

Linking COVID-19 and heme-driven pathophysiologies: A combined computational-experimental approach

Marie-T. Hopp

University of Bonn Faculty of Mathematics and Natural Sciences: Rheinische Friedrich-Wilhelms-Universität Mathematisch-Naturwissenschaftliche Fakultät <https://orcid.org/0000-0002-0883-7564>

Daniel Domingo-Fernandez

Fraunhofer Institute for Algorithms and Scientific Computing SCAI: Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen SCAI <https://orcid.org/0000-0002-2046-6145>

Yojana Gadiya

Fraunhofer Institute for Algorithms and Scientific Computing SCAI: Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen SCAI <https://orcid.org/0000-0002-7683-0452>

Milena S. Detzel

University of Bonn Faculty of Mathematics and Natural Sciences: Rheinische Friedrich-Wilhelms-Universität Mathematisch-Naturwissenschaftliche Fakultät <https://orcid.org/0000-0001-8763-306X>

Regina Graf

University of Bonn Faculty of Mathematics and Natural Sciences: Rheinische Friedrich-Wilhelms-Universität Mathematisch-Naturwissenschaftliche Fakultät <https://orcid.org/0000-0002-7560-9991>

Benjamin F. Schmalohr

University of Bonn Faculty of Mathematics and Natural Sciences: Rheinische Friedrich-Wilhelms-Universität Mathematisch-Naturwissenschaftliche Fakultät <https://orcid.org/0000-0001-6710-5251>

Alpha T. Kodamullil

Fraunhofer Institute for Algorithms and Scientific Computing SCAI: Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen SCAI <https://orcid.org/0000-0001-9896-3531>

Diana Imhof (✉ dimhof@uni-bonn.de)

University of Bonn Faculty of Mathematics and Natural Sciences: Rheinische Friedrich-Wilhelms-Universität Mathematisch-Naturwissenschaftliche Fakultät <https://orcid.org/0000-0003-4163-7334>

Martin Hofmann-Apitius

Fraunhofer Institute for Algorithms and Scientific Computing SCAI: Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen SCAI <https://orcid.org/0000-0001-9012-6720>

Research article

Keywords: COVID-19, heme, heme-binding motifs, hemolysis, inflammation, SARS-CoV-2, pathway networks

Posted Date: March 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-319973/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The SARS-CoV-2 outbreak has been declared a worldwide pandemic in 2020. Infection triggers the respiratory tract disease COVID-19, which is accompanied by serious changes of clinical biomarkers such as hemoglobin and interleukins. The same parameters are altered during hemolysis, which is characterized by an increase in labile heme. We present two computational-experimental approaches that aim at analyzing a potential link between heme-related and COVID-19 pathophysiologies. Herein, we performed a detailed analysis of the common pathways induced by heme and SARS-CoV-2 by superimposition of knowledge graphs covering heme biology and COVID-19 pathophysiology. Focus was laid on inflammatory pathways and distinct biomarkers as the linking elements. In a second approach, four COVID-19-related proteins, the host cell proteins ACE2 and TMPRSS2 as well as the viral protein 7a and S protein, were computationally analyzed as potential heme-binding proteins with an experimental validation. The results contribute to the understanding of the progression of COVID-19 infections in patients with different clinical backgrounds and might allow for a more individual diagnosis and therapy in the future.

Introduction

In the beginning of 2020, the coronavirus disease 19 (COVID-19) has been declared a pandemic of international concern and an unprecedented challenge for the country-specific health care systems [1]. COVID-19 is caused by infections with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and accompanied by pneumonia, acute respiratory distress syndrome (ARDS) associated with a cytokine storm, and death in the most severe cases [2–4]. The renin-angiotensin system (RAS), which is associated with hypertension, is directly associated with SARS-CoV-2 viral transmission [5–7]. The virus gains access to the host cell by docking of its spike proteins (S protein) to the membrane surface of the host cell, which occurs via the transmembrane protein angiotensin-converting enzyme 2 (ACE2), an essential part of RAS [8–10]. ACE2 is expressed on the cell surface of alveolar epithelial cells of the lungs [11]. The interaction between S protein and ACE2 involves several residues in the receptor-binding domain of the S protein and ACE2 that form hydrogen bonds, a hydrophobic interaction interface and a salt bridge [7, 12, 13]. Upon binding, the S protein is subjected to proteolytic cleavage by the host cell's transmembrane serine protease subtype 2 (TMPRSS2) [8]. An interplay between the viral proteins membrane protein (M protein), nucleocapsid envelope protein (E protein) and S protein may support the viral budding process [14, 15]. Protein 7a act as an accessory protein in virus-host interaction and virus particle formation, which is crucial before release of reproduced virus particles into surrounding areas [16–18]. Immunomodulating functions of protein 7a through interaction with CD14⁺ monocytes has been demonstrated as well [19].

Numerous studies provided information about the main symptoms, risk factors for severe disease progression, and clinical diagnostic values including blood routine, blood biochemistry, and infection-related biomarkers [2, 20–22]. Among other criteria, the patients' blood group seems to affect the disease progression [23–26]. Main symptoms are fever, cough, and fatigue, all presenting reactions of an

activated immune system [21]. The activation of the immune and the complement system is monitored by a variety of markers including increased values for interleukin (IL)-6 (52% of the patients), erythrocyte sedimentation rate (85%), serum ferritin (63%), and C-reactive protein (CRP) (86%) [2, 20, 27–30]. Furthermore, studies that monitored coagulation parameters of COVID-19 patients showed a tendency towards procoagulant states [22, 31–33] and an increased risk of venous thromboembolism [34]. This was indicated, for example, by higher levels of fibrin/fibrinogen degradation products, fibrinogen, platelet hyperreactivity, increased neutrophil extracellular trap formation, and lower antithrombin levels [22, 35, 36]. In some patients, however, an increased bleeding risk has been described, possibly occurring due to elevated tissue-type plasminogen activator [37]. Levels of hemoglobin and heme-scavenging proteins (i.e., hemopexin, albumin) were often changed in COVID-19 patients [2, 20, 21, 29, 38]. These clinical parameters are interrelated when viewed from the perspective of heme and its interaction radius [39–43]. Heme is the prosthetic group of hemoglobin, which is responsible for oxygen transport in the blood [44]. Under hemolytic conditions such as sickle cell disease (SCD) and β -thalassemia, hemoglobin is degraded, resulting in an accumulation of labile heme [44, 45]. As a consequence, the heme-scavenger hemopexin, yet also albumin, become saturated, which allows heme to act as a cytotoxic, procoagulant, complement activating and proinflammatory effector [39, 41, 43, 46, 47]. These responses are, in part, mediated through interaction of heme with proteins (e.g., tumor necrosis factor α (TNF α), Toll-like receptor 4 (TLR4) [41, 42, 48, 49]) by up-regulation of cytokines (e.g., IL-1 β , TNF α) or through ROS-dependent induction of signaling pathways (e.g., MAPK/ERK pathway, NF- κ B signaling) (Dutra and Bozza, 2014). Although these findings suggest a link between processes implicated in SARS-CoV-2 and those related to heme, there is still a lack of information and data. Despite of the recent efforts invested in research on SARS-CoV-2, knowledge of the molecular mechanisms responsible for the pathophysiology of the infection still remains scarce. Likewise, data concerning heme biology is underrepresented in bioinformatics resources such as pathway databases. This knowledge gap can be addressed by using custom-made models, such as Knowledge Graphs (KGs) [50]. Placing research data in the right context by employing KGs supports the elucidation of biochemical mechanisms, the generation of novel hypotheses, and the identification of targets for drug repurposing [50–53]. Two of our recent scientific publications focused on KGs around heme as well as COVID-19 [42, 54]. Although both studies are tangential, the generated KGs can subsequently be employed to investigate the overlap between hemolytic disorders and COVID-19. Therefore, two approaches are applied herein to explore a potential link between COVID-19 and heme-driven pathophysiology. We provide insights into pathways that might play a role when considering heme in the context of COVID-19 infections by superimposing the heme KG [42] and the COVID-19 KG [54]. Furthermore, we investigated the interaction of heme with select SARS-CoV-2 proteins and specific host cell proteins by applying our established strategy employing the heme-binding motif (HBM) prediction system HeMoQuest [55] and experimental studies to characterize the computationally predicted heme-binding sites in the viral and host proteins.

Methods/experimental

Modeling the interplay between SARS-CoV-2 and heme

In order to investigate the mechanisms linking SARS-CoV-2 and heme, we exploited the KGs generated in our previous work [42, 54]. We compiled the two KGs encoded in Biological Expression Language (BEL) using PyBEL [56] directly from their public repositories (i.e., <https://github.com/covid19kg> and <https://github.com/hemekg>) and superimposed their interactions onto a merged network. Given the high degree of expressivity of BEL that enables the representation of multimodal biological information, the KGs were not only enriched with molecular information, but also with interactions from the molecular level to phenotypes and clinical readouts. We leveraged this multimodal information to hypothesize the pathways that connect key molecules associated with SARS-CoV-2 and heme to the phenotypes observed in COVID-19 patients. Since both KGs comprise several thousands of interactions, manually inspecting all relations and evaluating the implication of the crosstalk between COVID-19 and heme is largely infeasible. Accordingly, this analysis primarily focuses on the set of nodes present in both KGs. Prior to this crosstalk analysis, we conducted a one-sided Fisher's exact test [57] to confirm the significance of the overlap between human proteins present in each of the KGs (p -value < 0.01). We then manually classified the set of overlapping nodes into four pathways based on their functional role: i) immune response - inflammation, ii) immune response - complement system, iii) blood and coagulation system, and iv) organ-specific diagnostic markers. Finally, upon superimposing the relations between the overlapping nodes from the Heme and the COVID KGs, we analyzed the signature similarities between each of the above-mentioned pathways. These relations are summarized in Fig. 1 and also shown in Additional files 1–4 together with their evidence and provenance information. In order to validate the hypotheses coming from the KG, we compared the relations emerging from the overlap between the two KGs with experimental data published in the context of COVID-19 [58]. The concordance of the expression patterns in these datasets with each relation is shown in Additional file 5.

