Secreted ORF8 is a pathogenic cause of severe Covid-19 and potentially targetable with select NLRP3 inhibitors

Xiaosheng Wu[1]*, Michelle K. Manske[1], Gordon Ruan[1], Kevin E. Nowakowski[1], Jithma P. Abeykoon[1], Xinyi Tang[1], Yue Yu[2], Taylor L. Witter[3], Vanessa Taupin[5], Jonas Paludo[1], Stephen M Ansell[1], Andrew D. Badley[4], Matthew J. Schellenberg[3], Thomas E. Witzig[1]*

[1]Division of Hematology, Department of Medicine, [2] Department of Quantitative Health Sciences, [3]Department of Biochemistry and Molecular Biology, [4]Division of infectious diseases, Department of Medicine, Mayo Clinic, Rochester, MN, USA; [5]Electron Microscopy Core, University of California San Diego, La Jolla, CA, USA

Correspondence:

Dr. Xiaosheng Wu (wu.xiaosheng@mayo.edu) or

Dr. Thomas E. Witzig (Thomas.witzig@mayo.edu)

Abstract: 134 words

Main text: 3005 words (methods excluded)

Main Figure: 5

Suppl Figure: 5

Main Table: 0

Suppl Table: 3

Reference: 31

ABSTRACT

Despite extensive research, the specific factor associated with SARS-CoV-2 infection that mediates the life-threatening inflammatory cytokine response in patients with severe Covid-19 remains unidentified. Herein we demonstrate that the virus-encoded Open Reading Frame 8 (ORF8) protein is abundantly secreted as a glycoprotein *in vitro* and in patients with newly diagnosed Covid-19. ORF8 specifically binds to the NOD-like receptor family pyrin domain-containing 3 (NLRP3) in $CD14^+/CD16^+$ monocytes to induce an inflammasomal cytokine response. The levels of ORF8 protein in the blood correlate with disease mortality in patients with acute infection, and the disease trajectory in patients with severe Covid-19. Furthermore, *in vitro* the ORF8-induced inflammasome response can be readily inhibited by the select NLRP3 inhibitor MCC950. Our results identify the pathogenic cause and mechanism of severe disease, and a potential new treatment of severe Covid-19.

INTRODUCTION

The global pandemic of Covid-19 caused by infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has affected over 260 million people and caused over 5.2 million deaths worldwide.[1] Numerous studies have shown that the production of pro-inflammatory cytokines including IL1b, IL6, IL8, IL18, and CCL2, is responsible for the life-threatening symptoms, and the levels of these cytokines correlate with disease severity.[2,3] It has also been demonstrated that the virus primarily targets nasal and lung epithelial cells expressing ACE2 and TMPRSS2 receptors,[4,5] while the proinflammatory cytokine responses are mediated by innate immune cells. However, it remains unidentified what intermediate factor from the initial localized infection induces the systemic exaggerated release of proinflammatory cytokines. Studies have shown that the viral load is tightly associated with cytokine levels and clinical outcomes,[6-8] and treatment with virus-lowering neutralizing antibodies [9,10] could readily reduce the pro-inflammatory cytokine production, mitigate symptoms, and improve clinical outcomes, suggesting this intermediate factor is likely a direct product of the viral replication.

Upon infection of human cells, the SARS-CoV-2 virus replicates its 29.9kb RNA genome and produces up to 29 possible viral proteins including 16 non-structural proteins (NSP1-16), four structural proteins - Spike (SPK), Membrane (MEM), Envelope (ENV) and Nucleocapsid (NUC) and nine accessory proteins (ORF3A, 3B, 6, 7A, 7B, 8, 9b, 9c, and 10). While only four structural proteins along with the RNA genome are assembled into the new viral particles, the other proteins are thought to be left behind after completing their functions in the viral replication, assembly, and release.[11] While many of the viral proteins may interact with host cell machineries to disrupt host cell

function,[12,13] we hypothesized that one or more of these proteins is secreted to the bloodstream and is capable of inducing the systematic inflammatory cytokine response. Herein, we demonstrate that SARS-CoV-2 encoded ORF8 is a glycoprotein that is abundantly secreted to culture supernatant *in vitro* and to blood circulation in patients with Covid-19. Glycosylated ORF8 stimulates $CD14^+/CD16^+$ monocytes to produce Covid-19 specific pro-inflammatory cytokines through an NLRP3-mediated inflammasome response, which is readily inhibited by select NLRP3 inhibitors.

RESULTS

Secretion of SARS-CoV-2 proteins from human cells

To determine the secretion potential of SARS-CoV-2 proteins from human cells, we transduced HEK293 cells with one of 22 lentiviral constructs available to us each expressing a SARS-CoV-2 protein (Table S1). The secreted viral proteins in the culture supernatants were pulled down with Strep-Tactin beads or anti-HA beads (for SPK protein), analyzed by Western blotting, and the levels of secreted proteins compared to their expression in whole cell lysates. As shown in Fig 1A most of the viral proteins were robustly expressed with expected molecular sizes except for NSP13 (not detectable in our system), ORF10 (38 aa, too small to be detected), MEM (forms large aggregates), and NSP4 (insoluble, detectable by immunofluorescence only, Fig. S1). While NSP2, 5, 7, 9, 10, 12, 14, 15, ORF8 and NUC proteins were all detected as secreted proteins, ORF8 was the single most robustly secreted protein (Fig 1B). This is consistent with our *in silico* analysis showing that all these secreted NSPs each possess an unconventional protein secretion signal (UPS) while ORF8 carries a classic protein signal sequence (SS) and a N-link glycosylation sites (^{78}NYTV) (Fig. S2). These suggest that ORF8 is glycosylated and secreted through the classic ER/Golgi protein secretion pathway. The up-shifting and smearing of ORF8 band on Western blot (Fig. 1B) is likely due to its glycosylation.

Secreted ORF8 protein induces the production of Covid-19 specific cytokines

We then queried if any of the secreted SARS-CoV-2 proteins from HEK293 cells are capable of inducing the expression of pro-inflammatory cytokines relevant to Covid-19. To that end, we treated human peripheral blood mononuclear cells (PBMCs) from healthy donors *in vitro* with conditioned media (CM) containing major secreted

NSP proteins or secreted ORF8 followed by assessment of cytokine mRNA expression in the treated cells using quantitative reverse transcription-polymerase chain reaction (qRT-PCR). We found none of the secreted NSP proteins had any cytokine induction activity (Fig. 2A) while ORF8 containing CM (ORF8-CM) specifically induced the mRNA expression of IL1b, IL6, IL8, and CCL2 up to 5-fold in PBMCs from select donors (Fig. 2B). Since IL1b, IL6, IL8, and CCL2 are among the key cytokines specifically elevated in patients with severe Covid-19,[14] our data suggest secreted ORF8 from locally infected cells is likely to be the specific viral factor responsible for systemic inflammatory cytokine response in patients with severe Covid-19.

We then asked if the glycosylation of ORF8 would impact its cytokine induction potential. Therefore, we compared the cytokine induction activities of unglycosylated recombinant ORF8 (ORF8-glyco^{null}) purified from *E. coli* and highly glycosylated ORF8 (ORF8-glyco^{hi}) purified from HEK293 cell supernatant. As shown in Fig 2C, Covid-19 related cytokine IL1b, IL-6, and IL-8 but not IL-2 were induced by ORF8-glyco^{hi} in a dose-dependent manner while ORF8-glyco^{null} showed no such activity whatsoever regardless the doses used, suggesting that ORF8 glycosylation is indeed necessary for the cytokine induction.

