



Molecular Mechanisms Behind Anti SARS-CoV-2 Action of Lactoferrin

Mattia Miotto^{1,2*}, Lorenzo Di Rienzo², Leonardo Bò², Alberto Boffi^{2,3}, Giancarlo Ruocco^{1,2} and Edoardo Milanetti^{1,2}

¹Department of Physics, University of Rome 'La Sapienza', Rome, Italy, ²Istituto Italiano di Tecnologia (IIT), Center for Life Nano Science, Rome, Italy, ³Department of Biochemical Sciences "A. Rossi Fanelli" Sapienza University, Rome, Italy

Despite the huge effort to contain the infection, the novel SARS-CoV-2 coronavirus has rapidly become pandemic, mainly due to its extremely high human-to-human transmission capability, and a surprisingly high viral charge of symptom-less people. While the search for a vaccine is still ongoing, promising results have been obtained with antiviral compounds. In particular, lactoferrin is regarded to have beneficial effects both in preventing and soothing the infection. Here, we explore the possible molecular mechanisms with which lactoferrin interferes with SARS-CoV-2 cell invasion, preventing attachment and/or entry of the virus. To this aim, we search for possible interactions lactoferrin may have with virus structural proteins and host receptors. Representing the molecular iso-electron surface of proteins in terms of 2D-Zernike descriptors, we 1) identified putative regions on the lactoferrin surface able to bind sialic acid present on the host cell membrane, sheltering the cell from the virus attachment; 2) showed that no significant shape complementarity is present between lactoferrin and the ACE2 receptor, while 3) two high complementarity regions are found on the N- and C-terminal domains of the SARS-CoV-2 spike protein, hinting at a possible competition between lactoferrin and ACE2 for the binding to the spike protein.

Keywords: lactoferrin, SARS-CoV-2, shape complementarity, spike (S) glycoprotein, sialic acid, molecular dynamics simulation

OPEN ACCESS

Edited by:

Stefano Piana-Agostinetti,
D. E. Shaw Research, United States

Reviewed by:

Antonella Di Plizio,
Technical University of Munich,
Germany

Matteo Dal Peraro,
École Polytechnique Fédérale de
Lausanne, Switzerland

*Correspondence:

Mattia Miotto
mattia.miotto@roma1.infn.it

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 17 September 2020

Accepted: 11 January 2021

Published: 15 February 2021

Citation:

Miotto M, Di Rienzo L, Bò L, Boffi A,
Ruocco G and Milanetti E (2021)
Molecular Mechanisms Behind Anti
SARS-CoV-2 Action of Lactoferrin.
Front. Mol. Biosci. 8:607443.
doi: 10.3389/fmolb.2021.607443

INTRODUCTION

Lactoferrin (Lf) is a versatile glycoprotein, which plays a key role in many biological functions Redwan et al. (2014). In this work, we focus on the Lf as a crucial player in natural immunity, since it has been proposed to play a strong antiviral activity against a wide range of RNA and DNA viruses Puddu et al. (1998), Andersen et al. (2001), Drobni et al. (2004), Marchetti et al. (2004), and Waarts et al. (2005). Lf is composed of a single chain of about 700 residues folded into two symmetrical lobes. Each lobe possesses a metal-binding site, able to bind iron but also other ions like Cu^{2+} , Zn^{2+} , and Mn^{3+} Baker and Baker (2005) and Giansanti et al. (2016).

This protein is present in saliva, tears, seminal fluid, white blood cells, and milk of mammals Niaz et al. (2019). From its discovery in 1939 Sorensen and Sorensen (1940) and Groves (1960) lactoferrin has been identified as the most important iron-binding protein in milk. Besides, in recent years, lactoferrin has been found involved in a multitude of biological processes. In fact, despite the name, the iron cargo capacity of Lf is not the prominent activity exerted by this molecule. Instead, it performs antioxidant, anti-inflammatory, and anticancer activities Caccavo et al. (2002) and Giansanti et al. (2016), together with a broad antimicrobial action against bacteria and fungi. The latter activity, in particular, is due to Lf's ability to reversibly bind two atoms of iron with high

affinity in the presence of bicarbonate. The iron-free form of Lf, apo-lactoferrin (apoLf), deprives bacteria of iron, thus inhibiting their metabolic activities *in vivo*.

Besides all the aforementioned activities, Lf has been demonstrated to prevent infection of a wide range of diverse viral species van der Strate et al. (2001).

Many viruses make use of glycans as attachment factors. Such glycans are chain-like structures composed of sugar molecules that decorate the surface of cellular constituents, like membrane proteins Varki (2016) and are involved in many cell functions like support, signaling, protein folding, and protection Bishop and Gagneux (2007) and Miotto et al. (2020).

Two important classes of glycans are sialosides (SIA), which contain sialic acid (SA), and glycosaminoglycans, like heparan sulfate (HF). See, for example, Langford-Smith et al. (2015) for more details. When the contact between the virus particle and these receptors is established, the viruses roll toward their specific viral receptor and subsequently enter the host cell, for instance by fusing with the host cell membrane Raman et al. (2016).

