collected using a nasal and oropharyngeal (NUCLISWAB) swab (Salubris, Turkey) and were placed in a tube with Universal Transport Media. After collection the samples were stored at -80 °C before proteomics analysis. The study group consists of 15 PCR positive and 15 negative cases. Informed consent was sought from each accepted to enter the study and ethical approval for the conduct of the study was given by Acibadem Mehmet Ali Aydinlar University Human Scientific and Ethical Review Committee (Approval ID: 2020-07/9).

2.2. COVID-19 RT-PCR Test

For molecular testing of SARS-CoV-2;

The extraction of nucleic acids from the samples was performed by a manual liquid phase method using Bio-Speedy Nucleic Acid Isolation Kit (Bioeksen, Turkey). Nucleic acid amplification test (NAT) was carried out by Bio-Speedy COVID-19 RT-qPCR Detection Kit (Bioeksen, Turkey) according to the manufacturer's instructions on RotorGene (Qiagen, Germany) Real-Time PCR instrument. The test briefly achieves a one-step reverse transcription (RT) and real-time PCR (qPCR) targeting SARS-CoV-2 specific RdRp (RNA-dependent RNA polymerase) gene fragment

2.3. LC-MSMS Analysis
A shorter sample preparation approach was applied to obtain the tryptic peptides for analyses. 50 ul of the naso-oropharyngeal transport solution was taken and lyophilized. The powder was reconstituted in 20ul of 50 mM Ammonium bicarbonate solution with 1 ul of protease inhibitor mixture (Thermo Scientific). The mixture was centrifuged at 14000 xg for 10 min and the supernatant was transferred to a clean Eppendorf tube. DTT was added to 10 mM final concentration and incubated at 60°C for 10 min. The mixture was alkylated in dark for 30 min with 20mM IAA. To the resulting mixture 1 ug of sequencing grade trypsin (Promega Gold) was added and incubated at 55°C for 1.5 hrs. The digest was acidified with 1 ul formic acid and transferred to an LC vial for injection. The samples were analyzed by the protocols in our previous studies [1]. Briefly, tryptic peptides were trapped on a Symmetry C18 (5µm,180µm i.d. × 20 mm) column and eluted with ACN gradient (4% to 40% ACN, 0.3 ul/min flow rate) with a total run time of 60 min on a CSH C18 (1.7 µm, 75 µm i.d. × 250 mm) analytical nano column. Data were collected in positive ion sensitivity mode using a novel data-independent acquisition mode (DIA) coined as SONAR [2] with a quadrupole transmission width of 24 Da. Progenesis-QIP (V.2.4 Waters) was used for data processing.

2.4. PPI Network and Pathway Analysis

Protein-protein interaction networks functional enrichment analysis was carried out using the STRING (http://string-db.org/, v11.0) database with the highest confidence interaction score level to identify possible pathways related to the identified proteins. Textmining, experiments, databases co-expression, neighborhood, gene fusion, and co-occurrence selected as active interaction sources. The minimum interaction score was set to high (high confidence=0.700). REACTOME (http://www.reactome.org) pathways analysis tool was also used for processing.

3. Results

3.1. Label-free Proteomics
Proteins were extracted from a small amount of naso-oropharyngeal samples, fast tryptic digestion was applied and followed with a 60 min reverse-phase separation. NanoLC-MSMS analysis provided the identification of 207 protein groups with high confidence (<1% FDR). Statistical analysis done in Progenesis QIP software identified 17 proteins to be statistically significantly expressed in patients’ naso-oropharyngeal samples (Table 1).

3.2. Bioinformatic Analysis

Pathway analysis of the significantly altered protein levels between COVID-19 positive and negative patients’ naso-oropharyngeal swab samples were analyzed using the STRING online database. The PPI network obtained, contained 17 differentially expressed proteins with 15 nodes (disconnected nodes were not shown) and 14 edges as shown in Figure 1A. The main cluster includes Neutrophil Elastase (ELANE), Azurocidin (AZU1), Myeloperoxidase (MPO), Myeloblastin (PRTN3), Cathepsin G (CTSG) and Transcobalamine-1 (TCN1). The abundance of these proteins were found to be increased in COVID-19 positive patient samples compared to negative ones. REACTOME pathway enrichment analyses of the differentially expressed proteins were performed. Nine of the proteins were primarily associated with the immune system pathway. Proteins clustered in the PPI network were mainly enriched in the neutrophile degranulation pathway (HSA-6798695, FDR=2.01E-7) which is a subpathway of the innate immune system (Figure 1B).