In a complimentary experiment, we also found that treatment of ORF8 producing HEK293 cells with the Golgi inhibitor Brefeldin A and Monensin (BFA/M) resulted in reduced (but not absent) glycosylation in secreted ORF8 (Supplemental Fig. 3), coinciding with enhanced cytokine induction activity of glycosylated ORF8. These data suggest that the cytokine induction activity of ORF8 is tied to the level and possibly the type of its glycosylation that is governed by the functionality of ER/Golgi. Since several SARS-CoV2 proteins (MEM, ENV, ORF3A, ORF7A) are localized to

ER/Golgi membranes and induce the formation of double membrane vesicles (DMV)-like structures resembling viral replication vesicles (Supplemental Fig. 4B). We hypothesize that the co-expression of these viral membrane targeting proteins alters ORF8 glycosylation and hence the cytokine induction activity of ORF8. Indeed, co-expression of one or more of these ER/Golgi targeting proteins with ORF8 in HEK293 cells did result in enhanced cytokine induction activities of ORF8 (Supplemental Fig. 4D, 4E) suggesting they can modulate the ability of ORF8 to induce cytokine during SARS-CoV2 infection.

Secreted ORF8 stimulates CD14+/CD16+ monocytes to produce Covid-19 specific cytokines.

Having identified that SARS-Cov2 factor ORF8 is capable of inducing proinflammatory cytokines, we then asked which blood cell subset is the primary target of ORF8. To that end, we performed single cell RNA sequencing (scRNA-Seq) analysis on PBMCs treated with ORF8-CM or control-CM. Single cell transcriptome-based cell clustering showed that cells in cluster 5 on the t-SNE plot (Fig 3A) expressing intermediate monocyte subset markers CD14, CD16, CD68, and HLA-DR were specifically enriched (Fig 3B).[15,16] More intriguingly, mRNAs of cytokines IL1b, IL8, and CCL2 were among the most induced genes expressed in these cells (Fig 3C) (sTable 3). Our data suggest that the intermediate monocyte ($CD14^+/CD16^+$) population in PBMCs is the primary target of ORF8 and the main producer of Covid-19 related cytokines.

To validate this finding, we stimulated purified $CD14^+/CD16^+$ alongside with purified $CD4^+$, $CD8^+$, $CD14^+/CD16^-$ subsets as well as the total PBMCs with ORF8-CM. We found that significant increases of IL1b, IL8, and CCL2 mRNAs were only detected in the intermediate monocyte subset as well as the total PBMCs (Fig 3D). This further

demonstrates that the intermediate monocyte subset is indeed the primary producer of pro-inflammatory cytokines in response to ORF8 stimulation.[16]

Secreted ORF8 directly binds to and activate NLRP3 mediated inflammasome response

To decipher the potential mechanism by which ORF8 induces $CD14^+/CD16^+$ monocytes to produce pro-inflammatory cytokines, we performed pathway enrichment analysis on cells in cluster 5 using the KEGG pathway analysis tool. The results showed that SAR-CoV-2 infection pathway (HSA05171) was the top ranked enriched pathway. Specifically, the components of the NOD-like receptor signaling sub-pathway (HSA04621) were significantly enriched in those cells. These components include the lysosomal peptidoglycan cleaving enzyme LYZ, lysosomal cathepsin family proteases (CTSD, CTSB, CTSS, CTSL, CTSH), pathogen associated molecular pattern (PAMP) recognition receptor NLPR3, cytokine precursors pro-IL1b and pro-IL18, as well as the cleaving enzyme Caspase-1 (CASP-1) were among the top expressed mRNAs identified (Supplemental Table 3, & Fig 4A). These results indicate that the NLRP3 mediated inflammasome response was activated in those monocytes upon encountering glycoprotein ORF8.

We then asked if ORF8 glycoprotein enters monocytes through a receptor-mediated internalization or a non-receptor-mediated process to initiate NLRP3-mediated inflammasome activation. To differentiate these two possibilities, we preincubated PBMCs with pure ORF8 protein on ice for 2 hours to saturate any potential ORF8 binding receptors on the cell surface followed by an extensive wash to remove unbound ligand ORF8. The cells were then incubated at 37°C to initiate cell activation in the presence or absence of additional ORF8. As shown in Fig 4B, robust IL1b mRNA expression was only observed in cells continuously exposed to additional

ORF8 but not those only pre-incubated with ORF8 regardless of activation time (3 or 8 hours). These results strongly suggest ORF8 likely enters monocytes through a non-receptor-mediated process, such as phagocytosis.

Additionally, we also performed a pulldown assay to identify if surface receptors TLR2, TLR4 and intracellular receptor NLRP3 would bind to ORF8. To that end, we incubated ORF8 protein beads with monocyte lysates prepared from four pre-screened donors - two ORF8 responsive and two ORF8 non-responsive. As shown in Fig 4C, neither TLR2 or TLR4 was detectable in ORF8 precipitates. However, NLRP3 was readily detected as ORF8 binding protein in one of the two ORF8 responders, suggesting that endogenous NLRP3 indeed physically binds to ORF8 in primary human monocytes.

To ascertain the binding nature between the ORF8 and NLRP3, we transfected ORF8 expressing HEK293 cells with Flag tagged NLRP3 constructs Flag-N3-FL encoding the full-length NLRP3, or one of the deletion mutants Flag-N3-N, Flag-N3-M, or Flag-N3-C each expressing the N-terminal PYD domain (aa1-90), the middle NACHT domain (aa91-710), or the C-terminal LRR domain (aa711-1034) (Fig 4D), respectively. Cell lysates were immunoprecipitated with Strep-Tactin beads for ORF8 protein. As presented in Fig 4E, ORF8 strongly co-precipitated with NLRP3-FL as expected. Interestedly, Flag-N3-M, Flag-N3-C but not Flag-N3-N also robustly co-precipitated with ORF8 suggest both the middle NACHT domain and C-terminal LRR domain of NLRP3 can independently bind to ORF8. These results clearly demonstrate a direct link between ORF8 and activation of the NLRP3 pathway.

Blood ORF8 protein levels corelate with the mortality and disease course of severe patients with Covid-19

Having demonstrated that ORF8 is secreted as a glycoprotein and capable of inducing inflammatory cytokine responses *in vitro*, we then asked if ORF8 is also secreted and glycosylated in blood from patients newly infected with SARS-CoV-2. By analyzing serum samples collected from patients within 4-7 days of Covid-19 diagnosis using Western blotting, we found that ORF8 protein was readily detected in 92% (23/25) of newly infected patients. To determine the glycosylation status of the serum ORF8, we compared the sizes of ORF8 from patients and from HEK293 lysate. Our results (Fig 5A) show that patient serum ORF8 is indeed glycosylated when compared to the ORF8 from HEK293 lysate with its extra Strep-II tag and uncleaved signal sequences. Due to the lack of a quantitative assay, we were unable to determine the absolute quantity of ORF8 in patient samples; however, our semi-quantitative Western blot could readily detect ORF8 in as little as 4 μl of patient serum suggesting the protein is present at significant levels in patient blood at the onset of Covid-19.