While the interaction between Lf and HF has been observed Lang et al. (2011), studies on its interaction with sialic acid derivatives are still missing. On the other hand, Lf has been reported to interact with virus structural proteins, S, M, and E Yi et al. (1997).

In general, depending on the specifics of the virus, lactoferrin prevents infection of the target cell by either 1) interfering with the attachment factor or 2) by binding to host cell molecules that the virus uses as a receptor or co-receptor (competition) or 3) by direct binding to virus particles, as described for herpesvirus Harmsen et al. (1995), polio- and rotavirus Superti et al. (2001) and McCann et al. (2003) and possibly human immunodeficiency virus Puddu et al. (1998). See, for example, Berlutti et al. (2011) for a more detailed discussion.

While we are writing this article, a novel virus, first observed in the autumn of 2019, has rapidly become pandemic. This virus, called SARS-CoV-2, belongs to the coronavirus family and causes a severe acute respiratory syndrome Huang et al. (2020) and Zhu et al. (2020), somewhat similar to those caused by two other coronaviruses, SARS-CoV and MERS-CoV, which crossed species in 2002–2004 Drosten et al. (2003), Ksiazek et al. (2003), and Zaki et al. (2012). In fact, SARS-CoV-2, similarly to SARS-CoV and MERS-CoV, attacks the lower respiratory system, thus provoking viral pneumonia. However, this infection can also lead to effects on the gastrointestinal system, heart, kidney, liver, and central nervous system Su et al. (2016), Prompetchara et al. (2020), and Zhu et al. (2020).

As for SARS-CoV Li et al. (2005a), Li et al. (2005b), and Li (2008) recent *in vivo* experiments confirmed that also SARS-CoV-2 cell entry is mediated by high-affinity interactions between the receptor-binding domain (RBD) of the virus S glycoprotein and the human-host Angiotensin-converting enzyme 2 (ACE2) receptor Zhou et al. (2020). The spike protein is located on the virus envelope and promotes the host attachment and fusion between the viral and cellular membrane. Kuo et al. (2000) and Graham and Baric (2010). Structural studies determined the structures of such protein

both in free form and bound to ACE2 Yan et al. (2020). Further studies investigate the possible interaction of SARS-CoV-2 to sialic acids Milanetti et al. (2020a), Liu et al. (2020), Robson (2020), and Vandelli et al. (2020), or heparan sulfate receptors Liu et al. (2020), both considered involved in SARS-CoV-2 as well as in other coronavirus infections Schwegmann-Weßels and Herrler (2006), Lang et al. (2011), Langford-Smith et al. (2015), Hulswit et al. (2019), and Tortorici et al. (2019).

While the use of Lf as an antiviral against previous coronaviruses infections has been poorly investigated, except for Lang et al. (2011), where evidence of an effect in the attachment process is shown, an increasing number of works hint at the potential antiviral effect of Lactoferrin against the novel SARS-CoV-2 infection Campione et al. (2020), Chang et al. (2020), and de Carvalho et al. (2020).

Here, we computationally investigate the possible molecular mechanisms behind the suggested antiviral action of lactoferrin. In particular, we make use of a recently developed computational protocol based on the 2D Zernike Polynomials, able to rapidly characterize the shape conformation of given protein regions Milanetti et al. (2020a). In this framework using a simple pairwise distance, it is possible to evaluate the similarity between 2 protein pockets or the shape complementarity between the binding regions of 2 interacting proteins.

To assess whether lactoferrin could influence the attachment factors, we investigated the ability of Lt to bind to sialic acid (SIA) or heparan sulfate (HS) receptors Langford-Smith et al. (2015), both considered involved in SARS-CoV-2 infection Milanetti et al. (2020a), Liu et al. (2020), Robson (2020), and Vandelli et al. (2020), as well as to other coronavirus infections Schwegmann-Weßels and Herrler (2006), Lang et al. (2011), Hulswit et al. (2019) and Tortorici et al. (2019).