Screening For Potential HBMS In Covid-19-related Proteins

HeMoQuest (<http://131.220.139.55/SeqDHBM/>) [55] was used to identify potential HBMs in proteins of SARS-Cov-2 (i.e., S protein, M protein, E protein, protein 7a) and human host cells (i.e., ACE2, TMPRSS2). HBM prediction was refined considering surface accessibility, glycosylation sites and the involvement in disulfide bonds.

HBM-peptide Synthesis, Purification And Heme-binding Analysis

To substantiate the screening for HBMs with experimental data, the potential motifs found in S protein (3), protein 7a (2), ACE2 (5), and TMPRSS2 (10) were synthesized as nonapeptides as earlier described [59, 60]. In brief, automated solid-phase peptide synthesis was performed by standard Fmoc/tBu strategy on Rink amide MBHA resin (loading 0.53 mmol/g). For peptides containing cysteine, the methylated form has been used in order to mimic disulfide bond engagement. Preparative reverse-phase high-performance liquid chromatography (RP-HPLC; PU-987, JASCO) was used to purify the crude products. Peptides were characterized by analytical RP-HPLC (LC-20A, Shimadzu) and either LC-ESI-MS (microTOF-Q III, Bruker Daltonics; UltiMate 3000 LC, ThermoScientific) or MALDI-TOF-MS (UltrafleXtreme, Bruker Daltonics) (Additional file 6). Furthermore, amino acid analysis was carried out (LC3000, Eppendorf-Biotronik). Using an established set-up [60], the heme-binding capacity of the 20 peptides was evaluated by UV/vis

spectroscopy. Heme (0.4 - 40 μM) was incubated with each peptide (20 μM) for 30 min in HEPES buffer (100 mM, pH 7.0). Afterwards, absorbance spectra (300 - 600 nm) were measured with a Multiskan GO spectrophotometer (Thermo Fisher Scientific). Difference spectra were generated through spectra subtraction of the single heme and peptide spectra from the ones of the heme-peptide complexes. Dissociation constants (K_D) were calculated by using GraphPad prism 8.0 [60, 61].

Results

COVID-19 progression severely diverges between affected patients with ARDS and other patients, which could even remain asymptomatic. Current research is thus focusing on explaining the reasons for these discrepancies considering the physical conditions and (pre-)existing illnesses of those affected. With regard to a possible interrelation between COVID-19 and heme, several options need to be regarded. Systemic hyperinflammation follows severe COVID-19 infection. It becomes manifest by an increase in the abundance of numerous cytokines (e.g., IL-2, IL-7, IL-6, TNF α) [3], which is indicative of cytokine release syndrome [62] and leads to elevated serum biomarkers in patients (e.g., CRP, lactate dehydrogenase (LDH), D-dimer, ferritin) [2, 30, 63, 64]. Several of these indications, however, were also reported for labile heme occurring in patients with hemolytic disorders [65, 66]. In addition, ARDS has been directly correlated with increased levels of labile heme [67]. Thus, the direct interaction of heme with viral surface proteins and host cell proteins should be considered as well. This can be exemplified with the earlier reported interactions of heme with e.g., Zika, Chikungunya, and HIV-1 viruses [68–71]. In the following, we present our results concerning a detailed analysis of common pathways identified from Heme KG and COVID-19 KG as well as the potential heme-protein interactions of respective viral and host cell candidates.

Effects Of Heme And Covid-19 Intersect At Inflammation

In order to shed light on the crosstalk and common pathways between heme and COVID-19, we investigated the overlap between our two KGs (i.e., heme KG [42] and the COVID-19 KG [54]) (Fig. 1A, B). While the Heme KG was generated from the analysis of 46 scientific articles specifically selected to explain inflammatory processes related to labile heme, the COVID-19 KG contains over 150 articles. The difference in the size of these KGs thus explains the disproportionate number of molecules they possess (Fig. 1C). Nonetheless, we observed that a significant number of proteins is shared, predominantly in three major systems, namely blood coagulation, complement and immune system. Among these 85 shared nodes, there are 45 clinical phenotypes, 35 proteins, 4 immune system specific cells, and 5 small molecules. 27 nodes belong to immune response evoking (pro-)inflammatory pathways, 4 to the complement system, and 22 to the blood coagulation system (Fig. 1). Moreover, we also noticed the presence of 7 clinical phenotypes related to organ dysfunction. Further, we individually investigated the four systems to reveal the common relations observed in each of the two KGs (Fig. 1, Additional files 1–4).

We compared the directionality of these relations (i.e., up-/down-regulation) against experimental data published in the context of COVID-19 [58]. We found that the vast majority of the observed dysregulations were concordant with our findings (Additional file 5). The largest consistency was found in inflammatory pathways (Fig. 1B, D) as indicated by a common set of inflammatory – mostly pro-inflammatory – molecules. These molecules are changed with respect to their levels due to expression and/or secretion or their activity as a consequence of both, high heme concentrations and COVID-19 infection, mediating inflammatory response. In particular, the pro-inflammatory cytokines TNF α , IL-1 β , IL-6, IL-8, and the anti-inflammatory cytokine IL-10, as well as proteins related to TLR4-mediated signaling pathways (i.e., CD14, MyD88, NF- κ B and TLR4) are influenced under both conditions (Fig. 1D, Additional file 1). Within the complement system (Fig. 1E, Table conditions, which are associated with high heme concentrations, thus leading to complement activation [41]). The same was observed in COVID-19 patients [27]. Furthermore, other complement factors, including C5a and C1q, were reported to be activated by heme [41]. So far, an increase of the activation of these proteins was not described for COVID-19. Finally, the number of neutrophils is positively correlated with both heme and COVID-19 infection. However, heme induces neutrophil activation through a ROS-dependent mechanism [39], a pathway that is not yet discussed in the context of COVID-19. The blood and coagulation system is pronounced by the connecting proteins ferritin and albumin (Fig. 1F, Additional file 3). Both conditions lead to reduced levels of ferritin, a protein involved in iron uptake and release. Same applies for albumin in COVID-19 patients [2, 20, 21, 38]. Moreover, albumin is known as one of the common heme scavengers, neutralizing heme's toxic effects up to a certain extent [41]. As indicated by the impact on different components of the blood coagulation system, such as plasminogen or fibrin in case of COVID-19 and heme, respectively, both conditions can influence hemostasis. With regard to the impact on platelets, a decreased platelet count was observed in COVID-19 patients [28], whereas for heme an induction of platelet aggregation was described [43]. Finally, a trend towards elevated levels of organ-specific diagnostic markers (i.e., LDH and bilirubin) is shared by both KGs (Fig. 1G, Additional file 4).

Heme-binding Ability Of Proteins Of Sars-cov-2 And Host Cells

Numerous interesting target proteins of the virus surface and the host cell outer membrane were linked with pathological effects of SARS-CoV-2, including E protein, S protein, M protein and protein 7a as well as the human proteins ACE2 and TMPRSS2. All proteins contain at least an extracellular, surface-exposed part and are thus accessible for interaction with heme [72–75]. This led us to examine these proteins for potential heme-binding sites. It is known that regulatory heme binding occurs on surface-exposed HBMs. We identified potential HBMs in all target proteins using the recently published machine-learning web application HeMoQuest [55], which predicts HBMs from primary structure and was trained on a large array of heme-binding peptides. Screening of the amino acid sequences of S protein, protein 7a, ACE2 and TMPRSS2 resulted in 50, 6, 21, and 32 potential HBMs, respectively. M protein and E protein were dismissed as candidates, since no suitable HBMs were found. HBMs, which are part of the transmembrane or intravirion/intracellular domains, were removed from the selection, too. In addition, we excluded motifs in which the central coordinating residue was involved in disulfide bonds or where adjacent residues were glycosylated in the protein. After this refinement of the hits, we identified 24