Given its activity in inflammatory cytokine induction *in vitro*, and its significant presence in the blood of newly infected patients, we then asked if the ORF8 protein is the direct cause of life-threatening symptoms of Covid-19 patients. If so, the ORF8 levels should correlate with disease severity in real-time, and possibly the disease outcome. To that end, we examined the correlation between blood ORF8 levels and patients' survival outcome in our cohort of 25 hospitalized patients. As shown in Fig 5B by Kaplan-Meier analysis, after 120 days of follow-up, all seven fatalities were exclusively associated with the ORF8-high group (scored 2+ or 3+ on Western) while all patients in ORF8-low group (scored 0 or 1+) had mild symptoms and quick recovery without any events of death. These results demonstrate a clear correlation between the

ORF8 load and disease severity in newly infected patients. We then monitored the ORF8 levels at various time points during the disease course in two patients with prolonged Covid-19. As demonstrated in Fig. 5C, both patients had persistent severe symptoms requiring treatment in the intensive care unit (ICU) and their blood ORF8 levels remained high. At the time of recovery and discharge from the hospitals their blood ORF8 levels had fallen to sub-detectable levels. These results suggest that ORF8 level is closely and timely associated with severe COVID-19 disease course. Collectively, these data support that the ORF8 protein is indeed the pathogenic cause of severe Covid-19 in patients.

ORF8 mediated cytokine induction is targetable by select NLRP3 inhibitors

Given its causative role in Covid-19, we sought to determine whether the ORF8-initiated NLPR3-mediated inflammasome pathway can be targeted pharmacologically. Therefore, we tested three investigational NLRP3 inhibitors, MCC950, Tranilast, and OLT117 (Dapansutrile) on IL1b production in the monocytic cell line THP-1 as well as PBMCs from ORF8 responsive donors. Each of these inhibitors represents a unique class of NLRP3 inhibitors. As shown in Fig 5D, the NLRP3 inhibitor MCC950 showed effective inhibition on the production of IL1B in both THP-1 cells and PBMCs (blue bars), while Tranilast and OLT117 (hatched bars) exhibited moderate to marginal effect in PBMCs but no effect in THP-1 cells. Our results suggest that ORF8 mediated IL1b production can be effectively inhibited by NLRP3 inhibitor MCC950 at a nanomolar dose range, and possibly by other classes of inhibitors as well, demonstrating that targeting ORF8/NLRP3 axis is a promising strategy for treating patients with symptomatic Covid-19, an urgent unmet need for the current pandemic. Indeed, a trial of OLT117 for patients within 7 days of COVID-19 infection is ongoing (NCT04540120) in the US.

DISCUSSION

Despite the enormous success in the development and implementation of effective vaccines to prevent SARS-CoV-2 infection, Covid-19 remains a major health, social, and economic threat to society. First, a significant population in the world remains unvaccinated and will continue to be impacted by the pandemic. Second, immune response may wane with time requiring revaccination, or new emerging variants may invalidate the current vaccines. Third, many patients especially those undergoing cancer treatment, are poorly protected due to their compromised immunity.[17,18] It is most likely that the number of Covid-19 cases will remain significant for some time, and effective treatments for Covid-19 are the most urgent unmet need of the current time.

SARS-CoV-2 accessory protein ORF8, a 121 amino acid protein, is the least conserved protein in the betacoronavirus family.[19,20] Here we report that ORF8 is a major secreted viral glycoprotein both *in vitro* and in patients with newly diagnosed Covid-19. Our data, summarized in Supplemental Figure 5, support the concept that the ORF8 protein in blood stimulates circulating $CD14^+/CD16^+$ monocytes to initiate immediate and systematic cytokine responses. This is consistent with findings by others that monocytes are the single most affected subset in blood after SARS-CoV-2 infection.[21] A low level of $CD16^+$ monocytes in patients infected with SARS-CoV-2 is associated with severe disease,[22] and is likely the result of inflammation-activated pyroptosis of monocytes mediated by activated Gasdermin D. Our findings that secreted ORF8 is a key disease-causing viral factor and can be readily targeted with NLRP3 inhibitors offer a new treatment option to combat the current pandemic, especially those with persistently high ORF8 levels and severe disease.

The NLRP3 mediated inflammasome activation is a crucial part of innate immune response to pathogen-associated molecular patterns (PAMPs) or host danger associated patterns (DAMPs). When no threat signal is present, NLRP3 is in an inactive state where the middle ATPase activity-possessing NACHT domain (Fig 4D) is masked by its C-terminal LRR domain. Upon infection, the inflammation pathway is initiated with the breakdown of pathogenic peptidoglycan by lysozyme (LYZ) [23] and cathepsin proteases (CTSB, CTSL, etc.) [24] followed by NLRP3 receptor mediated recognition of PAMPs/DAMPs leading to the production of pro-IL1b and pro-IL18. Both cytokine precursors are subsequently cleaved by caspase 1 to release mature cytokines IL-1β/IL-18. Both NACHT and LRR domains are involved in NLRP3 activation through various post-translational modifications and protein recruitment.[25] We use IL1b expression level as a sensitive readout for the NLRP3 pathway activation since IL-18 is expressed at much lower levels. Our results suggest that ORF8 enters monocytes through a non-receptor mediated process followed by direct binding to the NACHT and/or LRR domains of inflammasome receptor protein NLRP3 to activate IL1b/IL18-mediated inflammatory responses. It is not known if the binding interaction with both domains are mutually exclusive or inclusive, necessary and/or sufficient, for NLRP3 activation. Further studies are needed to determine the ORF8-NLRP3 binding mechanics.

Due to the lack of effective Covid-19 animal models, it is challenging to genetically test the pathogenic role and targetability of ORF8 protein *in vivo*. However, a SARS-CoV-2 variant (Δ382) found in Taiwan and Singapore with a complete loss of the ORF8 gene serves as a tailor-made natural genetic model for validating our findings.[26,27] Young et al analyzed the clinical data of a cohort of 92 patients infected with the wildtype virus and 29 patients infected with Δ382 variant, and found that patients infected with the wildtype SARS-Cov2 exhibited much higher levels of pro-

inflammatory cytokines than those infected with the Δ382 variant. Furthermore, no patients with the Δ382 variant required supplemental oxygen whereas it was required in 26 (28%) patients infected with wildtype virus. These observations provide yet another line of evidence supporting the link between ORF8 and Covid-19 disease severity, and an *in vivo* genetic proof of the potential role to target ORF8 to prevent severe Covid-19. Additionally, various clinical studies and trials have shown that the IL1 receptor antagonist Anakinra significantly reduced the mortality risk in hospitalized patients with moderate to severe inflammation symptoms.[28,29] We believe that the data from our present study suggests that targeting NLRP3 or ORF8 upstream of the IL1 receptor will provide more effective therapeutic benefit to patients with severe Covid-19 symptoms, and perhaps could also provide prophylactic benefit to patients at risk of developing severe Covid-19.

MATERIALS AND METHODS

Patient Samples

Serum samples collected from Mayo Clinic patients with documented Covid-19 diagnosis. These patients consented to the collection of serum for the Covid-19 Research Task Force Specimen Biobank. Samples for these studies were requested and approved by the Task Force Review Committee and the Mayo Clinic Institutional Review Board.

SARS-CoV-2 protein constructs and their expression in human cells

Lentiviral constructs expressing SARS-CoV-2 proteins originally described by Gordon, et al [12] were purchased from Addgene (Table S1). Construct DNAs were used to make stable expression cell lines in HEK293 cells by lentiviral transduction, transient expression in HeLa cells using lipofectamine method or transient expression in HEK293F cells with HyCell TransFx media (Cytiva) method, respectively.