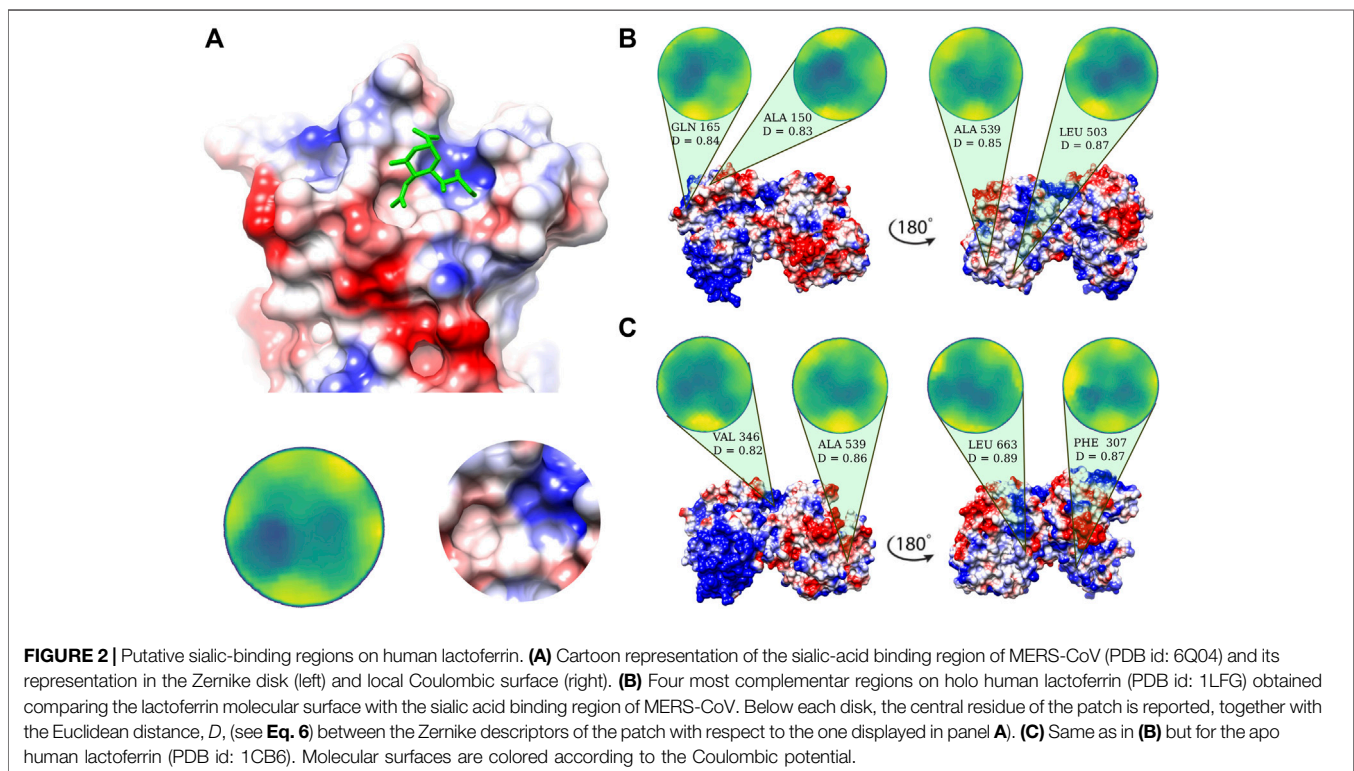
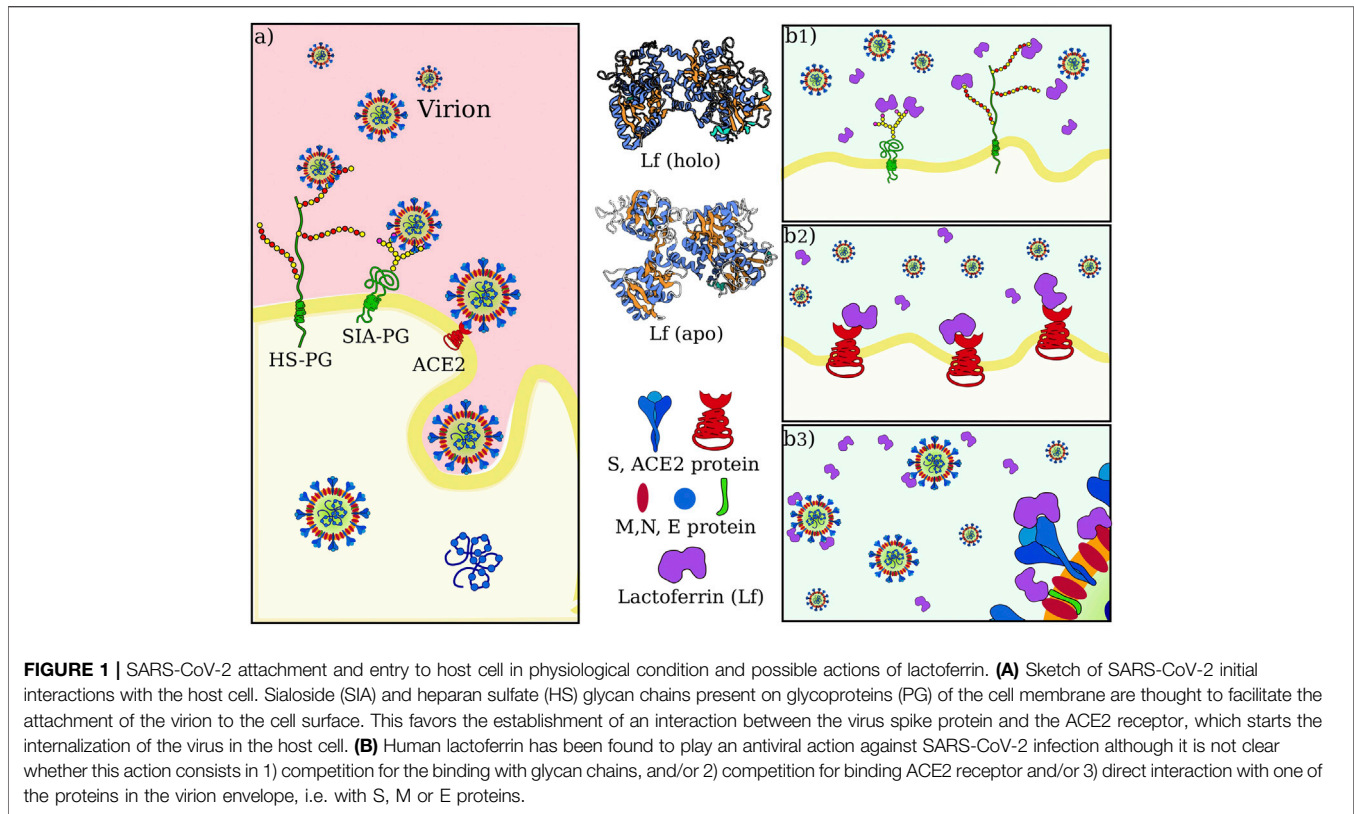
We moreover checked for a possible direct interaction between LF and ACE2 receptor, which could inhibit the interplay between spike and ACE2 binding necessary to the virus infection.