motifs in S protein, two in protein 7a, 15 in ACE2, and 14 in TMPRSS2 (Fig. 2). These motifs were then manually screened for surface accessibility using the available protein X-ray or EM structures (Fig. 2). Consequently, three motifs for S protein, two motifs for protein 7a, five motifs for ACE2 and ten motifs for TMPRSS2 remained (Fig. 2). These were synthesized as nonapeptides for studying their heme-binding capacity, which enabled experimental classification and validation of predicted HBMs. The potential HBMs in S protein are all located in the N-terminal domain of the S1 subunit (Fig. 2A) [76, 77]. The first occurring sequence **FLGVY¹⁴⁴YHKN** is the most promising HBM, which is based on a YYH motif and further equipped with phenylalanine at position 4, two additional hydrophobic amino acids (Val, Leu), and a net charge of +2, all beneficial for heme binding [78]. This prediction was validated by UV/vis binding studies, which revealed a high heme-binding affinity ($K_D = 0.96 \pm 0.47 \mu\text{M}$; Fig. 3A) for this motif. According to the best fit, a stoichiometry of 1:2 peptide to heme was determined. While on peptide level heme binding at two sites is conceivable, on the protein level heme binding to two coordination sites within the motif is rather unlikely due to steric hindrance by adjacent residues of the protein. The other two potential HBMs **IYSKH²⁰⁷TPIN** and **LHRSY²⁴⁸LTPG** contain a Y/H-based motif with two spacers between the potential coordinating residues (e.g., YXXH), which have been shown to be less favorable for heme binding [78]. These motifs showed only a moderate ($K_D = 7.87 \pm 0.64 \mu\text{M}$) and very low heme-binding affinity ($K_D = 26.09 \pm 4.75 \mu\text{M}$; Additional file 7), respectively, rendering both unsuitable for strong heme binding to the S protein. In protein 7a, only two overlapping motifs were predicted, which is not surprising due to the small size of 121 amino acids (Fig. 2B). Both, **DGVKH⁷³VYQL** and **VKHVY⁷⁵QLRA**, possess a HXY motif [78] and three hydrophobic residues, rendering it a moderate heme binder and, in turn, protein 7a as a less interesting candidate for interaction with heme. While motif **DGVKH⁷³VYQL** did not bind heme at all *in vitro*, heme binding was observed for the second motif (**VKHVY⁷⁵QLRA**). These motifs are only shifted by two amino acids, and thus overlapping, but based on the N-terminal asparagine acid residue, the first motif (**DGVKH⁷³VYQL**) has a more acidic character which can abolish binding of heme. For the second motif (**VKHVY⁷⁵QLRA**) no binding affinity could be determined (n. sat.), but binding was observed by a shift of the heme Soret band to $\sim 419 \text{ nm}$ (Fig. 3B). Heme binding to protein 7a might thus be possible but is not very pronounced. The analysis of ACE2 revealed five potential HBMs in total, two of which representing promising H/Y motifs (Fig. 2C). The most interesting HBM is **LTAHH³⁷⁴EMGH** comprising a HXXXH motif, which was recently shown to exhibit high heme-binding affinity [78]. The central H³⁷⁴ is immediately adjacent to the site that is essential for cleavage by ADAM17 and part of the zinc(II) ion binding site of ACE2 [79–81]. However, *in vitro* only a very low heme-binding affinity was observed, which excludes this motif as a suitable HBM on the protein level. Although the occurrence of three histidines may be favorable for heme binding, the presence of E³⁷⁵ is detrimental. Since it is the zinc(II) ion binding site within ACE2, the present zinc(II) ion might abolish heme binding as well. The second interesting motif is **PLYEH²³⁹LHAY**, since it contains the efficient HXH motif [78] with further advantageous aromatic tyrosines (Y²³⁷, Y²⁴³) and hydrophobic leucines (L²³⁶, L²⁴⁰). Again, affinity is limited due to the acidic E²³⁸, which resulted in a moderate heme-binding affinity ($K_D = 4.04 \pm 1.20 \mu\text{M}$). The third motif **SFIRY⁵¹⁵YTRT** has a YY motif. Although such motifs were shown to be less favorable

with respect to affinity [78], this motif displays two basic residues that support heme binding to the motif with high affinity ($K_D = 0.60 \pm 0.33 \mu\text{M}$; Fig. 3C). The remaining two motifs QAAKH⁵³⁵EGPL and AMRQY⁶⁵⁴FLKV are less promising because they only contain one coordinating amino acid. However, while QAAKH⁵³⁵EGPL could be excluded as a potential HBM due to no heme binding *in vitro*, moderate heme-binding affinity to AMRQY⁶⁵⁴FLKV ($K_D = 5.01 \pm 0.78 \mu\text{M}$; Additional file 7) has been demonstrated. As such, ACE2 possesses one high and two moderately affine HBMs in total. The largest number of motifs (10 in total) was identified in TMPRSS2 (Fig. 2D). Seven of these motifs contain only one coordinating amino acid, including those with an additional but hindered cysteine residue. However, one of these, RDMGYKNNF, exhibits a high heme-binding affinity ($K_D = 0.94 \pm 0.38 \mu\text{M}$), suggesting it as the most promising motif in TMPRSS2 (Fig. 3D). The others either did not show heme binding, possessed only a very low affinity towards heme or a K_D value could not be determined (Additional file 7). Two further overlapping motifs (KVISH³³⁴PNYD and SHPNY³³⁷DSKT) were found in the protease domain of TMPRSS2 (Fig. 2D). Such motifs (HXXY) have been demonstrated to be less favorable, as aforementioned for two motifs from the S protein. *In vitro*, KVISH³³⁴PNYD bound heme with moderate affinity, whereas SHPNY³³⁷DSKT did not bind. Interestingly, the motif KVISH³³⁴PNYD represents a potential heme-binding sites directly in the catalytic protease domain of TMPRSS2. Within the scavenger receptor cysteine-rich (SRCR) domain of the enzyme [72], the interesting motif KKL^{YH}²²⁷SDAC was found. It features an YH motif of intermediate heme-binding affinity on the peptide level, however, of markedly improved affinity on the protein level as earlier demonstrated for IL-36 α and activated protein C (APC) [78, 82, 83]. In TMPRSS2, it has a high positive net charge and a hydrophobic leucine, which likely leads to high heme-binding affinity. With UV/vis binding studies binding to the motif was observed, however, the heme-binding affinity could not be determined due to missing saturation (Additional file 7).

Discussion

SARS-CoV-2 and its associated disease COVID-19 still keeps the world in suspense. Patients being most severely affected suffer from pneumonia, acute respiratory distress syndrome, and death [2–4]. While COVID-19 patients often exhibit high levels of proinflammatory markers as well as an activation of the complement and the coagulation system, hemoglobin and albumin levels have been reported to be remarkably low [3, 20, 27, 29]. In contrast, there is evidence for higher level of hemopexin and haptoglobin in COVID-19 patients [29]. These affected clinical parameters have generated a debate about the role of heme in the context of COVID-19 that has not been conclusively explained to date [84, 85]. With this work, we intend to provide deeper insights into a potential correlation between SARS-CoV-2 infection, COVID-19 and the effects of heme. Such a connection would be in line with studies that already described the impact of heme in the context of different viruses [68, 69, 71]. Lecerf et al. reported on the interaction of heme with antibodies (Abs) resulting in the induction of new antigen binding specificity and acquisition of binding polyreactivity to gp120 HIV-1 in 24% of the antibodies from different B-cell subpopulations of seronegative individuals [68]. The transient interaction of heme with a fraction of circulating Abs that might change their antigen binding repertoire was suggested as another possible regulatory function of

heme [68]. In addition, the novel antigen specificities of these circulating Abs was proposed to be recruited only in case of certain pathological conditions that might depend on extracellular heme as occurring in hemolytic disorders [68]. A similar report by Gupta et al. revealed the heme-mediated induction of monoclonal immunoglobulin G1 antibodies that acquired high-affinity reactivity towards antigen domain III of the Japanese encephalitis virus (JEV) E glycoprotein that exhibited neutralizing activity against dominant JEV genotypes [69]. In both cases, heme was found to confer novel binding specificities to the respective Abs without changing the binding to their cognate antigen and, as a consequence of the contact with heme, the anti-inflammatory potential of these Abs was substantially increased [69]. Finally, the inactivation of different arthropod-borne viruses like Dengue, Yellow Fever, Zika, Chikungunya, and others by porphyrin treatment has been described to occur through targeting of the viral envelope and thus, the early steps of viral infection [70, 71]. All together, these studies advocate for studying the impact of heme in coronavirus-infected patients (Fig. 4). We herein investigated the possibility of a direct interaction of heme with SARS-CoV-2 surface proteins and their human counterparts ACE2 and TMPRSS2. Our analysis revealed that heme binding conferred by HBMs is possible for all of these proteins. In case of S protein, the location of the most promising HBM correlates with the important S1 subunit of the protein, which is responsible for binding to ACE2 (Fig. 3A). This potential heme interaction would be of a transient nature, as has been observed for other heme-binding proteins such as IL-36 α and APC [82, 83]. In order to further proof these predictions, UV/vis binding studies with motif-derived peptides and heme were conducted. In the past, heme-binding studies on the level of protein-derived peptides as a tool for the evaluation of motifs derived from proteins was successful to characterize heme binding, as can be exemplified with IL-36 α and APC [82, 83]. In addition, heme binding to the S1 subunit of the S protein has been analyzed with surface plasmon resonance spectroscopy recently, which revealed a binding affinity of ~ 7 nM [86]. The study focused on bilirubin binding to the S protein, yet lacked information about heme-binding sites, stoichiometry and functional consequences of heme binding. However, the observation of direct binding to the S1 subunit clearly supports our findings, since all of the herein identified and validated motifs are found within the S1 subunit (Fig. 3A, Additional file 7). For protein 7a the central heme-binding motif VKHVV⁷⁵QLRA was confirmed (Fig. 3B, Additional file 7). The function of this protein is not yet fully understood. However, a contribution to the COVID-19 pathogenesis through induction of apoptotic pathways is expected [87], and, thus, could be affected by heme binding as well. From the pool of potential HBMs that were predicted for ACE2, three motifs turned out as suitable motifs with moderate to high heme-binding affinity (Fig. 3C, Additional file 7). All of them are part of the catalytically active zinc metallopeptidase domain of the enzyme. Transient heme binding might thus lead to a change of the catalytic activity of the enzyme, as earlier observed for example for APC [83] or hemolysin C [88]. The same applies for TMPRSS2. One high affine HBM is part of the SRCR domain and two moderately affine HBMs are found within the peptidase domain. Future binding and functional studies on the protein level for COVID-19-infection-related proteins are required to enable a complete assessment of the suggested interactions. Thereby, the correct nature of these proteins, including glycosylation, folding and multimerization should be considered, whose availability is currently restricted or not existing.