Stimulation PBMCs or pure subsets and cytokine detection

Fresh PBMCs were isolated from healthy blood donors who are 18 years or older (relevant to Covid-19 patients) by Ficoll density centrifugation. For the cytokine induction assay, we used a 24 well plate and resuspended $3x10^6$ PMBCs or pure subsets isolated using negative selection kits (StemCell Technology) in 1.2ml RPMI/10% FCS, and 0.3ml of ORF8 (20% v/v), or NSP containing conditioned medium that had been extensive dialyzed against two changes of PBS and finally fresh RPMI in a dialysis cassette with the molecular weight cutoff of 3,000 (ThermoFisher). If purified proteins were used in a smaller volume, fresh RPMI was added to achieve a final volume of 1.5ml. The cells were incubated for 3 days, then supernatants were harvested for cytokine protein detection with custom Luminex cytokine detection kit (ThermoFisher),

342 and the cells were collected for cytokine mRNA detection using qRT-PCR. Total RNAs
343 isolated using Trizol methods, following cDNA synthesis using supermix III kits
344 (ThermoFisher), the expression of various cytokines was measured by qPCR using
345 cyber green method. The final relative quantity of mRNA was extrapolated after
346 normalization with threshold cycle (Ct) of house-keeping gene beta-Actin or HPRT1,
347 and Ct of untreated control sample. All PCR primers are listed in Table S2.
348
349 *Single cell RNA sequencing of PBMCs treated with ORF8 or control medium*
350 About 3 million PMBCs were treated with either control media or ORF8-CM for
351 3 days, cells were collected, enriched for viable cells to greater than 90% using the dead
352 cell removing kit (Miltenyi Biotec, Cat#130-090-101), submitted for cell capturing and
353 single cell RNA sequencing. Barcoded Gel Beads were thawed from -80°C to room
354 temperature and the cDNA master mix was prepared according to the manufacture's
355 instruction for Chromium Single Cell 3′ v2 library kit (10x Genomics). The target
356 capture number of cells was 4000, so a volume containing more than 5000 live cells was
357 combined with cDNA master mix. The cell suspension and Chromium Single Cell 3′ v2
358 library master mix (10x Genomics), gel beads, and partitioning oil were added to a
359 Chromium Single Cell A chip which was then loaded into the Chromium Controller for
360 capturing and cDNA synthesis. The resulting cDNA libraries were pooled and
361 sequenced as 100x2 paired end reads on an Illumina HiSeq 4000 using HiSeq 3000/4000
362 sequencing kit.
363
364 *Single cell RNA sequencing Data Analysis*
365 We used the Seurat package (v4.0.1) to perform integrated analyses of single cells.
366 Genes expressed in < 3 cells and cells that expressed < 200 genes and >40% mitochondria
367 genes were excluded for downstream analysis in each sample. For scRNA-Seq analysis,

we followed the Seurat integration workflow and the comparative analysis workflow. Each dataset was normalized and the top 2000 Highly Variable Genes (HVGs) across cells were selected. The datasets were integrated based on "anchors" identified between datasets before Principal Component Analysis (PCA) was performed to do linear dimensional reduction. Shared Nearest Neighbor (SNN) Graph was constructed to identify clusters on the low-dimensional space (top 50 statistically significant principal components (PCs)). Enriched marker genes in each cluster conserved across all samples were identified, and differentially expressed genes between two selected conditions were detected using the default Wilcoxon Rank Sum test at cluster level. KEGG pathway enrichment analysis was performed on the genes that were differentially expressed significantly using the ClusterProfiler package.

Detection of SARS-CoV-2 protein secretion

To detect protein secretion, HEK293 cells expressing the protein of interest were seeded at 70% confluency in a T25 flask in 4ml of DMEM/10% FCS. After 48 hours of culturing, the supernatants were collected and cleared for cells and debris by centrifugation at 3000 rpm for 5 min. Four mL of cleared supernatants were then incubated overnight with 40ul of a 50% slurry of StrepTactin (or anti-HA for SPK protein) beads. The beads were then washed twice with PBS and then resuspended in 40 ml Laemmli sample buffer with 5% BME. The amount of sample equivalent to 1ml of culture supernatant was loaded into a 4-20% gradient gel followed by western blotting per routine probing with anti-Strep II antibody at 1:1000 dilution (#71590, Millipore) and developed on a Li-Cor scanner.

Expression and purification of ORF8 from HEK293F cell supernatant

Transfected HEK293F cells were grown for 3-4 days at a cell density below $6x10^6$ cells per ml. Cells were pelleted by centrifugation and supernatant containing secreted ORF8 was supplemented with Complete-EDTA free protease inhibitor cocktail (Roche), 1mM BME and 1 mL of biotin blocking solution (BioLock). The media was loaded on a 5 mL StrepTactin Sepharose High Performance column (GE Healthcare) in PBS buffer and eluted with a linear gradient of 0-50% 5 mM Biotin in PBS. Fractions containing ORF8 were pooled, concentrated by ultrafiltration (Amicon) and then purified using a Superdex 200 Increase 10/300 GL (Cytiva) in PBS, followed by anion exchange chromatography using a Source 15Q 4.6/10 column (Cytiva) with a gradient of 0–500 mM NaCl in 20mM Tris pH 7.5. Fractions containing ORF8 were pooled and concentrated by ultrafiltration and stored at 4°C.

Expression and purification of ORF8 from E. Coli

DNA encoding SARS CoV-2 ORF8 codon-optimized for expression in *E. coli* was cloned into pMCSG7 [30] and transformed into Rosetta2 cells (EMD), and protein expression was induced with 100 μM IPTG overnight at 15°C. *E. coli* culture was pelleted, lysed, and protein purified as described previously [31], except that cells were lysed by sonication with a Branson sonicator in place of Dounce homogenization. The concentrated, refolded protein was loaded on a Superdex S75 size-exclusion column equilibrated in PBS. Fractions containing ORF8 were pooled and concentrated by ultrafiltration (Amicon). The His-tag was removed by incubation with TEV protease. Untagged ORF8 was run on a Source 15Q 4.6/10 anion exchange column (Cytiva) equilibrated in wash buffer (20 mM Tris pH 7.5) and eluted with a linear gradient of 0-100% 500 mM NaCl in 20 mM Tris pH 7.5. Monomer and dimer peaks were pooled and concentrated.

Detection of ORF8 protein in sera of patients newly infected with SARS-CoV-2

For detection ORF8 protein in serum samples from patients newly infected with SARS-CoV-2, typically four microliters of serum directly mixed with 37 microliters of 4x Laemmli buffer with 5% of BME, boiled for 3 min prior to load to 10-20% SDS-PAGE gels, the blots were probed with 1:5000 diluted rabbit anti-ORF8 antibody (MyBioSource, Cat# MBS3014575) at 4°C overnight. The images were developed using HRP.

Detection of ORF8 protein in culture supernatants with Luminex method

To detect cytokines in the culture supernatant we used two methods – the mRNA results from qPCR and a custom Luminex detection kit for IL1b, IL6, IL8, and TNFa) (R&D systems, Minneapolis, MN). The mRNA results from qPCR of the same experiment were compared with the proteins data from Luminex.

Analysis of ORF8 expression and survival in patients with Covid-19

Overall survival was calculated from the time of Covid-19 diagnosis, using the Kaplan-Meir method on JMP 14.0 software (SAS Institute, Cary, NC).

ACKNOWLEDGEMENT

We thank The Mayo Clinic Genome Analysis Core, Bioinformatics Core, Microscopy and Cell Analysis Core, The CMM Electron Microscopy Facility of University of California San Diego (supported by NIH equipment grant 1S10OD023527-01) for their expert service, Mayo Clinical Blood Bank for providing fresh leukocyte cones of healthy donor blood, and the Mayo Clinic Center for Individualized Medicine for providing access to their biobank data. We also thank the University of Iowa/Mayo Clinic Lymphoma SPORE (NIH/NCI grant P50 CA97274) and the Predolin Foundation Biobank for sharing some of their resources for this project.