Finally, we investigate the possible interaction between Lf and the three proteins present on the SARS-CoV-2 membrane, i.e. the spike (S), membrane (M), and envelope (E) proteins.

RESULTS

The entry of the virus inside the host cells requires the occurrence of a sequence of molecular interactions. Sialoside (SIA) and/or heparan sulfate (HF) chains mediate the attachment of the virion to the cell surface. Once in the proximity of the cellular receptor (ACE2), SARS-CoV-2 spike protein binds to the receptor and initiates the internalization process. **Figure 1A** shows a sketch of the mechanism.

The observed antiviral action of Lf may consist of interference in one or more of those steps. We thus investigate in the next three sections the possibility of direct binding between lactoferrin and sialic acid, ACE2, and SARS-CoV-2 Spike protein, respectively (see **Figure 1B**).



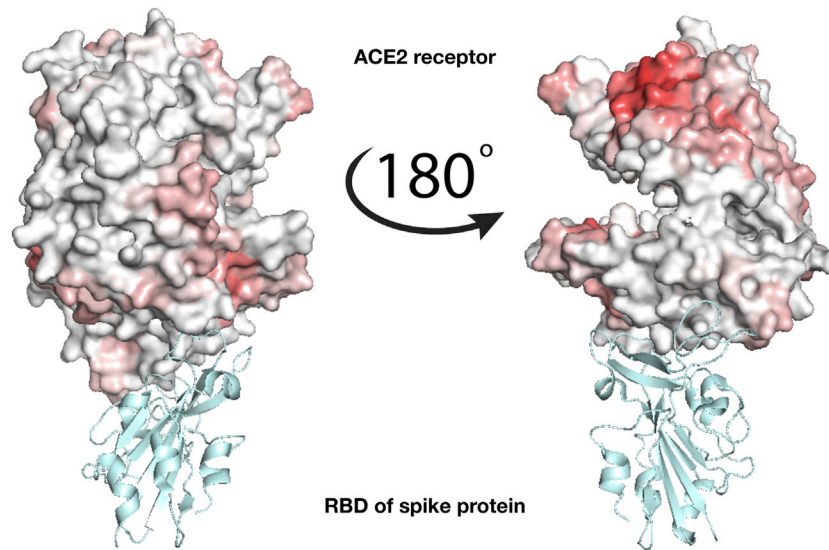


FIGURE 3 | Analysis of the binding between lactoferrin and the ACE2 receptor. Molecular representation of the experimentally solved complex of the ACE2 receptor with SARS-CoV-2 spike receptor-binding domain, RBD (PDB id: 6m17). ACE2 molecular surface is colored according to its binding propensity to bind lactoferrin regions. Dark red indicates high binding propensity while white means no interaction. The spike RBD is colored in cyan.

Interaction with Sialic Acid

Possible binding regions for sialic acid, the terminal molecule of the SIA chains on human lactoferrin are investigated on the basis of the procedure described in Milanetti et al. (2020a), i.e. we select the portion of the molecular surface of the MERS-CoV spike protein in interaction with sialic acid, experimentally solved in Park et al. (2019) (Figure 2A), and we search for similar patches on the Lf molecular surface. Within the same strategy, we search for similar patches on the surface of human lactoferrin (see *Methods* for more details).

In the 2D Zernike framework, the geometrical shape of a protein surface patch is compactly summarized in a set of ordered numerical descriptors, whose number - 121 in our case - modulate the detail of the description. Dealing with ordered numerical descriptors, the comparison between different protein patches can be performed with a Euclidean distance. Figures 2B,C show the four most geometrically similar patches identified in the Apo and Holo form of Lf. An *a posteriori* check of the electrostatic potential on the patches, allows us to select only some of the possible solutions identified based on the shape comparison analysis, i.e. the ones having also a similar electrostatic surface with the SIA binding site on the MERS-CoV spike.

The region on the Lf surface identified as the most similar to the MERS region interacting with sialic acid, both in shape and in electrostatics, is the one centered on VAL 346.