Apart from investigating the direct impact of heme on proteins at the interface of the virus-host cell interaction, we also explored similarities between relevant pathways characterizing the respective pathologies, i.e. labile heme occurrence in hemolytic conditions and COVID-19 disease progression (Fig. 4). Interestingly, we found several commonalities in the literature: Both, hemolytic conditions and COVID-19, have been found to trigger inflammatory pathways. COVID-19 patients often develop respiratory distress syndrome, which is accompanied by a cytokine storm, and thus an activation of the immune system [3]. Clinically, this is manifested by an increase in the levels of a wide range of cytokines, including TNF α , IL-1 β , IL-6 and IL-8 [3], and the activation of the complement system (e.g. C3) [27, 89]. The hemoglobin level is often decreased in COVID-19 patients [29, 62], but the underlying molecular mechanism of this phenomenon is not yet identified [84, 90, 91]. However, the lower levels seem to correlate with increased levels of the iron-storage protein ferritin. Ferritin is also upregulated during hemolytic diseases as a consequence of hemoglobin degradation and the associated increase of oxidative stress, for example induced by heme [92]. Hemolytic disorders, such as malaria and ischemia reperfusion, are associated with an excess of labile heme and are, as in COVID-19 infection, often accompanied by inflammatory events [46, 66]. Therefore, similar clinical parameters are observed under these conditions [66]. Moreover, several studies have reported that heme directly binds or induces TNF α , IL-1 β , and IL-8, and triggers numerous inflammatory pathways (e.g. NF- κ B signaling) [39, 42]. Taken together, these clinical observations suggest a correlation between both processes, which we aimed to analyze by superimposing the two KGs of both pathophysiologies [42, 54]. We overlaid Heme KG with COVID-19 KG to identify similarities, as successfully shown in the earlier reported approach for type II Diabetes and Alzheimer's disease [50, 51, 53]. Indeed, the results of the knowledge-driven analysis revealed a core of similar shared molecular patterns. The majority of these were related to the three major systems inflammation, complement and coagulation system. As expected, inflammation was the most emphasized and shared system, suggesting several processes that are commonly mediated by both heme and in COVID-19 pathogenesis. The TLR4 signaling pathway was previously shown to play an important role in heme-mediated inflammatory processes [42, 49]. Interestingly, this pathway with its components TLR4, MyD88, and NF- κ B was pronounced in the overlay of the heme KG and COVID-19 KG. The TLR4 pathway belongs to the innate immune system, and thus results in the production of several proinflammatory cytokines, such as TNF α , IL-1 β , and IL-6 [42]. TNF α and IL-1 β can further stimulate the release of inflammatory mediators, such as IL-8. Exactly the same proteins have emerged as common key molecules in our analysis. Clinical observations revealed their upregulation in COVID-19 patients as well as during hemolytic events [3, 39, 42], which highlights even more the TLR4 signaling pathway in both situations. Interestingly, TNF α and IL-1 β were reported to be capable of regulating platelet aggregation. This supports the common link of both pathologies to blood coagulation [93]. Blood parameters, such as hemoglobin and albumin levels, may allow for a direct correlation between COVID-19 progression and heme availability, since they are inevitably connected to the processing of hemoglobin and heme, respectively, under hemolytic conditions [41, 46]. At the current state of research, there is no explanation for the decreased levels of hemoglobin in COVID-19. It might be conceivable that it is due to a rapid turnover of red blood cells, which would lead to a degradation of hemoglobin and, in turn, to an increase of heme. SARS-CoV-2 infections were recently associated with anemia, hemolytic and/or

hemorrhagic conditions [94–98]. Moreover, increased heme levels ($\sim 20 \mu\text{M}$; comparable to hemolytic conditions [99, 100] were detected in COVID-19 patients [101], which further supports the importance of the results gained within this study and with respect to a need of developing tailor-made diagnosis and therapeutic strategies for these patients.

Conclusion

In summary, commonalities between COVID-19 pathophysiology and heme-driven complications under hemolytic conditions were investigated using a computational approach and a possible link supported by *in vitro* studies of select heme-ligand interactions derived from COVID-19 related proteins.

First, a detailed analysis of common pathways has been carried out through the superimposition of two knowledge graphs, i.e., “Heme KG” [42] and “COVID-19 KG” [54] that encompass information of heme-driven and COVID-19 pathophysiologies on a molecular basis. In support of the obviously common clinical parameters, several biomarkers emerged that belong to proinflammatory pathways (cytokines, such as *TNF α*), the complement system (e.g., C3), the blood and coagulation system (e.g., ferritin, platelets) as well as organ-specific markers (e.g., LDH).

In a second approach, four proteins that participate in the entry of SARS-CoV-2 into host cells, i.e., the two viral proteins S protein and protein 7a as well as the two human host cell proteins ACE2 and TMPRSS2 were predicted as potential heme-binding proteins (out of six analyzed COVID-19-related proteins). The heme-binding potential of these proteins has been evaluated with the help of the web application HeMoQuest [55] and manual refinement, which revealed two potential HBMs on the surface of protein 7a, three for S protein, five for ACE2 and 10 motifs for TMPRSS2.

Finally, heme binding to these motifs was further analyzed *in vitro* using respective peptide models (HBMs). UV/vis spectroscopy allowed for the evaluation of the heme-binding capacity and affinity of the heme association with the individual motifs. The presence of at least one highly favorable HBM for each of the four proteins was confirmed.

In conclusion, the results of this study draw attention to a relationship that is plausible based on the current characterization of COVID-19 pathogenesis by clinical parameters. A correlation between the symptoms of COVID-19 infection and the consequences of excess heme does not necessarily have to be relevant for each patient, but in specific cases it may correlate or even cause a more severe progression of the disease due to pre-existing hemolytic conditions or hemolysis-provoking events.

List Of Abbreviations

Abs antibodies

ACE2 angiotensin-converting enzyme 2

APC activated protein C

ARDS acute respiratory distress syndrome

BEL biological expression language

COVID-19 coronavirus disease 2019

E protein nucleocapsid envelope protein

ERK extracellular signal-regulated kinase

Fmoc fluorenylmethoxycarbonyl protecting group

gp glycoprotein

HBM heme-binding motif

HIV-1 human immunodeficiency virus, type 1

IL interleukin

JEV Japanese encephalitis virus

K_D dissociation constants

KG knowledge graph

LC-ESI-MS liquid chromatography electrospray ionization mass spectrometer

LDH lactate dehydrogenase

MALDI-TOF-MS matrix-assisted laser desorption/ionization time of flight mass spectrometer

MAPK mitogen-activated protein kinase

MBHA resin 4-methylbenzhydramine resin

M protein membrane protein

MyD88 myeloid differentiation primary response protein 88

NF- κ B nuclear factor kappa-light-chain-enhancer of activated B cells

RAS renin-angiotensin system

ROS reactive oxygen species

RP-HPLC reverse-phase high-performance liquid chromatography

SARS-CoV-2 severe acute respiratory syndrome coronavirus 2

SCD sickle cell disease

S protein spike glycoprotein

SRCR scavenger receptor cysteine-rich

tBu tert-Butyloxycarbonyl protecting group

TLR4 Toll-like receptor 4

TMPRSS2 transmembrane serine protease subtype 2

TNF α tumor necrosis factor α

UV/vis ultraviolet/visible

Declarations

Availability of data and materials: The datasets supporting the conclusions of this article are included within the article and its additional files. The herein used knowledge graphs and the web application HeMoQuest are publicly accessible at <https://github.com/covid19kg>, <https://github.com/hemekg> and <http://131.220.139.55/SeqDHBM/>, respectively.

Competing interests: The authors declare that they have no competing interests.

Funding: This work has been supported by the MAVO and ICON programs of the Fraunhofer Society. Financial support by the University of Bonn (Argelander grant to M.-T.H.) is gratefully acknowledged.

Authors' contributions: M.H.A. and D.I. designed and planned the project. M.-T.H., D.D.-F., Y.G., M.S.D, R.G., and B.F.S. performed the experiments and collected the data. Data analysis was carried out by all authors. The manuscript was written through the contribution of all authors.

Acknowledgement: The authors like to thank Lara Balaom and Luise Bödecker for technical assistance.

References

1. Cucinotta D, Vanelli M (2020) WHO declares COVID-19 a pandemic. *Acta Biomed* 91:157–160. <https://doi.org/10.23750/abm.v91i1.9397>
2. Zhou F, Yu T, Du R, et al (2020) Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395:1054–1062. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)

3. Ye Q, Wang B, Mao J (2020) The pathogenesis and treatment of the 'Cytokine Storm' in COVID-19. *J Infect* 80:607–613. <https://doi.org/10.1016/j.jinf.2020.03.037>
4. Ragab D, Salah Eldin H, Taeimah M, et al (2020) The COVID-19 cytokine storm: What we know so far. *Front Immunol* 11:1446. <https://doi.org/10.3389/fimmu.2020.01446>
5. Hanff TC, Harhay MO, Brown TS, et al (2020) Is there an association between COVID-19 mortality and the renin-angiotensin system? - A call for epidemiologic investigations. *Clin Infect Dis*. <https://doi.org/10.1093/cid/ciaa329>
6. Mascolo A, Scavone C, Rafaniello C, et al (2020) Renin-angiotensin system and coronavirus disease 2019: A narrative review. *Front Cardiovasc Med* 7:143. <https://doi.org/10.3389/fcvm.2020.00143>
7. Yi C, Sun X, Ye J, et al (2020) Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. *Cell Mol Immunol* 17:621–630. <https://doi.org/10.1038/s41423-020-0458-z>
8. Hoffmann M, Kleine-Weber H, Schroeder S, et al (2020) SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181:271-280.e8. <https://doi.org/10.1016/j.cell.2020.02.052>
9. Ali A, Vijayan R (2020) Dynamics of the ACE2–SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms. *Sci Rep* 10:14214. <https://doi.org/10.1038/s41598-020-71188-3>
10. Xue X, Shi J, Xu H, et al (2021) Dynamics of binding ability prediction between spike protein and human ACE2 reveals the adaptive strategy of SARS-CoV-2 in humans. *Sci Rep* 11:3187. <https://doi.org/10.1038/s41598-021-82938-2>
11. Ren X (2006) Analysis of ACE2 in polarized epithelial cells: surface expression and function as receptor for severe acute respiratory syndrome-associated coronavirus. *J Gen Virol* 87:1691–1695. <https://doi.org/10.1099/vir.0.81749-0>
12. Wu K, Peng G, Wilken M, et al (2012) Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus. *J Biol Chem* 287:8904–8911. <https://doi.org/10.1074/jbc.M111.325803>
13. Zhai X, Sun J, Yan Z, et al (2020) Comparison of severe acute respiratory syndrome coronavirus 2 spike protein binding to ACE2 receptors from human, pets, farm animals, and putative intermediate hosts. *J Virol* 94:e00831-20. <https://doi.org/10.1128/JVI.00831-20>
14. Bianchi M, Benvenuto D, Giovanetti M, et al (2020) Sars-CoV-2 envelope and membrane proteins: Structural differences linked to virus characteristics? *Biomed Res Int* 2020:1–6. <https://doi.org/10.1155/2020/4389089>
15. Boson B, Legros V, Zhou B, et al (2021) The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles. *J Biol Chem* 296:100111. <https://doi.org/10.1074/jbc.RA120.016175>
16. Fielding BC, Gunalan V, Tan THP, et al (2006) Severe acute respiratory syndrome coronavirus protein 7a interacts with hSGT. *Biochem Biophys Res Commun* 343:1201–1208. <https://doi.org/10.1016/j.bbrc.2006.03.091>