AUTHOR CONTRIBUTION

XW and TEW designed the experiments, analyzed, and interpreted results, and wrote the manuscript; XW, MKM, KEN, XT, TLW, VT, and MJS conducted experiments and analyzed data; XW, JPA, YY, XT, and MJS analyzed data and generated figures; ADB, JPA, GR, JP, and TEW provided clinical samples; JPA, GR, JP, TLW, SMA, MJS discussed and interpreted results; all authors read and approved the manuscript.

FIGURE LEGEND

Figure 1. Secretion property of SARS-CoV-2 proteins. A. Western image of cell lysates for validating SARS-CoV-2 protein expression in HEK293 cells. The blots were probed with anti-Strep II tag antibody and anti-HA (for SPK protein). **B.** Western image of SARS-CoV-2 protein pulldown using Strep Tactin beads and anti-HA-beads (for SPK) from the culture supernatants for validating SARS-CoV-2 protein secretion from HEK293 cells. The blots were also probed with anti-Strep II tag antibody and anti-HA (for SPK protein).

Figure 2. Secretion of ORF8 and induction of pro-inflammatory cytokines is glycosylation dependent. A. Conditioned media containing major secreted NSP proteins do not induce Covid-19 specific cytokines in PBMCs; B. Conditioned media containing secreted ORF8 specifically induced the Covid-19 specific cytokines (highlighted in yellow) IL1b, IL6, IL8, and CCL2 in PBMCs of some health donors. **C.** Pure highly glycosylated (yellow highlighted) but not unglycosylated ORF8 (unhighlighted) stimulates PBMCs to produce covid-19 specific (IL1b, IL6, IL8) cytokines but not IL2.

Figure 3. Single cell RNA sequencing of ORF8 treated PBMCs. A. tSNE maps of cell clusters using concatenated scRNA-Seq data from PBMC samples of two responding donors. Cluster 5 shows cells respond to ORF8 treatment. **B.** Violin plots showing the expression of intermediate monocyte markers CD14, CD16, CD68 and HLA-DRB1. **C.** Violin plots showing the expression of pro-inflammatory cytokines IL1B, IL8, and CCL2 mRNAs in cells of cluster 5. **D.** Validation study showing cytokine mRNA expression

in total PBMCs, pure $CD4^+$, $CD8^+$, $CD14^+/CD16^-$ and $CD14^+/CD16^+$ subsets stimulated with ORF8 containing conditioned medium.

Figure 4. ORF8 induces proinflammatory cytokines through activation of NLRP3 mediated inflammasome pathway. A. Violin plots showing the heightened expression of major NLRP3 pathway components LYZ, CTSB, NLPR3 in cluster 5 monocytes; **B.** ORF8 induces IL1b production through non-surface receptor mediated process. PBMCs were pre-incubated on ice with ORF8 for 2 hours followed by washing and incubating at 37ºC to activate the cytokine production in the presence or absence of additional ORF8. **C.** ORF8 directly binds to NLRP3 but not TLR2/4 in primary monocytes; **D.** Schematic drawing of NLRP3 deletion constructs for mapping ORF8 binding domains of NLPR3. **E.** Mapping NLRP3 domains that binds to ORF8 by coimmunoprecipitation using ORF8-StrepTacin™-beads. Western images of total cell lysates (left) used as expression controls and Co-IP Western images (right) showing full-length or mutant NLRP3 bind to ORF8.

Figure 5. Detection of ORF8 protein serum of patients with onset of Covid-19, correlates with outcome and disease course, and is targetable with NLRP3 inhibitors. **A.** Representative Western blot image showing intact ORF8 protein is detectable in majority of newly infected Covid-19 patient using as little as 4.0 ul of serum sample. ORF8 protein in patient sera shown as glycosylated as well; **B.** Kaplan-Meier curve showing fatality in hospital patients is associated with higher serum level of ORF8; **C.** Western images showing ORF8 levels at different time points correlate with disease course in patients with prolonged clinical courses. D. Targetability of ORF8 mediated IL1b expression with various NLRP3 inhibitors in THP1 cells and human primary PBMCs. Abbreviations: Dx = diagnosis; ED = emergency department; ICU = intensive care unit

REFERENCE

1. Johns Hopkins University Coronavirus Resource Center (2021) Retrived (November 23, 2021) from: https://coronavirus.jhu.edu/map.html.
2. Arunachalam, P.S., *et al.* Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* **369**, 1210-1220 (2020).
3. Schultze, J.L. & Aschenbrenner, A.C. COVID-19 and the human innate immune system. *Cell* **184**, 1671-1692 (2021).
4. Qi, F., Qian, S., Zhang, S. & Zhang, Z. Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses. *Biochem Biophys Res Commun* **526**, 135-140 (2020).
5. Zou, X., *et al.* Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection. *Front Med* **14**, 185-192 (2020).
6. Chen, X., *et al.* Detectable Serum Severe Acute Respiratory Syndrome Coronavirus 2 Viral Load (RNAemia) Is Closely Correlated With Drastically Elevated Interleukin 6 Level in Critically Ill Patients With Coronavirus Disease 2019. *Clin Infect Dis* **71**, 1937-1942 (2020).
7. Fajnzylber, J., *et al.* SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nat Commun* **11**, 5493 (2020).
8. Lucas, C., *et al.* Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463-469 (2020).
9. Chen, P., *et al.* SARS-CoV-2 Neutralizing Antibody LY-CoV555 in Outpatients with Covid-19. *N Engl J Med* **384**, 229-237 (2021).
10. Weinreich, D.M., *et al.* REGN-COV2, a Neutralizing Antibody Cocktail, in Outpatients with Covid-19. *N Engl J Med* **384**, 238-251 (2021).
11. Mariano, G., Farthing, R.J., Lale-Farjat, S.L.M. & Bergeron, J.R.C. Structural Characterization of SARS-CoV-2: Where We Are, and Where We Need to Be. *Front Mol Biosci* **7**, 605236 (2020).
12. Gordon, D.E., *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459-468 (2020).
13. Gordon, D.E., *et al.* Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **370**(2020).
14. Huang, C., *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497-506 (2020).
15. Lambert, C., Preijers, F., Yanikkaya Demirel, G. & Sack, U. Monocytes and macrophages in flow: an ESCCA initiative on advanced analyses of monocyte lineage using flow cytometry. *Cytometry B Clin Cytom* **92**, 180-188 (2017).
16. Kapellos, T.S., *et al.* Human Monocyte Subsets and Phenotypes in Major Chronic Inflammatory Diseases. *Front Immunol* **10**, 2035 (2019).
17. Roeker, L.E., *et al.* COVID-19 vaccine efficacy in patients with chronic lymphocytic leukemia. *Leukemia* (2021).
18. Greenberger, L.M., *et al.* Antibody response to SARS-CoV-2 vaccines in patients with hematologic malignancies. *Cancer Cell* **39**, 1031-1033 (2021).
19. Ceraolo, C. & Giorgi, F.M. Genomic variance of the 2019-nCoV coronavirus. *J Med Virol* **92**, 522-528 (2020).