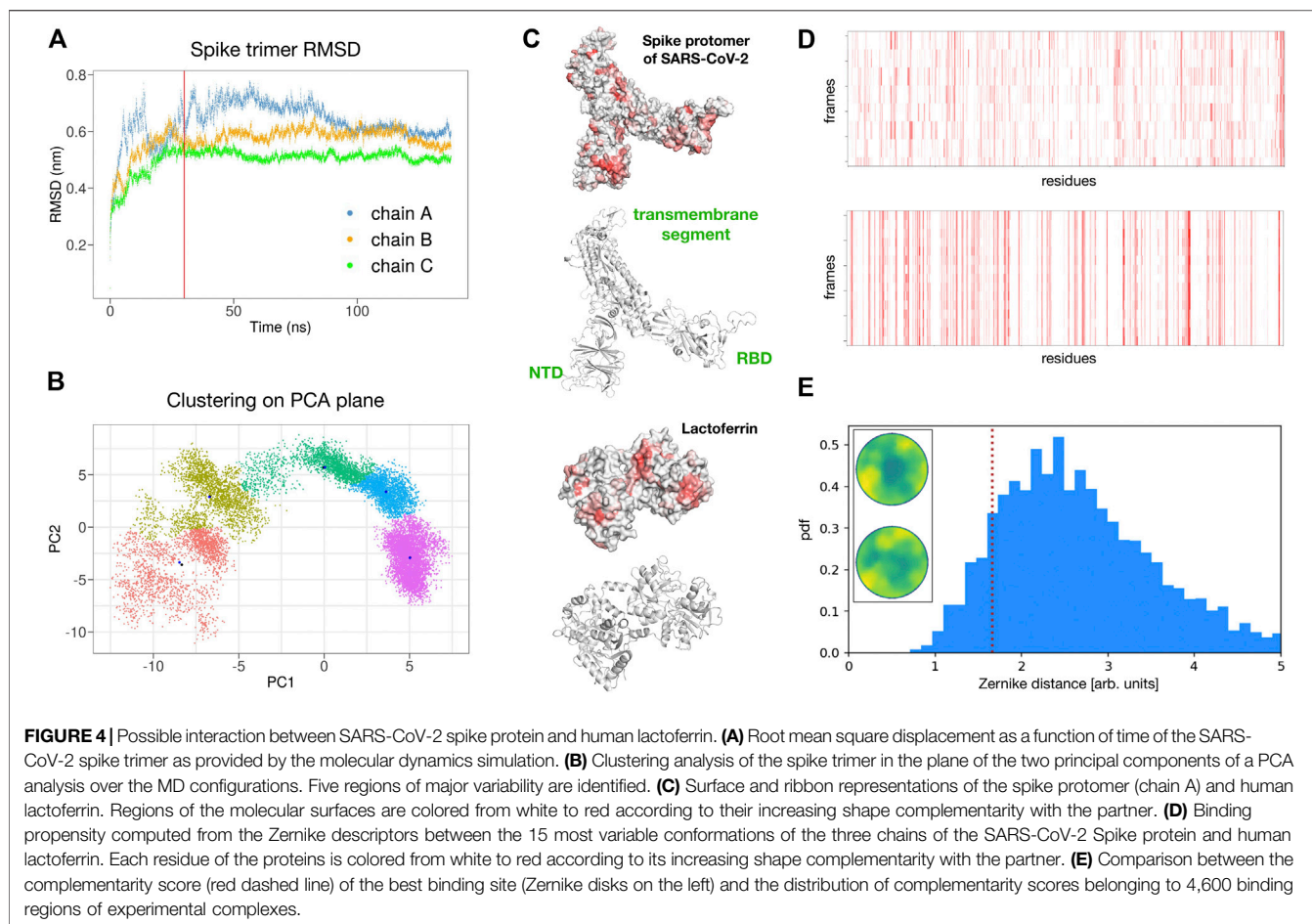
Interaction with ACE2

To check whether the action of lactoferrin can be ascribed to a competition with the virion spike proteins in binding directly the ACE2 receptor (Figure 1), we performed a blind search of the molecular surfaces of both ACE2 and human Lf to identify possible binding regions having a meaningful shape

complementarity. Under this hypothesis, if the interaction between the Lf and the ACE2 receptor occurs, Lf could hinder the molecular binding between the spike protein of SARS-CoV-2 and the corresponding ACE2 receptor. Figure 3 shows the molecular surface of the ACE2 receptor colored according to its propensity to bind regions of the Lf protein. The redder the region, the greater the shape complementarity between that region and another one found on the surface of the putative molecular partner, i.e. holo lactoferrin. As one can see from Figure 3, a complementary region is indeed present, however, it is located far from the binding site of the spike (grey in the figure) and in a part of ACE2 that looks toward the membrane. To better visualize the result, we have represented two points of view of the binding between spike and ACE2, one rotated 180° with respect to the other. On the other hand, we can see that the ACE2 region interacting with the SARS-CoV-2 spike protein has no low shape complementarity with Lf regions.

Interaction with Virion Membrane Proteins

A third possible mechanism at the basis of the observed antiviral activity of Lf could be ascribed to a direct interaction with the membrane proteins present on the virion envelope. In particular, SARS-CoV-2 presents three different kinds of proteins on its membrane, i.e. S, M, and E proteins Bianchi et al. (2020) and Seah et al. (2020). While the 3D structure of the S protein has been determined - even if some loop regions in the S1 sub-unit are not solved - unfortunately, no structures are available for the E and M proteins. Thus, the molecular surfaces of Lf were compared with those of the three proteins in order to check whether an interaction with lactoferrin is possible. For this analysis, we adopted the same computational procedure used in the previous paragraph.