17. Kwon S, Kim E, Jung YS, et al (2020) Post-donation COVID-19 identification in blood donors. *Vox Sang*. <https://doi.org/10.1111/vox.12925>
18. Rosenthal SH, Kagan RM, Gerasimova A, et al (2020) Identification of eight SARS-CoV-2 ORF7a deletion variants in 2,726 clinical specimens. *bioRxiv*. <https://doi.org/10.1101/2020.12.10.418855>
19. Zhou Z, Huang C, Zhou Z, et al (2020) Structural insight reveals SARS-CoV-2 Orf7a as an immunomodulating factor for human CD14+ monocytes. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.3699795>
20. Chen N, Zhou M, Dong X, et al (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *Lancet* 395:507–513. [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7)
21. Guan W, Ni Z, Hu Y, et al (2020) Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 382:1708–1720. <https://doi.org/10.1056/NEJMoa2002032>
22. Han H, Yang L, Liu R, et al (2020) Prominent changes in blood coagulation of patients with SARS-CoV-2 infection. *Clin Chem Lab Med*. <https://doi.org/10.1515/cclm-2020-0188>
23. Latz CA, DeCarlo C, Boitano L, et al (2020) Blood type and outcomes in patients with COVID-19. *Ann Hematol* 99:2113–2118. <https://doi.org/10.1007/s00277-020-04169-1>
24. Li J, Wang X, Chen J, et al (2020) Association between ABO blood groups and risk of SARS-CoV-2 pneumonia. *Br J Haematol* 190:24–27. <https://doi.org/10.1111/bjh.16797>
25. Wu B-B, Gu D-Z, Yu J-N, et al (2020) Association between ABO blood groups and COVID-19 infection, severity and demise: A systematic review and meta-analysis. *Infect Genet Evol* 84:104485. <https://doi.org/10.1016/j.meegid.2020.104485>
26. Zhao J, Yang Y, Huang H, et al (2020) Relationship between the ABO blood group and the coronavirus disease 2019 (COVID-19) susceptibility. *Clin Infect Dis*. <https://doi.org/10.1093/cid/ciaa1150>
27. Risitano AM, Mastellos DC, Huber-Lang M, et al (2020) Complement as a target in COVID-19? *Nat Rev Immunol* 20:343–344. <https://doi.org/10.1038/s41577-020-0320-7>
28. Wang R, Pan M, Zhang X, et al (2020) Epidemiological and clinical features of 125 Hospitalized Patients with COVID-19 in Fuyang, Anhui, China. *Int J Infect Dis* 95:421–428. <https://doi.org/10.1016/j.ijid.2020.03.070>
29. Whetton AD, Preston GW, Abubeker S, Geifman N (2020) Proteomics and informatics for understanding phases and identifying biomarkers in COVID-19 disease. *J Proteome Res* 19:4219–4232. <https://doi.org/10.1021/acs.jproteome.0c00326>
30. Zhang X, Cai H, Hu J, et al (2020) Epidemiological, clinical characteristics of cases of SARS-CoV-2 infection with abnormal imaging findings. *Int J Infect Dis* 94:81–87. <https://doi.org/10.1016/j.ijid.2020.03.040>
31. Tang N, Li D, Wang X, Sun Z (2020) Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. *J Thromb Haemost* 18:844–847. <https://doi.org/10.1111/jth.14768>

32. Ji H-L, Zhao R, Matalon S, Matthay MA (2020) Elevated plasmin(ogen) as a common risk factor for COVID-19 susceptibility. *Physiol Rev* 100:1065–1075. <https://doi.org/10.1152/physrev.00013.2020>
33. Grobler C, Maphumulo SC, Grobbelaar LM, et al (2020) Covid-19: The rollercoaster of fibrin(ogen), D-Dimer, von Willebrand factor, P-selectin and their interactions with endothelial cells, platelets and erythrocytes. *Int J Mol Sci* 21:5168. <https://doi.org/10.3390/ijms21145168>
34. Giannis D, Ziogas IA, Gianni P (2020) Coagulation disorders in coronavirus infected patients: COVID-19, SARS-CoV-1, MERS-CoV and lessons from the past. *J Clin Virol* 127:104362. <https://doi.org/10.1016/j.jcv.2020.104362>
35. Manne BK, Denorme F, Middleton EA, et al (2020) Platelet gene expression and function in patients with COVID-19. *Blood* 136:1317–1329. <https://doi.org/10.1182/blood.2020007214>
36. Leppkes M, Knopf J, Naschberger E, et al (2020) Vascular occlusion by neutrophil extracellular traps in COVID-19. *EBioMedicine* 58:102925. <https://doi.org/10.1016/j.ebiom.2020.102925>
37. Zuo Y, Warnock M, Harbaugh A, et al (2021) Plasma tissue plasminogen activator and plasminogen activator inhibitor-1 in hospitalized COVID-19 patients. *Sci Rep* 11:1580. <https://doi.org/10.1038/s41598-020-80010-z>
38. Yang X, Yu Y, Xu J, et al (2020) Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 8:475–481. [https://doi.org/10.1016/S2213-2600\(20\)30079-5](https://doi.org/10.1016/S2213-2600(20)30079-5)
39. Dutra FF, Bozza MT (2014) Heme on innate immunity and inflammation. *Front Pharmacol* 5:115. <https://doi.org/10.3389/fphar.2014.00115>
40. Kühl T, Imhof D (2014) Regulatory Fe (II/III) heme: The reconstruction of a molecule's biography. *ChemBioChem* 15:2024–2035. <https://doi.org/10.1002/cbic.201402218>
41. Roumenina LT, Rayes J, Lacroix-Desmazes S, Dimitrov JD (2016) Heme: Modulator of plasma systems in hemolytic diseases. *Trends Mol Med* 22:200–213. <https://doi.org/10.1016/j.molmed.2016.01.004>
42. Humayun F, Domingo-Fernández D, Paul George AA, et al (2020) A computational approach for mapping heme biology in the context of hemolytic disorders. *Front Bioeng Biotechnol* 8:74. <https://doi.org/10.3389/fbioe.2020.00074>
43. Hopp M-T, Imhof D (2021) Linking labile heme with thrombosis. *J Clin Med* 10:427. <https://doi.org/10.3390/jcm10030427>
44. Ascenzi P, Bocedi A, Visca P, et al (2005) Hemoglobin and heme scavenging. *IUBMB Life* 57:749–759. <https://doi.org/10.1080/15216540500380871>
45. Gouveia Z, Carlos AR, Yuan X, et al (2017) Characterization of plasma labile heme in hemolytic conditions. *FEBS J* 284:3278–3301. <https://doi.org/10.1111/febs.14192>
46. Chiabrando D, Vinchi F, Fiorito V, et al (2014) Heme in pathophysiology: A matter of scavenging, metabolism and trafficking across cell membranes. *Front Pharmacol* 5:61. <https://doi.org/10.3389/fphar.2014.00061>

47. Martins R, Maier J, Gorki A-D, et al (2016) Heme drives hemolysis-induced susceptibility to infection via disruption of phagocyte functions. *Nat Immunol* 17:1361–1372. <https://doi.org/10.1038/ni.3590>
48. Kupke T, Klare JP, Brügger B (2020) Heme binding of transmembrane signaling proteins undergoing regulated intramembrane proteolysis. *Commun Biol* 3:73. <https://doi.org/10.1038/s42003-020-0800-0>
49. Janciauskiene S, Vijayan V, Immenschuh S (2020) TLR4 signaling by heme and the role of heme-binding blood proteins. *Front Immunol* 11:. <https://doi.org/10.3389/fimmu.2020.01964>
50. Kodamullil AT, Younesi E, Naz M, et al (2015) Computable cause-and-effect models of healthy and Alzheimer's disease states and their mechanistic differential analysis. *Alzheimer's Dement* 11:1329–1339. <https://doi.org/10.1016/j.jalz.2015.02.006>
51. Karki R, Kodamullil AT, Hofmann-Apitius M (2017) Comorbidity analysis between Alzheimer's Disease and type 2 diabetes mellitus (T2DM) based on shared pathways and the role of T2DM drugs. *J Alzheimer's Dis* 60:721–731. <https://doi.org/10.3233/JAD-170440>
52. Emon MAEK, Kodamullil AT, Karki R, et al (2017) Using Drugs as Molecular Probes: A Computational Chemical Biology Approach in Neurodegenerative Diseases. *J Alzheimer's Dis* 56:677–686. <https://doi.org/10.3233/JAD-160222>
53. Karki R, Madan S, Gadiya Y, et al (2020) Data-driven modeling of knowledge assemblies in understanding comorbidity between type 2 diabetes mellitus and Alzheimer's Disease. *J Alzheimer's Dis* 78:87–95. <https://doi.org/10.3233/JAD-200752>
54. Domingo-Fernández D, Baksi S, Schultz B, et al (2020) COVID-19 Knowledge Graph: a computable, multi-modal, cause-and-effect knowledge model of COVID-19 pathophysiology. *Bioinformatics* btaa834:. <https://doi.org/10.1093/bioinformatics/btaa834>
55. Paul George AA, Lacerda M, Syllwasschy BF, et al (2020) HeMoQuest: a webserver for qualitative prediction of transient heme binding to protein motifs. *BMC Bioinformatics* 21:124. <https://doi.org/10.1186/s12859-020-3420-2>
56. Hoyt CT, Konotopez A, Ebeling C (2018) PyBEL: A computational framework for Biological Expression Language. *Bioinformatics* 34:703–704. <https://doi.org/10.1093/bioinformatics/btx660>
57. Fisher RA (1992) Statistical methods for research workers. In: *Breakthroughs in statistics*. pp 66–70
58. Blanco-Melo D, Nilsson-Payant BE, Liu W-C, et al (2020) Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* 181:1036-1045.e9. <https://doi.org/10.1016/j.cell.2020.04.026>
59. Kühn T, Sahoo N, Nikolajski M, et al (2011) Determination of Hemin-Binding Characteristics of Proteins by a Combinatorial Peptide Library Approach. *ChemBioChem* 12:2846–2855. <https://doi.org/10.1002/cbic.201100556>
60. Kühn T, Wißbrock A, Goradia N, et al (2013) Analysis of Fe(III) heme binding to cysteine-containing heme-regulatory motifs in proteins. *ACS Chem Biol* 8:1785–1793. <https://doi.org/10.1021/cb400317x>
61. Pîrnău A, Bogdan M (2008) Investigation of the interaction between flavonoids and human serum albumin. *Rom J Biophys* 18:49–55. <https://doi.org/10.1016/j.molstruc.2004.05.026>