20. Tan, Y., Schneider, T., Leong, M., Aravind, L. & Zhang, D. Novel Immunoglobulin Domain Proteins Provide Insights into Evolution and Pathogenesis of SARS-CoV-2-Related Viruses. *mBio* **11**(2020).
21. Notarbartolo, S., *et al.* Integrated longitudinal immunophenotypic, transcriptional and repertoire analyses delineate immune responses in COVID-19 patients. *Sci Immunol* **6**(2021).
22. Vasse, M., *et al.* A low level of CD16(pos) monocytes in SARS-CoV-2 infected patients is a marker of severity. *Clin Chem Lab Med* (2021).
23. Ragland, S.A. & Criss, A.K. From bacterial killing to immune modulation: Recent insights into the functions of lysozyme. *PLoS Pathog* **13**, e1006512 (2017).
24. Swanson, K.V., Deng, M. & Ting, J.P. The NLRP3 inflammasome: molecular activation and regulation to therapeutics. *Nat Rev Immunol* **19**, 477-489 (2019).
25. Niu, T., *et al.* NLRP3 phosphorylation in its LRR domain critically regulates inflammasome assembly. *Nat Commun* **12**, 5862 (2021).
26. Gong, Y.N., *et al.* SARS-CoV-2 genomic surveillance in Taiwan revealed novel ORF8-deletion mutant and clade possibly associated with infections in Middle East. *Emerg Microbes Infect* **9**, 1457-1466 (2020).
27. Young, B.E., *et al.* Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. *Lancet* **396**, 603-611 (2020).
28. Kyriazopoulou, E., *et al.* Early treatment of COVID-19 with anakinra guided by soluble urokinase plasminogen receptor plasma levels: a double-blind, randomized controlled phase 3 trial. *Nat Med* **27**, 1752-1760 (2021).
29. Kyriazopoulou, E., *et al.* Effect of anakinra on mortality in patients with COVID-19: a systematic review and patient-level meta-analysis. *Lancet Rheumatol* **3**, e690-e697 (2021).
30. Stols, L., *et al.* A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* **25**, 8-15 (2002).
31. Flower, T.G., *et al.* Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. *Proc Natl Acad Sci U S A* **118**(2021).

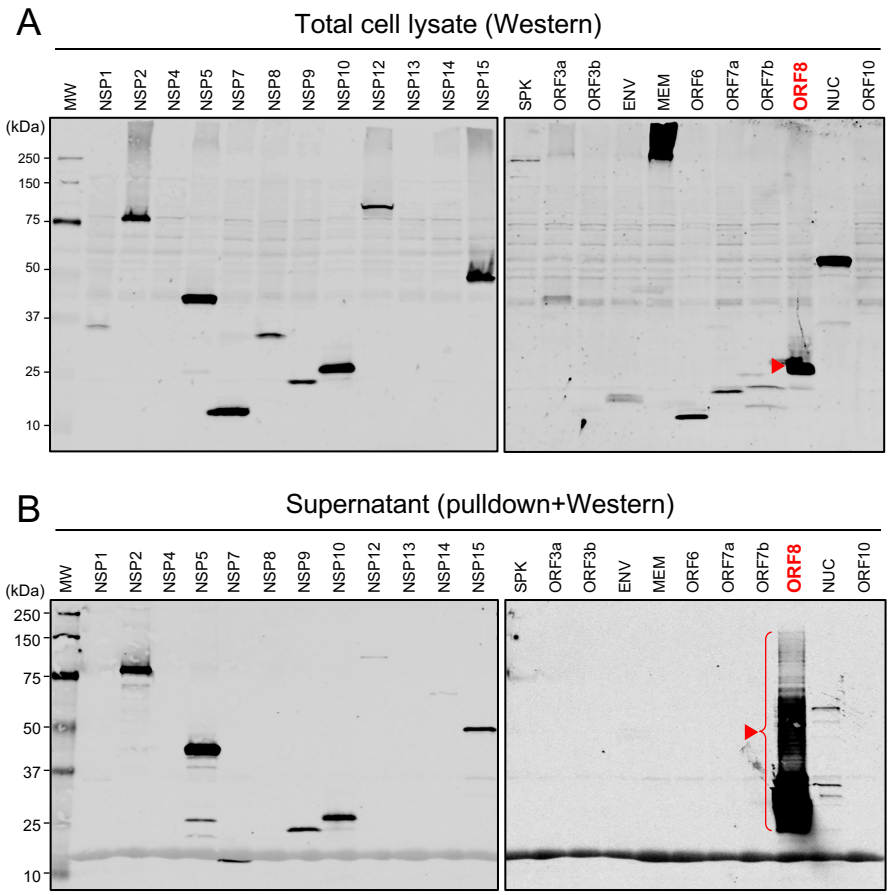

Figure 1. Secretion property of SARS-CoV-2 proteins. A. Western image of cell lysates for validating SARS-CoV-2 protein expression in HEK293 cells. The blots were probed with anti-Strep II tag antibody and anti-HA (for SPK protein). **B.** Western image of SARS-CoV-2 protein pulldown using Strep Tactin beads and anti-HA-beads (for SPK) from the culture supernatants for validating SARS-CoV-2 protein secretion from HEK293 cells. The blots were also probed with anti-Strep II tag antibody and anti-HA (for SPK protein).