For both S, M, and E proteins, we sample their whole molecular surface and compared all the possible patches with those of Lf. In this way, all molecular surfaces, both membrane proteins, and Lf ones are colored according to the corresponding binding propensity.

E and M presented a possible region of interactions located in the intra-membrane region (data not shown). The most robust and relevant result of this analysis regards the compatibility between the spike and lactoferrin. According to our findings, Lf presents two regions of high complementarity with one portion of the C-terminal domain of the spike S1 subunit and another located in the N-terminal part.

To test the reliability of the found signals, we performed a molecular dynamics simulation of the spike trimer (see *Methods* for details) and sampled five configurations for each of the three chains at equilibrium (see **Figure 4A**).

In particular, for each chain, we performed a Principal Component Analysis (PCA) on the frames of the dynamics and thus projected them on the plane identified by the two principal components. Upon clustering these points we obtain five subgroups. For each subgroup, we extract the centroidal configuration. Since distant points in the PCA plane correspond to the different 3D structure, picking one point from each identified cluster assured that the selected configurations have

high structural differences between them in the explored configuration space (see **Figure 4B**).

Remarkably, repeating the blind search for complementarity regions on the 15 surfaces of the extracted spike monomers, we found conservation of the signal over the conformational noise. **Figure 4C**) shows the identified regions and their conservation in the different sampled frames.

Among the found regions, the one involving the spike C-terminal domain is the one with higher shape complementarity. In particular, according to our method, human lactoferrin could interact with the spike protein with the surface region centered in residue ALA 539. Alternatively, the spike protein may interact with the Lf protein using a molecular patch centered in the residue PHE 490. Since the spike is usually heavily glycosylated (see for example *Casalino et al. (2020)* for details), we checked whether the identified region could be hidden by glycans. The closest glycosylation site (ASN 343) has a distance of 34 Å from the found region, thus the interaction with Lt is not expected to resent from the glycan shield. furthermore, it must be pointed out that the complementarity achieved by the identified patches is comparable with those of experimentally solved complexes. Indeed, analyzing the shape complementarity of over 4,600 x-ray protein-protein complexes (see *Methods* for

details), we have the distribution shown in **Figure 4D**, where the lower the distance the higher the complementarity of the binding region. The red dotted line shows the complementarity found between the spike and Lt. As it is evident the proposed patch is characterized by values of shape complementarity typical of experimental complexes.

Finally, to further support this result with an independent and external methodology to our approach, we performed a completely blind molecular docking analysis between the spike protein and the Lf protein. To this end, the Zdock server was used as a state of the art of molecular docking software [Chen et al. \(2003\)](#). As per default settings, only the first 10 docking poses have been selected and the predicted contacts analyzed. In particular, five out of ten poses show bindings involving the spike region involve the region we found with our protocol when the residues are defined in contact if their C-alpha atoms have a distance less than 8 Å.

DISCUSSION AND CONCLUSIONS

In the last decade, the Zernike formalism has been widely applied for the characterization of molecular surfaces [Venkatraman et al. \(2009\)](#), [Kihara et al. \(2011\)](#), [Di Rienzo et al. \(2017\)](#), [Daberduku and Ferrari \(2019\)](#), and [Di Rienzo et al. \(2020\)](#).

Very recently, we developed a new representation, based on the 2D Zernike polynomials, which allows an extremely efficient, fast, and completely unsupervised description of the local geometrical shape, allowing for easy comparison between different regions of molecules. Through this compact description, it is possible both to analyze the similarity between two different regions - suggesting, for example, a similar ligand for binding regions - and to study the complementarity between interacting surfaces [Milanetti et al. \(2020a\)](#).

Here, we used our novel method to shed some light on the molecular mechanisms that may support the antiviral action of human lactoferrin against SARS-CoV-2 infection. In particular, we focused on the early stages of the infection, i.e. the attachment and entry of the virus to the host cell, when lactoferrin can interfere with the virus-host interaction without the need to be internalized in the cell. We thus tried to establish whether lactoferrin could compete with the virus in binding to sialic acid, the sticky end of sialoside chains, which has been suggested to mediate the attachment of SARS-CoV-2 to the host cell. Interestingly, comparing the binding region of sialic acid in the MERS-CoV coronavirus with patches on the lactoferrin surface, we found possible spots on both apo and holo forms of Lf, which could compete in forming low affinity but high avidity interactions.

We then proceeded to test the hypothesis of an interaction between lactoferrin and the primary SARS-CoV-2 protein receptor, ACE2. A blind search for complementarity regions highlighted a hot-spot in a region that in physiological conditions is oriented toward the membrane, while no significative complementarity is present in the ACE2 region involved in the interaction with the virus spike protein. At

last, we analyzed the three membrane proteins on the virus envelope, i.e. the E, M, and S ones. Similarly to ACE-2, both E and M presented possible interacting regions in portions of the surface, that are buried in the virion membrane under normal conditions (data not shown). On the other hand, the spike protein showed two main hot spots, one in the N-terminal domain of the S1 subunit and another in the C-terminal one. Those two regions are robust to molecular noise, as the signal endures using different configurations sampled from a molecular dynamics simulation and each of the three chains of the trimer. Notably, the most complementary region is the one in the C-terminal region, the one involved in the spike-ACE2 interaction. Notably, this region is far from known spike glycosylation sites and therefore even more available for the binding. This aspect is quite important if we consider that the spike is usually heavily glycosylated, with a large fraction of the protein surface buried under a glycan shield [Casalino et al. \(2020\)](#). Thus our finding suggests a possible competition between ACE2 and lactoferrin for the binding of the SARS-CoV-2 spike, which may explain the observed antiviral action.