62. Huang C, Wang Y, Li X, et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395:497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
63. Chen T, Wu D, Chen H, et al (2020) Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ* 368:m1091. <https://doi.org/10.1136/bmj.m1091>
64. Young BE, Ong SWX, Kalimuddin S, et al (2020) Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. *J Amer Med Assoc* 323:1488–1494. <https://doi.org/10.1001/jama.2020.3204>
65. Litalien C, Proulx F, Mariscalco MM, et al (1999) Circulating inflammatory cytokine levels in hemolytic uremic syndrome. *Pediatr Nephrol* 13:840–845. <https://doi.org/10.1007/s004670050712>
66. Barcellini W, Fattizzo B (2015) Clinical applications of hemolytic markers in the differential diagnosis and management of hemolytic anemia. *Dis Markers* 2015:635670. <https://doi.org/10.1155/2015/635670>
67. Aggarwal S, Lazrak A, Ahmad I, et al (2020) Reactive species generated by heme impair alveolar epithelial sodium channel function in acute respiratory distress syndrome. *Redox Biol* 36:101592. <https://doi.org/10.1016/j.redox.2020.101592>
68. Lecerf M, Scheel T, Pashov AD, et al (2015) Prevalence and gene characteristics of antibodies with cofactor-induced HIV-1 specificity. *J Biol Chem* 290:5203–5213. <https://doi.org/10.1074/jbc.M114.618124>
69. Gupta N, de Wispelaere M, Lecerf M, et al (2015) Neutralization of Japanese Encephalitis Virus by heme-induced broadly reactive human monoclonal antibody. *Sci Rep* 5:16248. <https://doi.org/10.1038/srep16248>
70. Assunção-Miranda I, Cruz-Oliveira C, Neris RLS, et al (2016) Inactivation of Dengue and Yellow Fever viruses by heme, cobalt-protoporphyrin IX and tin-protoporphyrin IX. *J Appl Microbiol* 120:790–804. <https://doi.org/10.1111/jam.13038>
71. Neris RLS, Figueiredo CM, Higa LM, et al (2018) Co-protoporphyrin IX and Sn-protoporphyrin IX inactivate Zika, Chikungunya and other arboviruses by targeting the viral envelope. *Sci Rep* 8:9805. <https://doi.org/10.1038/s41598-018-27855-7>
72. Mendes de Oliveira G, Valle Garay A, Araújo Souza A, et al (2018) Structural characterization and crystallization of human TMPRSS2 protease. *Biophys J* 114:567a. <https://doi.org/10.1016/j.bpj.2017.11.3102>
73. Mousavizadeh L, Ghasemi S (2020) Genotype and phenotype of COVID-19: Their roles in pathogenesis. *J Microbiol Immunol Infect* in press: <https://doi.org/10.1016/j.jmii.2020.03.022>
74. Walls AC, Park Y-J, Tortorici MA, et al (2020) Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 181:281-292.e6. <https://doi.org/10.1016/j.cell.2020.02.058>
75. Hänel K, Willbold D (2007) SARS-CoV accessory protein 7a directly interacts with human LFA-1. *Biol Chem* 388:1325–1332. <https://doi.org/10.1515/BC.2007.157>
76. Ou X, Liu Y, Lei X, et al (2020) Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* 11:1620.

<https://doi.org/10.1038/s41467-020-15562-9>

77. Huang Y, Yang C, Xu X, et al (2020) Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19. *Acta Pharmacol Sin* 41:1141–1149. <https://doi.org/10.1038/s41401-020-0485-4>
78. Syllwasschy BF, Beck MS, Družeta I, et al (2020) High-affinity binding and catalytic activity of His/Tyr-based sequences: Extending heme-regulatory motifs beyond CP. *Biochim Biophys Acta - Gen Subj* 1864:129603. <https://doi.org/10.1016/j.bbagen.2020.129603>
79. Towler P, Staker B, Prasad SG, et al (2004) ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J Biol Chem* 279:17996–18007. <https://doi.org/10.1074/jbc.M311191200>
80. Heurich A, Hofmann-Winkler H, Gierer S, et al (2014) TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J Virol* 88:1293–1307. <https://doi.org/10.1128/JVI.02202-13>
81. Lan J, Ge J, Yu J, et al (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 581:215–220. <https://doi.org/10.1038/s41586-020-2180-5>
82. Wißbrock A, Goradia NB, Kumar A, et al (2019) Structural insights into heme binding to IL-36α proinflammatory cytokine. *Sci Rep* 9:16893. <https://doi.org/10.1038/s41598-019-53231-0>
83. Hopp M-T, Alhanafi N, Paul George AA, et al (2021) Molecular insights and functional consequences of the interaction of heme with activated protein C. *Antioxid Redox Signal* 34:32–48. <https://doi.org/10.1089/ars.2019.7992>
84. Liu W, Li H (2020) COVID-19: Attacks the 1-beta chain of hemoglobin and captures the porphyrin to inhibit human heme metabolism. *ChemRxiv*. <https://doi.org/10.26434/chemrxiv.11938173>
85. Wagener FADTG, Pickkers P, Peterson SJ, et al (2020) Targeting the heme-heme oxygenase system to prevent severe complications following COVID-19 infections. *Antioxidants* 9:540. <https://doi.org/10.3390/antiox9060540>
86. Rosa A, Pye VE, Graham C, et al (2021) SARS-CoV-2 recruits a haem metabolite to evade antibody immunity. *medRxiv*. <https://doi.org/10.1101/2021.01.21.21249203>
87. Michel CJ, Mayer C, Poch O, Thompson JD (2020) Characterization of accessory genes in coronavirus genomes. *Virol J* 17:131. <https://doi.org/10.1186/s12985-020-01402-1>
88. Peherstorfer S, Brewitz HH, Paul George AA, et al (2018) Insights into mechanism and functional consequences of heme binding to hemolysin-activating lysine acyltransferase HlyC from *Escherichia coli*. *Biochim Biophys Acta - Gen Subj* 1862:1964–1972. <https://doi.org/10.1016/j.bbagen.2018.06.012>
89. Skendros P, Mitsios A, Chrysanthopoulou A, et al (2020) Complement and tissue factor–enriched neutrophil extracellular traps are key drivers in COVID-19 immunothrombosis. *J Clin Invest* 130:6151–6157. <https://doi.org/10.1172/JCI141374>

90. Daniel Y, Hunt BJ, Retter A, et al (2020) Haemoglobin oxygen affinity in patients with severe COVID-19 infection. *Br J Haematol* 190:e126–e127. <https://doi.org/10.1111/bjh.16888>
91. DeMartino AW, Rose JJ, Amdahl MB, et al (2020) No evidence of hemoglobin damage by SARS-CoV-2 infection. *Haematologica* 105:2769–2773. <https://doi.org/10.3324/haematol.2020.264267>
92. Belcher JD, Beckman JD, Balla G, et al (2010) Heme degradation and vascular injury. *Antioxid Redox Signal* 12:233–248. <https://doi.org/10.1089/ars.2009.2822>
93. Bar J, Zosmer A, Hod M, et al (1997) The regulation of platelet aggregation in vitro by interleukin-1beta and tumor necrosis factor-alpha: changes in pregnancy and in pre-eclampsia. *Thromb Haemost* 78:1255–1261
94. Lazarian G, Quinquenel A, Bellal M, et al (2020) Autoimmune haemolytic anaemia associated with COVID-19 infection. *Br J Haematol* 190:29–31. <https://doi.org/10.1111/bjh.16794>
95. Capes A, Bailly S, Hantson P, et al (2020) COVID-19 infection associated with autoimmune hemolytic anemia. *Ann Hematol* 99:1679–1680. <https://doi.org/10.1007/s00277-020-04137-9>
96. Conti CB, Henchi S, Coppeta GP, et al (2020) Bleeding in COVID-19 severe pneumonia: The other side of abnormal coagulation pattern? *Eur J Intern Med* 77:147–149. <https://doi.org/10.1016/j.ejim.2020.05.002>
97. Agarwal A, Vishnu VY, Vibha D, et al (2020) Intracerebral hemorrhage and SARS-CoV-2: association or causation. *Ann Indian Acad Neurol* 23:261–264. https://doi.org/10.4103/aian.AIAN_362_20
98. Sahu KK, Borogovac A, Cerny J (2021) COVID-19 related immune hemolysis and thrombocytopenia. *J Med Virol* 93:1164–1170. <https://doi.org/10.1002/jmv.26402>
99. Reiter CD, Wang X, Tanus-Santos JE, et al (2002) Cell-free hemoglobin limits nitric oxide bioavailability in sickle-cell disease. *Nat Med* 8:1383–1389. <https://doi.org/10.1038/nm1202-799>
100. Oh J-YJ-Y, Hamm J, Xu X, et al (2016) Absorbance and redox based approaches for measuring free heme and free hemoglobin in biological matrices. *Redox Biol* 9:167–177
101. Su W-L, Lin C-P, Hang H-C, et al (2021) Desaturation and heme elevation during COVID-19 infection: A potential prognostic factor of heme oxygenase-1. *J Microbiol Immunol Infect* 54:113–116. <https://doi.org/10.1016/j.jmii.2020.10.001>
102. Zhang C, Zheng W, Huang X, et al (2020) Protein structure and sequence reanalysis of 2019-nCoV genome refutes snakes as its intermediate host and the unique similarity between its spike protein insertions and HIV-1. *J Proteome Res* 19:1351–1360. <https://doi.org/10.1021/acs.jproteome.0c00129>