Figure 1

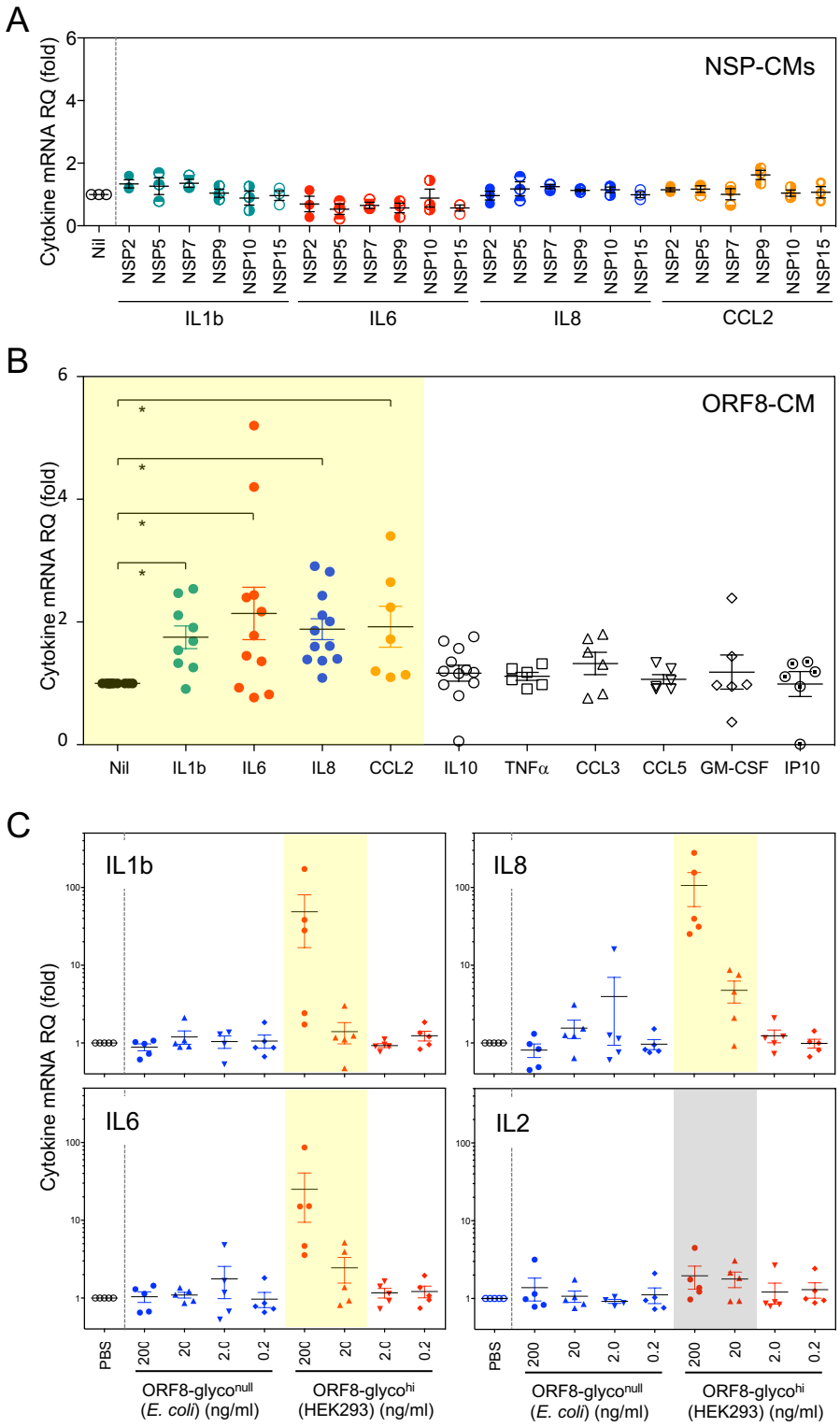

Figure 2. Secretion of ORF8 and induction of pro-inflammatory cytokines is glycosylation dependent. A. Conditioned media containing major secreted NSP proteins do not induce Covid-19 specific cytokines in PBMCs; **B.** Conditioned media containing secreted ORF8 specifically induced the Covid-19 specific cytokines (highlighted in yellow) IL1b, IL6, IL8, and CCL2 in PBMCs of some health donors . **C.** Pure highly glycosylated (yellow highlighted) but not unglycosylated ORF8 (unhighlighted) stimulates PBMCs to produce covid-19 specific (IL1b, IL6, IL8) cytokines but not IL2 .

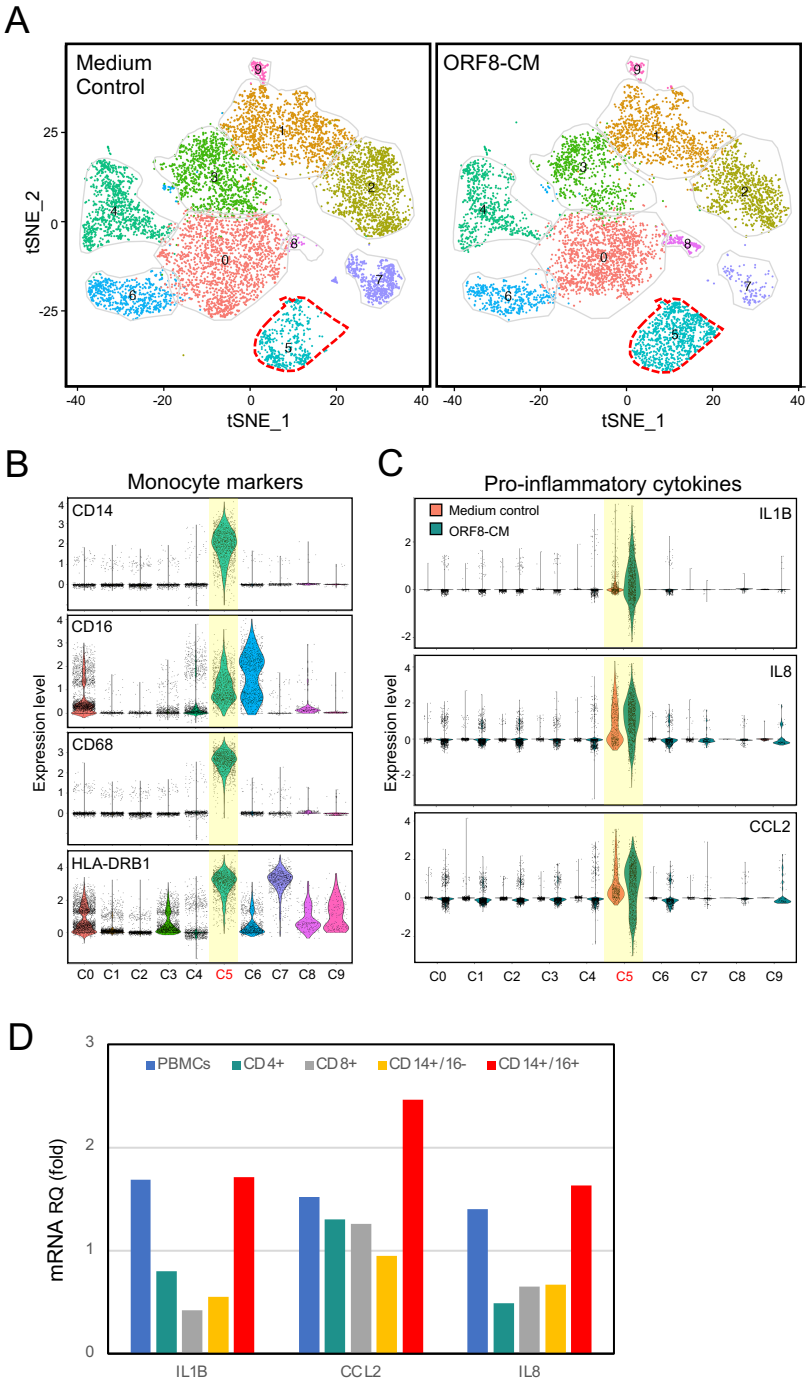

Figure 3. Single cell RNA sequencing of ORF8 treated PBMCs. A. tSNE maps of cell clusters using concatenated scRNA-Seq data from PBMC samples of two responding donors. Cluster 5 shows cells respond to ORF8 treatment. **B.** Violin plots showing the expression of intermediate monocyte markers CD14, CD16, CD68 and HLA-DRB1. **C.** Violin plots showing the expression of pro-inflammatory cytokines IL1B, IL8, and CCL2 mRNAs in cells of cluster 5. **D.** Validation study showing cytokine mRNA expression in total PBMCs, pure $CD4^+$, $CD8^+$, $CD14^+/CD16^-$ and $CD14^+/CD16^+$ subsets stimulated with ORF8 containing conditioned medium.