4. Discussion

Neutrophils play a vital role in airway infections and may also define the disease outcome and carry out various functions in viral infections ranging from phagocytosis, degranulation to the generation of neutrophil extracellular traps (NETs) [3]. Neutrophils also interact with the immune cells and secreting cytokines participate in eliciting the antiviral response.
Degranulation mechanism is part of the innate immune system which is necessary for the
fight against infection and provides the neutrophils with the necessary tools. In COVID-19
patients, it was argued that the interlink between the innate and adaptive immune system may
increase the severity and mortality of the disease [4]. Granules in neutrophils carry
synthesized antimicrobial molecules and the content differs depending on the stage of
maturation [5], depending on their contents they are named as azurophilic, gelatinase
granules, specific, and secretory vesicles [6].

In SARS-CoV-2 patients’ naso-oropharyngeal samples, we have identified azurophilic
granule (AG) proteins like Myeloperoxidase (MPO), elastase (ELANE), cathepsin G (CTSG),
azurocidin 1 (AZU1) and proteinase 3 (PRTN3) to be highly overexpressed. Azurophilic
neutrophils undergo limited exocytosis when activated [7] and their primary role is believed
to be killing and degradation of engulfed microbes in the phagolysosome [8]. Although
neutrophil degranulation having a positive impact on clearance of SARS-CoV-2 from the
nasopharynx, in the inflammation microenvironment the process is broken and may lead to
tissue damage by the secreted molecules from the granules. naso-oropharyngeal samples of
COVID-19 patients demonstrate an increased level of both PRTN3 and MPO. PRTN3, a
serine protease enzyme that is involved in granulocyte differentiation is expressed in the
neutrophil granulocytes [9]. Promyelocyte proliferation is inhibited by PR3 expression that
leads to neutrophil differentiation [10]. PRTN3 has various catalytic functions like cleavage
of extracellular matrix and antiinflammatory proteins and activation of cytokines. MPO which
is also expressed in neutrophil granulocytes can generate chlorine oxidants to ill infection
agents [11]. Anti-neutrophil cytoplasmic antibodies (ANCA) autoantibodies specifically
target azurophil proteins like PRTN3 and MPO and their role in numerous autoimmune
disorders are discussed but primary focuses is systemic vasculitis where blood vessels are
destroyed via inflammation [10]. ANCA with cytoplasmic staining pattern against PRTN3 (c-ANCA) is seen in granulomatosis with polyangiitis (GPA) whereas perinuclear staining pattern against MPO (p-ANCA) is seen in eosinophilic granulomatosis with polyangiitis (EGPA), microscopic polyangiitis (MPA), and pauci-immune idiopathic crescentic glomerulonephritis [12]. Vascular endothelium in GPA patients are damaged by neutrophils that actively participate in autoimmune response [13]. Vasculitis risk in patients increases with PRTN3 positive neutrophils [14]. PRTN3 is expressed on the surface of the neutrophils and the ANCAs that target PRTN3 causes pro-inflammatory effects [9]. The increased PRTN3 and MPO production in the neutrophils of COVID-19 patients are the target for anti-neutrophil cytoplasmic antibodies. Increased PRTN3 was shown to have a negative effect on the resolution of inflammation that causes immune system deregulation [15]. In a mouse study, ANCAs that target MPO were injected through IV and it triggered extra capillary glomerulonephritis [16]. Neutrophil apoptosis is necessary in the resolution of inflammation [17]. The resolution of inflammation and tissue repair processes are aided by the cytokines released from macrophages during phagocytosis of the neutrophils [10]. The dysregulation in the macrophage orchestrated phagocytosis and neutrophil apoptosis mechanisms leads to inflammation [18]. Patients with ANCA originated vasculitis are more prone to apoptosis as a result of ANCA activation [19]. PRTN3 is reported as a signal that sustains the inflammatory microenvironment [9]. The overexpression of PRTN3 observed in COVID-19 patients may impair the resolution of inflammation during neutrophil apoptosis. It is argued that the recruited neutrophils to the site of inflammation get activated and the release of uncontrolled mediators may cause tissue damage. The inflammation seen in vasculitis is converted to necrotizing vasculitis via the release of toxic mediators due to early neutrophil apoptosis and also by decreased level of apoptotic neutrophil clearance by macrophages [10]. PRTN3 was
also reported to interfere with complement protein C1-q mediated clearance of Apoptotic Cells [20].