MATERIALS AND METHODS

Datasets

The protein, whose structures are analyzed in this paper are:

- Human lactoferrin, in the apo (PDB id: 1CB6) and holo (PDB id: 1LFG) forms.
- ACE2, in its apo state (PDB id: 1R42).
- Unbound SARS-CoV-2 spike protein: modeled by I-TASSER server [Roy et al. \(2010\)](#).

In a nutshell, the used server works in three steps: 1) starting from the submitted amino acid sequence, it first tries to retrieve template proteins of similar folds (or super-secondary structures) from the PDB library. Then, 2) the continuous fragments excised from the PDB templates are reassembled into full-length models. In cases where no appropriate template is identified, I-TASSER builds the whole structure by ab initio modeling. Finally, 3) the algorithm refines the models by removing steric clashes as well as checking the global topology of the found models, optimizing the hydrogen-bonding network.

The 10 structural analogs (as identified by TM-align [Zhang and Skolnick \(2005\)](#)) found by I-tasser are PDB id: 5X58, 6NZK, 6NB3, 3JCL, 5I08, 6CV0, 5SZS, 5WRG, 6UTK, 6B7N. The first (out of five) model returned by the server has been used as a starting conformation for the molecular dynamics simulation. We note that the obtained model has an average RMSD of ~ 1.2Å with recently published experimental structures (see e.g. PDB id: 6VSB, 6VXX, and 6VYB).

• SARS-CoV-2 M protein, modeled using I-Tasser [Roy et al. \(2010\)](#). The 10 structural analogs (as identified by TM-align [Zhang and Skolnick \(2005\)](#)) are PDB id: 3W6Q, 4OUA, 4HE8, 4YU5, 1JS8, 3RKO, 2Y9W, 1TJ7, 2RF7, 4BED. The first (out of five) model returned by the server has been used for the analyses.

Additionally, we compared the structure of the M protein provided by I-tasser, with the one recently predicted by AlphaFold [Senior et al. \(2020\)](#). We found that while almost all

the secondary structures are correctly predicted by both algorithms, the overall fold of the protein is better reproduced in the AlphaFold model, even if it lacks the first 10 residues of the N-terminal domain, which is the part of the M protein that protrudes from the virion membrane Thomas (2020).

- SARS-CoV-2 E protein, modeled using I-Tasser Roy et al. (2010). Here, we directly provided the experimentally solved SARS-CoV E protein (PDB id: 2mm4) as a template for homology modeling. The first (out of five) model returned by the server has been used for the analyses. Notably, comparing the model structure with the recently solved trans-membrane region of SARS-CoV-2 E protein (PDB id: 7k3g), we obtained an RMSD of 2.6 Å.

To set a reference for the measured complementarities, a dataset of protein-protein complexes experimentally solved in x-ray crystallography is taken from Gainza et al. (2020). We only selected pair interactions regarding chains with more than 50 residues. The Protein-Protein dataset is therefore composed of 4605 complexes. For each complex, the binding region is identified as the portions of the two protein molecular surfaces distant less than 3 Å.

Computation of Molecular Surfaces

For each protein of the dataset (x-ray structure in PDB format Berman et al. (2003)), we use DMS Richards (1977) to compute the solvent-accessible surface, using a density of 5 points per Å² and a water probe radius of 1.4 Å. The unit normals vector, for each point of the surface, was calculated using the flag $-n$.

Patch Definition and Complementarity Evaluation

A molecular surface is represented by a set of points in the three-dimensional space. We define a surface patch, as the group of points that fall within a sphere of radius $R_s = 6\text{Å}$, centered on one point of the surface. Once the patch is selected,

- we fit a plane that passes through the points and reorient the patch in such a way to have the z-axis perpendicular to the plane and going through the center of the plane.
- we define the angle θ as the largest angle between the perpendicular axis and a secant connecting a given point C on the z-axis to any point of the patch. C is then set in order that $\theta = 45^\circ$. r is the distance between C and a surface point.
- build a square grid and associate each pixel with the mean r of the points inside it. This 2D function can be expanded on the basis of the Zernike polynomials (see next section).

Once a patch is represented in term of its Zernike descriptors, the similarity between that patch and another one can be simply measured as the Euclidean distance between the invariant vectors. The relative orientation of the patches before the projection in the unitary circle must be considered. In fact, if we search for similar regions we must compare patches that have the same orientation once projected in the 2D plane, i.e. the solvent-exposed part of the surface must be oriented in the same direction for both patches, for example as the positive z-axis. If instead, we want to assess the complementarity between two patches, we must orient the patches contrariwise, i.e. one patch with the solvent-exposed

part toward the positive z-axis (“up”) and the other toward the negative z-axis (“down”).