Figures

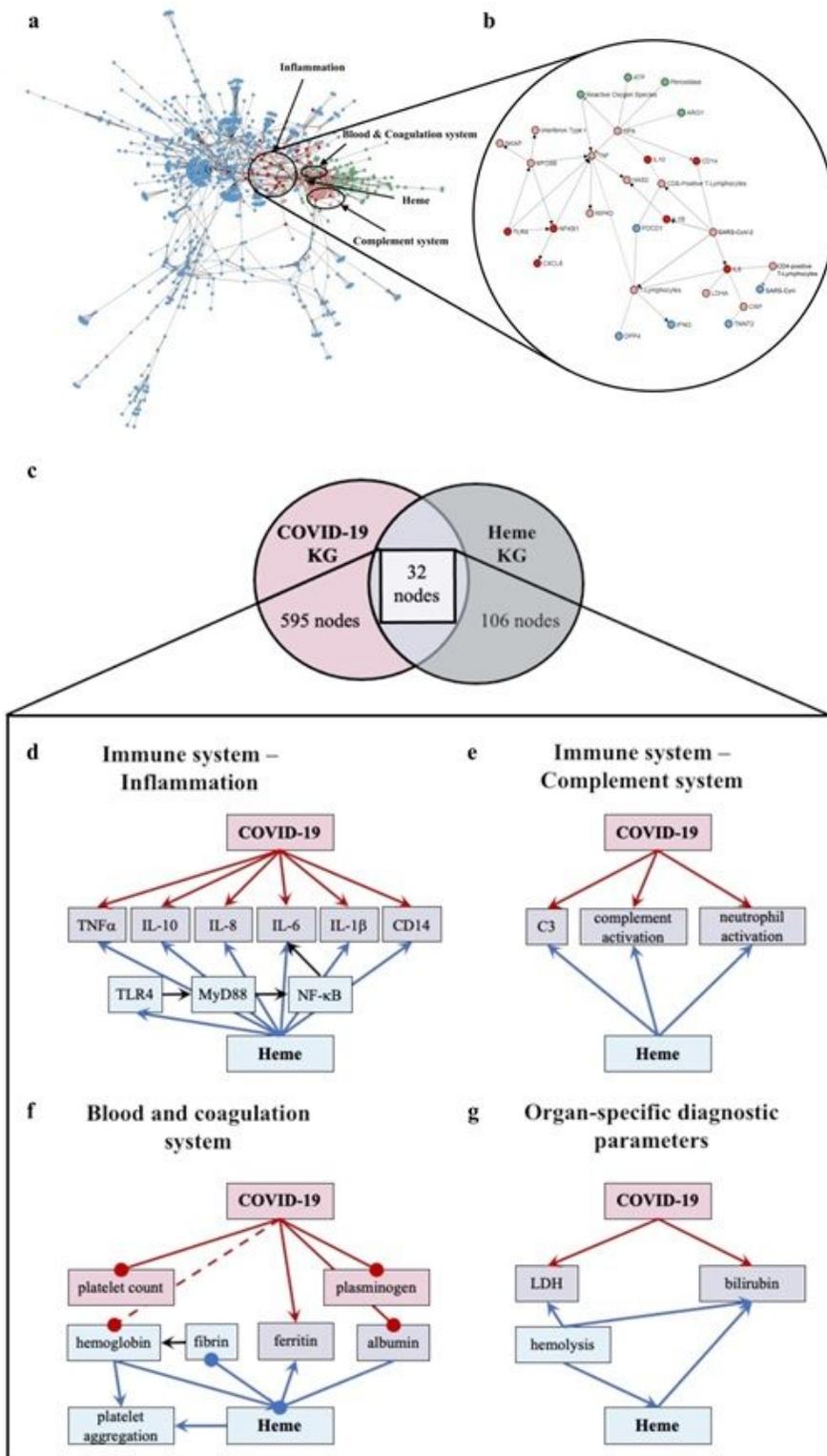


Figure 1

Overlap between the COVID-19 KG [54] and Heme KG [42] reveals shared biochemical pathways. a) The network displays the largest area of overlap between the two KGs. Node coloring denotes whether a particular entity is present exclusively in the COVID-19 KG (blue), in the Heme KG (green), or in both (red). Neighbors of overlapping nodes are colored in light red. The parts of the network corresponding to inflammation, blood coagulation and complement system are circled. b) High resolution network view of

the inflammation system with directionality of the relations. Solid edges represent increase and decrease relations, while dashed edges represent correlations. c) Overlap between the two KGs based on human proteins represented as a Venn diagram. The numbers of nodes that only present proteins are depicted (COVID-19 KG: 595 (red), Heme KG: 106 (blue), overlap between both: 32 (violet)). The two KGs overlap in the following systems: d) immune response - inflammation, e) immune response - complement system, f) blood and coagulation system, and g) organ-specific diagnostic markers. The hemoglobin level is often decreased in COVID-19 patients [29, 62] and is depicted (dashed red line). C3: complement component 3, IL: interleukin, LDH: lactate dehydrogenase.

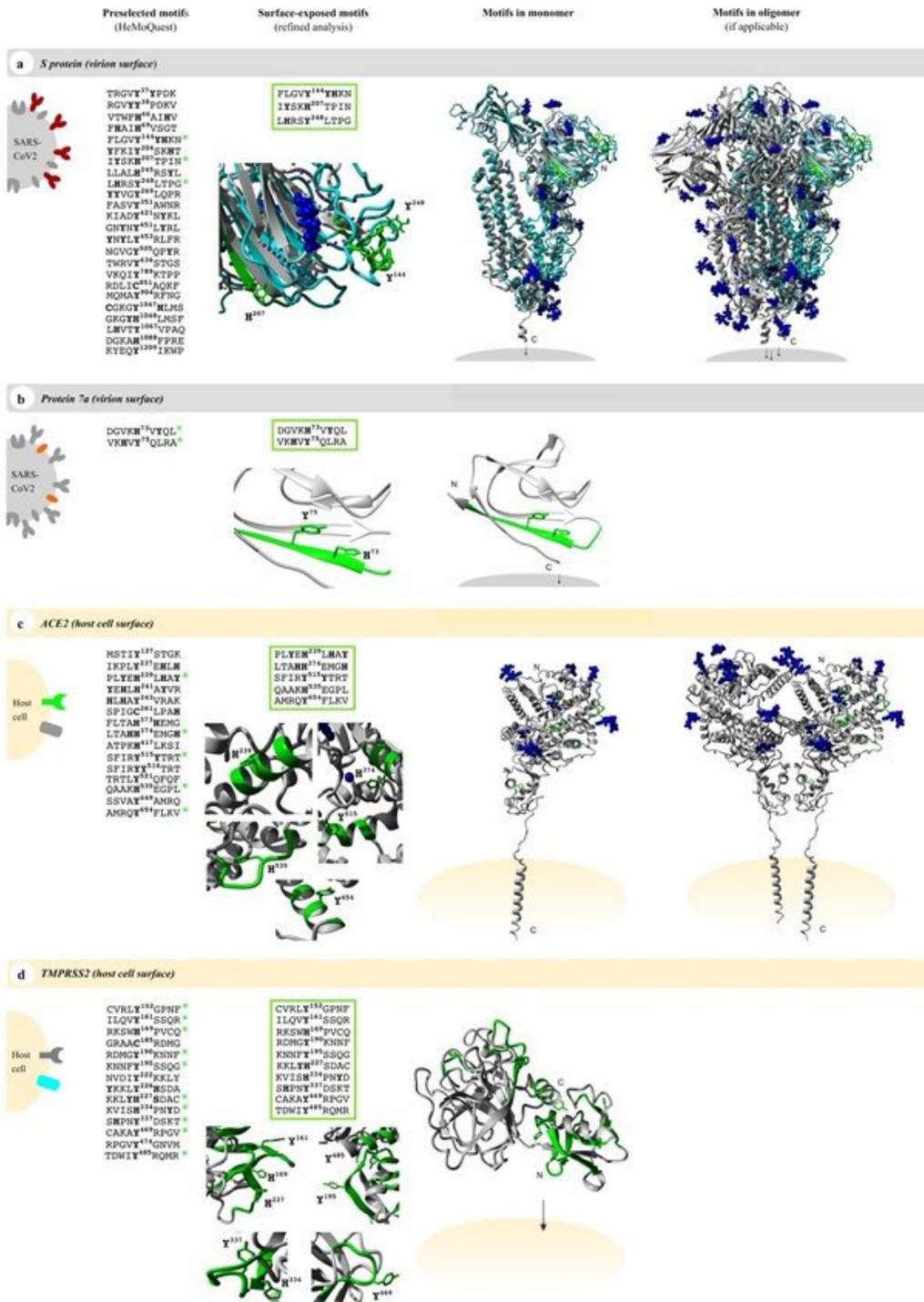


Figure 2

Potential heme-binding proteins on the virus and host cell surface. Four COVID-19-related proteins (first column, left), namely the virus proteins (grey) a) S protein (red) and b) protein 7a (orange) as well as the host cell proteins (yellow) c) ACE2 (green) and d) TMPRSS2 (turquoise), represent potential heme-binding proteins. All motifs predicted by HeMoQuest [55] are shown excluding those with modifications (glycosylation, disulfide bonds) or located in intracellular or virion domains. A refined analysis

considering the surface accessibility of the motifs resulted in the following number of motifs: 3 (S protein), 2 (protein 7a), 5 (ACE2), and 10 (TMPRSS2) (third column). The motifs are highlighted in a zoom-in below (green; third column), as well as in the available monomer (fourth column) and oligomer (fifth column) structures, if applicable (S protein, homology model from C-I-TASSER [102]; protein 7a, PDB: 6W37; ACE2, PDB: 6M18; TMPRSS2, Swiss-model: O15393). Within the oligomers, the motifs were only depicted in one of the monomers (green). Since some surface-exposed motifs in the S protein were not covered by the available EM structure (PDB: 6VXX), motifs were highlighted in the monomer (turquoise), which was then superimposed with the trimer (PDB: 6VXX). Blue: glycosylation sites and ions (if applicable).