NETosis is a type of programmed cell death where neutrophil extracellular traps are formed [21]. NETs were identified in 2004 and they are often overlooked as drivers of severe pathogenic inflammation [22]. NETs have pathogen killing properties and include strands of DNA wrapped with histones and are enriched with neutrophil proteins like MPO and ELANE [23]. PRTN3, ELANE, and CTSG are coined as neutrophil serine proteases (NSP) and have tissue degrading and microbial killing effects [24]. NSP’s active role in neutrophil-associated lung inflammatory and tissue-destructive diseases has been reported [25]. Increased expressions of PRTN3, ELANE, and CTSG was reported in COPD patients [26]. In a mouse model study ELANE or PRTN3 was introduced to the trachea which caused tissue destruction and enlargement in airspace [27]. ELANA expression was also implicated in the impairment of host defense resulting in a decrease in mucociliary clearance of bacteria and also pathogens phagocytosis [28]. CD2, CD4, and CD8 can be cleaved on the surface of T-cells by ELANE and CTSG that dysregulate T-cell Function [29]. It was reported that patients that lack alpha-1 antitrypsin (α1-Pi) which is the physiological inhibitor of PRTN3 and ELANE carries a high risk of developing emphysema [30]. Azurophil granules also carry Cathepsin D which was reported to be up-regulated in the pulmonary macrophages in a mouse model of cigarette smoking [31]. Severe COPD cases exhibited MPO positive cells as a signal of neutrophil activation [32]. On the other hand in a mouse model of influenza it was shown that the inflammation damage was reduced by the absence of MPO [33]. IL-32, a proinflammatory cytokine with four isoforms, cleavage by PRTN3 propagates cytokine activity and triggers IL-1beta, TNF-alpha, IL-6, and chemokines [34]. It was argued that the targeted inhibition of PR3 or silencing of IL-32 by an inactive form of PRTN3 may halt the IL-32 driven immune dysregulation [34]. NSPs like ELANE, CTSG, and PRTN3 can cleave pro-IL-1beta to
bioactive IL-1beta [35]. Caspase-1/Interleukin-1 converting enzyme (ICE) cleaves proteins
like precursors of the inflammatory cytokines interleukin 1β and interleukin 18 into their mature biologically active forms [36] and CTSG regulates Caspase-1 in this pathway [37]. ICE has an active role in cell immunity as an initiator of inflammatory response so once activated it triggers the formation of active IL-1beta which is secreted from the cell that induces inflammation in the neighboring cells [38]. It was recently reported that infection in COVID-19 patients with acute respiratory syndrome showed release of the pro-inflammatory cytokines like IL-1beta and IL-6 [39]. Through the available literature, we can see that the up-regulation of various proteins observed in the naso-oropharyngeal swab samples of COVID-19 patients is tightly interconnected with the immune response.

5. Conclusions

The alterations of various proteins in SARS-CoV-2 infected patients’ naso-oropharyngeal samples depict the molecular changes that govern the host antiviral defense system. The available literature for many respiratory diseases are very helpful in linking altered protein expressions to viral pathogenesis. Obtained data provided us an important view of SARS-CoV-2 molecular changes on the protein level in the infection site. Statistically significant protein alterations of PRTN3, MPO, ELANE, CTSG, AZU1, and TCN1 dysregulation is important in the early phases of infection and may be targets for anti-SARS-CoV-2 therapeutics. Further research may show a link between the level of these proteins with disease severity and may be used as prognostic markers. Modulating the dysregulated proteins like PRTN3 or MPO may promote an anti-inflammatory response to alleviate SARS-CoV-2 symptoms. Also targeting NETs to dampen the out-of-control host response as a treatment may increase the survival rate by reducing the number of patients who require mechanical ventilation in ICU. We posit here that excess NETs may elicit the severe multi-organ
consequences of COVID-19 via their known effects on tissues and the immune, vascular, and coagulation systems. Targeting NSPs and NETs in COVID-19 patients should therefore be considered by the biomedical community.

**Funding**

No specific funding was received from profit or non-profit organizations.