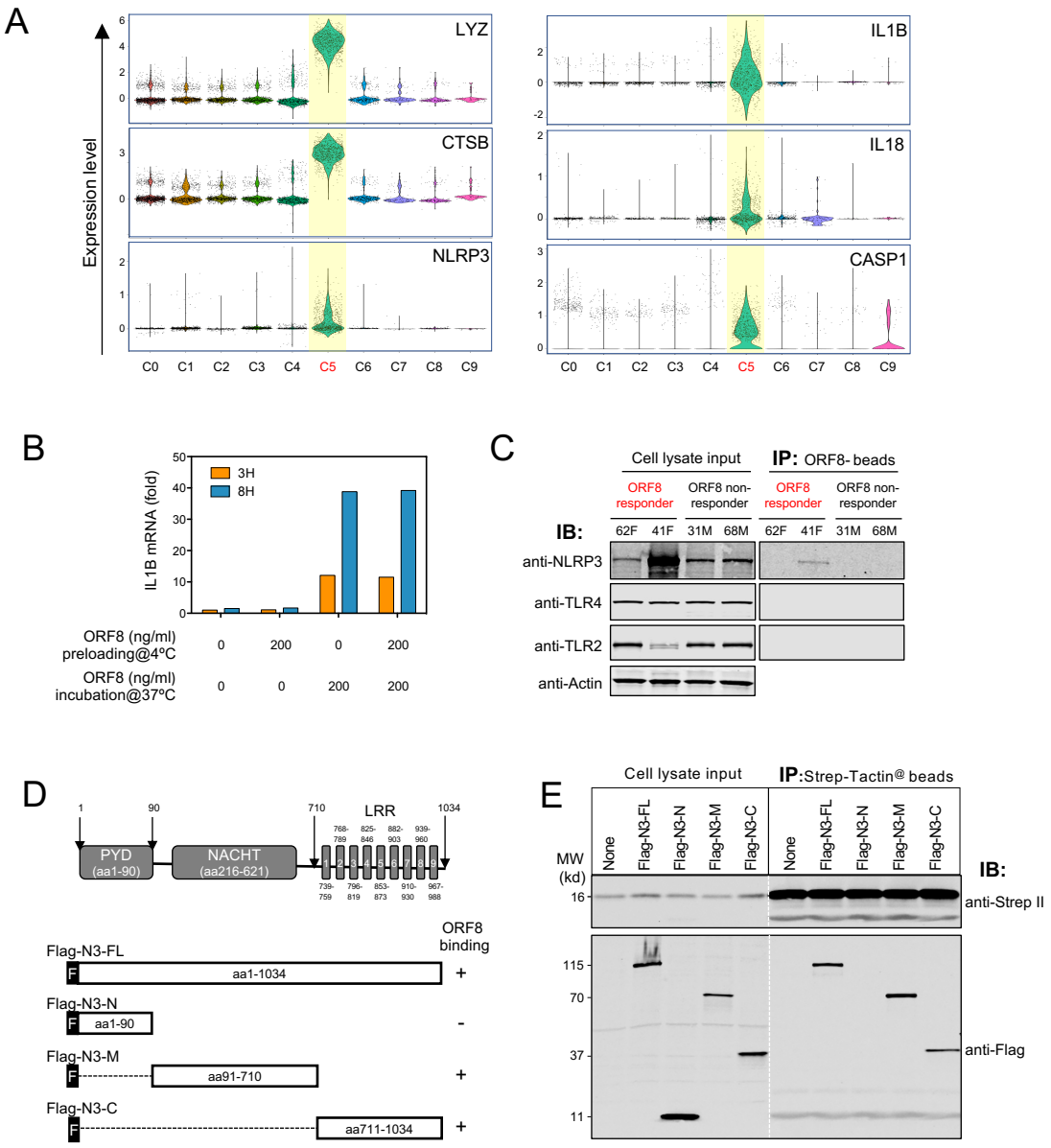

Figure 4. ORF8 induces proinflammatory cytokines through activation of NLRP3 mediated inflammasome pathway. A. Violin plots showing the heightened expression of major NLRP3 pathway components LYZ, CTSB, NLPR3 in cluster 5 monocytes; **B.** ORF8 induces IL1b production through non-surface receptor mediated process. PBMCs were pre-incubated on ice with ORF8 for 2 hours followed by washing and incubating at 37ºC to activate the cytokine production in the presence or absence of additional ORF8. **C.** ORF8 directly binds to NLRP3 but not TLR2/4 in primary monocytes; **D.** Schematic drawing of NLRP3 deletion constructs for mapping ORF8 binding domains of NLPR3. **E.** Mapping NLRP3 domains that binds to ORF8 by coimmunoprecipitation using ORF8-StrepTacin™-beads. Western images of total cell lysates (left) used as expression controls and Co-IP Western images (right) showing full-length or mutant NLRP3 bind to ORF8.

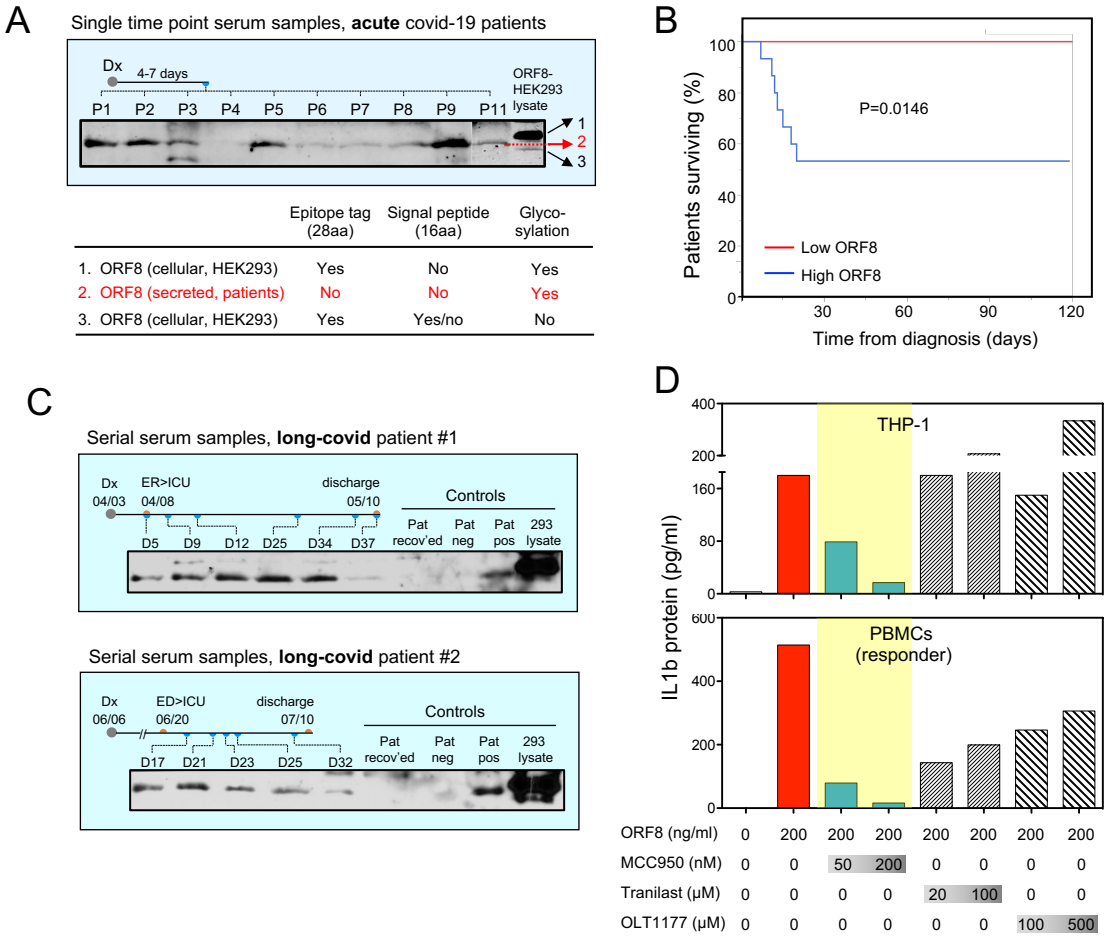

Figure 5. Detection of ORF8 protein serum of patients with onset of Covid-19, correlates with outcome and disease course, and is targetable with NLRP3 inhibitors. A. Representative Western blot image showing intact ORF8 protein is detectable in majority of newly infected Covid-19 patient using as little as 4.0 ul of serum sample. ORF8 protein in patient sera shown as glycosylated as well; **B.** Kaplan-Meier curve showing fatality in hospital patients is associated with higher serum level of ORF8; **C.** Western images showing ORF8 levels at different time points correlate with disease course in patients with prolonged clinical courses. D. Targetability of ORF8 mediated IL1b expression with various NLRP3 inhibitors in THP1 cells and human primary PBMCs. Abbriviations: Dx = diagnosis; ED = emergency department; ICU = intensive care unit