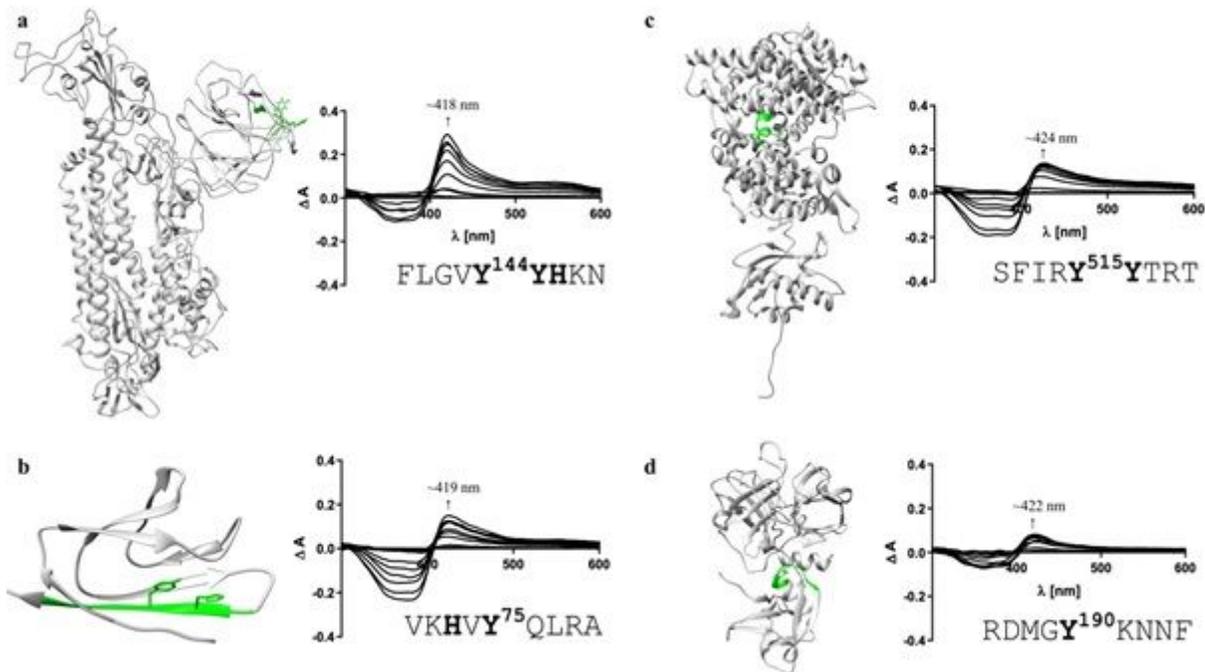


Figure 3

Heme-binding capacity of the most promising HBMs in SARS-CoV-2 and host cell proteins. Heme (0.4 - 40 μM)-binding properties of the motif-derived peptides (10 μM) were investigated by UV/vis spectroscopy. The locations of the most promising HBMs within the proteins' structures (extracellular part; left) and the difference spectra (right) are depicted. a) With a KD of $0.96 \pm 0.47 \mu\text{M}$ (right), FLGVY144YHKN (green) is the highest affine HBM in the S protein. b) For protein 7a (PDB: 6W37), VKHVV75QLRA was confirmed as a HBM (KD determination not possible). c) Within human ACE2, the motif SFIRY515YTRT (green) exhibits the highest heme-binding affinity (KD = $0.60 \pm 0.33 \mu\text{M}$). d) For TMPRSS2, one high-affine HBM in the SRCR domain has been identified (RDMGY190KNNF; KD = $0.94 \pm 0.38 \mu\text{M}$). HBMs that emerged only as moderate heme binders are not shown (Figure S1).

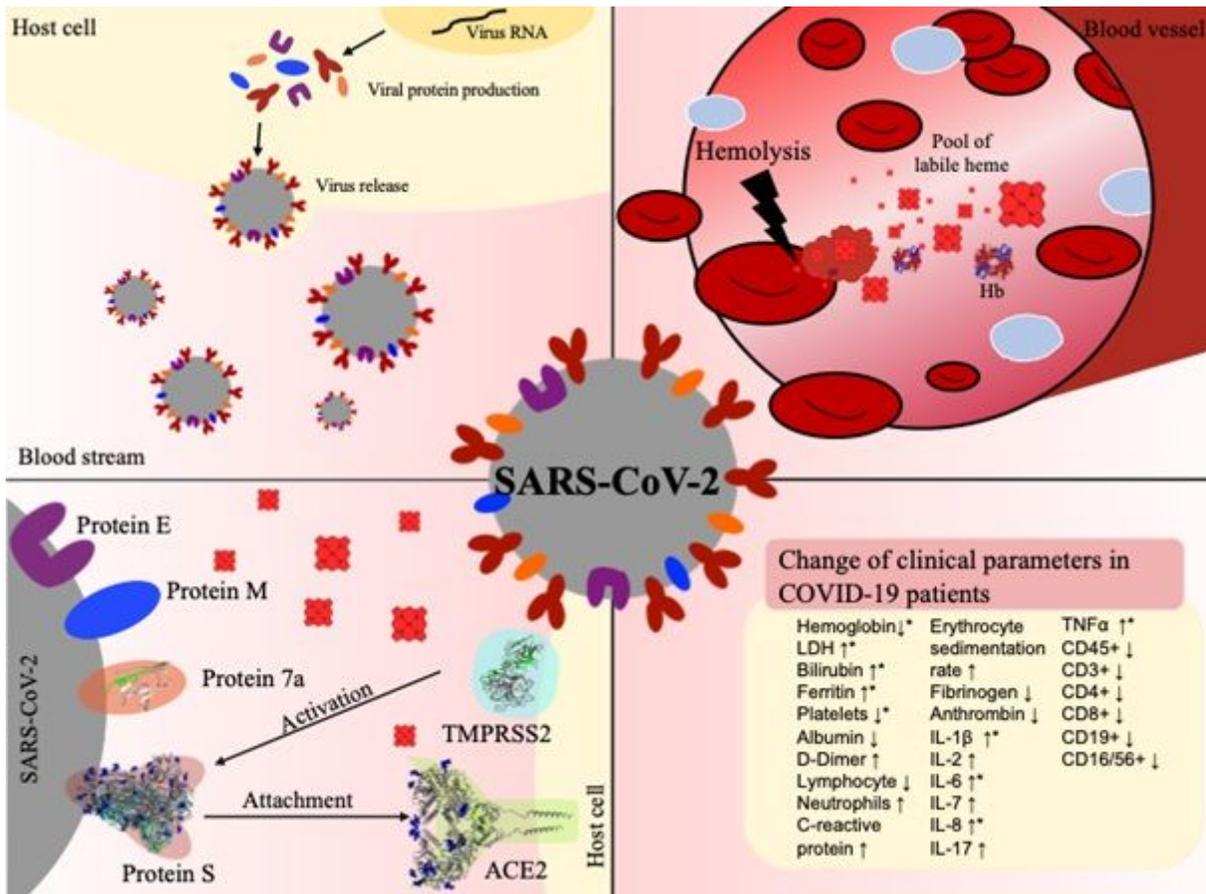


Figure 4

COVID-19 infection and hemolysis show common changes of clinical parameters. Top left: Virus (grey) release after conquering the host cell (yellow) and taking over its protein synthesis machinery. Top right: In case of hemolysis, erythrocyte lysis occurs in a blood vessel, leading to degradation of hemoglobin (Hb) and, thus, to an excess of labile heme. Bottom left: Interaction of virus and host cell before cell entry, involving in particular the viral S protein and protein 7a as well as the human proteins TMPRSS2 and ACE2. Bottom right: Prominent changes of clinical parameters in patients suffering from COVID-19 infection (↑increase, ↓decrease). The terms depicted by an asterisk, have been reported in both, hemolysis and COVID-19 infection [20, 66]. PDBs: 1GZX (hemoglobin), 6VXX (S protein), 6M18 (ACE2), 6W37 (protein 7a).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [HoppetalJCheminfAdditionalfile1.docx](#)
- [HoppetalJCheminfAdditionalfile2.docx](#)
- [HoppetalJCheminfAdditionalfile3.docx](#)
- [HoppetalJCheminfAdditionalfile4.docx](#)

- [HoppetalJCheminfAdditionalfile5.docx](#)
- [HoppetalJCheminfAdditionalfile6.docx](#)
- [HoppetalJCheminfAdditionalfile7.docx](#)