**Credit authorship contribution statement**

All authors have contributed to the study design and execution of the methodology. Sample collection, data acquisition, and data processing was done by Ahmet Tarik Baykal, Mete Bora Tuzuner, Emel Akgun, Betul Sahin, Mustafa Serteser, Ibrahim Unsal, Meltem Kilercik, Canan Kulah and Hacer Nur Cakiroglu. All authors have contributed and approved the manuscript.

**Conflict of Interest/Disclosure Statement**

The authors have declared no conflict of interest.

**Abbreviations**

CE, collision energy; DIA, data-independent acquisition; FA, formic acid; UPLC, Ultra performance liquid chromatography; UPX, Universal protein extraction kit.

**Appendix A. Supplementary data**

Supplementary Table 1. Protein quantification data

Supplementary Table 2. Protein identification data

**References**


[23] Yanbao Yu, Keehwan Kwon, Tamara Tshitrin, Shiferaw Bekele, Patricia Sikorski, Karen E. Nelson, and


<table>
<thead>
<tr>
<th>ACCESSION</th>
<th>UNIQUE</th>
<th>P value</th>
<th>FOLD CHANGE (POS/NEG)</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q9BQE3</td>
<td>2</td>
<td>0.0234</td>
<td>0.41</td>
<td>Tubulin alpha-1C chain</td>
</tr>
<tr>
<td>P20160</td>
<td>4</td>
<td>0.0413</td>
<td>2.03</td>
<td>Azurocidin</td>
</tr>
<tr>
<td>P01876</td>
<td>12</td>
<td>0.0073</td>
<td>2.04</td>
<td>Immunoglobulin heavy constant alpha 1</td>
</tr>
<tr>
<td>P29401</td>
<td>2</td>
<td>0.0337</td>
<td>2.10</td>
<td>Transketolase</td>
</tr>
<tr>
<td>P09104</td>
<td>2</td>
<td>0.0075</td>
<td>2.11</td>
<td>Gamma-enolase</td>
</tr>
<tr>
<td>P20061</td>
<td>3</td>
<td>0.0294</td>
<td>2.48</td>
<td>Transcobalamin-1</td>
</tr>
<tr>
<td>P04004</td>
<td>2</td>
<td>0.0001</td>
<td>2.65</td>
<td>Vitronectin</td>
</tr>
<tr>
<td>P02790</td>
<td>7</td>
<td>0.0000</td>
<td>2.71</td>
<td>Hemopexin</td>
</tr>
<tr>
<td>P0DOX7</td>
<td>3</td>
<td>0.0127</td>
<td>2.83</td>
<td>Immunoglobulin kappa light chain</td>
</tr>
<tr>
<td>P08246</td>
<td>2</td>
<td>0.0035</td>
<td>2.91</td>
<td>Neutrophil elastase</td>
</tr>
<tr>
<td>Q9UKL4</td>
<td>2</td>
<td>0.0011</td>
<td>3.18</td>
<td>Gap junction delta-2 protein</td>
</tr>
<tr>
<td>Q16695</td>
<td>2</td>
<td>0.0003</td>
<td>3.20</td>
<td>Histone H3.1t</td>
</tr>
<tr>
<td>P01871</td>
<td>2</td>
<td>0.0003</td>
<td>3.44</td>
<td>Immunoglobulin heavy constant mu</td>
</tr>
<tr>
<td>P08311</td>
<td>3</td>
<td>0.0097</td>
<td>3.67</td>
<td>Cathepsin G</td>
</tr>
</tbody>
</table>
Table 1. Significantly altered protein identification list.

<table>
<thead>
<tr>
<th>Protein ID</th>
<th>Count</th>
<th>P-value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P05164</td>
<td>9</td>
<td>0.0050</td>
<td>Myeloperoxidase</td>
</tr>
<tr>
<td>P00450</td>
<td>4</td>
<td>0.0020</td>
<td>Ceruloplasmin</td>
</tr>
<tr>
<td>P24158</td>
<td>2</td>
<td>0.0023</td>
<td>Myeloblastin</td>
</tr>
</tbody>
</table>

Figure 1. STRING and Reactome analysis of identified proteins. A- Protein-protein interaction network. B- Diagram of Initiate Immune System pathway (HSA-168249, FDR=2.3E-5) from Reactome, showing a significant enrichment for neutrophile degranulation pathway (HSA-6798695, FDR=2.01E-7) having the most hits.