2D Zernike Polynomials and Invariants

Each function of two variables, $f(r, \phi)$ (polar coordinates) defined inside the region $r < 1$ (unitary circle), can be decomposed in the Zernike basis as

$$f(r, \phi) = \sum_{n=0}^{\infty} \sum_{m=0}^{m=n} c_{nm} Z_{nm} \quad (1)$$

with

$$c_{nm} = \frac{(n+1)}{\pi} \langle Z_{nm} | f \rangle = \frac{(n+1)}{\pi} \int_0^1 dr r \int_0^{2\pi} d\phi Z_{nm}^*(r, \phi) f(r, \phi). \quad (2)$$

being the expansion coefficients, while the complex functions, $Z_{nm}(r, \phi)$ are the Zernike polynomials. Each polynomial is composed by a radial and an angular part,

$$Z_{nm} = R_{nm}(r) e^{im\phi}. \quad (3)$$

where the radial part for any n and m , is given by

$$R_{nm}(r) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} r^{n-2k} \quad (4)$$

Since for each couple of polynomials, the following relation holds

$$\langle Z_{nm} | Z_{n'm'} \rangle = \frac{\pi}{(n+1)} \delta_{nn'} \delta_{mm'} \quad (5)$$

the complete set of polynomials forms a basis and knowing the set of complex coefficients, $\{c_{nm}\}$ allows for a univocal reconstruction of the original image (with a resolution that depends on the order of expansion, $N = \max(n)$). Since the modulus of each coefficient ($z_{nm} = |c_{nm}|$) does not depend on the phase, i.e. it is invariant for rotations around the origin of the unitary circle, the shape similarity between two patches can be assessed by comparing the Zernike invariants of their associated 2D projections. In particular, we measured the similarity between patch i and j as the Euclidean distance between the invariant vectors, i.e.

$$d_{ij} = \sqrt{\sum_{k=1}^{M=121} (z_i^k - z_j^k)^2} \quad (6)$$

Detection of Similar Regions

The search for possible regions of interaction with sialic acid on the lactoferrin surface has been performed similarly to what done in Milanetti et al. (2020a). First, we identified the interaction region of sialic acid on the MERS-CoV surface. Next, we isolated the surface patch selecting all surface points within a sphere of radius 6 Å centered around residue I-132 of the N-terminal domain of MERS-CoV spike (see Figure 2A) and computed the corresponding Zernike descriptors. Then, we repeated a similar procedure using each point of the lactoferrin surfaces

(both apo and holo forms) as centers for the sphere. Finally, we evaluate the Euclidean distance between the Zernike descriptors of the MERS-CoV disk (reference patch) and those associated with the Lt surface points. All patches were studied in the “up” orientation (see previous sections). In **Figures 2B,C**, we show the four regions having lower Euclidean distance with respect to the reference patch.

Detection of Complementar Regions

To evaluate whether two proteins present regions with high shape complementarity (and thus of possible interaction), we first computed the Zernike descriptors of the patches centered in all the points of the two surfaces, choosing for one protein the “up” orientation and for the other the “down” orientation. Next, we computed the Euclidean distance between all possible couples of patches among the two proteins. We thus associate to each point of the two surfaces the minimum distance value observed - *the binding propensity* - between the considered point and all points of the other surface. After all surface points are associated with their binding propensity, we performed a smoothing process to highlight the signal in specific regions characterized mostly by low distance values. In this process each point is associated with the mean value of the points in its neighborhood: the basic idea is that the interacting region should be made up mostly of elements with high complementarity and therefore a high average value of binding propensity values (see Milanetti et al. (2020b, a) for further details). For both the patch definition and the smoothing process we adopted a sphere radius of 6 Å.

Molecular Dynamics Simulations

The starting structure of the SARS-CoV-2 spike trimeric complex was taken from the model structure proposed by the I-Tasser server Roy et al. (2010). All steps of the simulation were performed using Gromacs 2019.3 Spoel et al. (2005). Topologies of the system were built using the CHARMM-27

force field Brooks et al. (2009). The protein was placed in a dodecahedral simulation box, with periodic boundary conditions, filled with 131793 TIP3P water molecules Jorgensen et al. (1983). We checked that each atom of the trimer was at least at a distance of 1.1 nm from the box borders. The addition of three sodium counterions rendered the systems electroneutral. The final system, consisting of 448,572 atoms, was first minimized with 2,064 steps of steepest descent. Relaxation of water molecules and thermalization of the system in NVT and NPT environments were run each for 0.1 at 2 fs time-step. The temperature was kept constant at 300 K with v-rescale algorithm Bussi et al. (2007); the final pressure was fixed at one bar with the Parrinello-Rahman algorithm Parrinello and Rahman (1980) which guarantees a water density of 1,004 kg/m³, close to the experimental value. LINCS algorithm Hess et al. (1997) was used to constraint h-bonds.

Finally, the systems were simulated with a 2 fs time-step for 140 ns in periodic boundary conditions, using a cut-off of 12 Å for the evaluation of short-range non-bonded interactions and the Particle Mesh Ewald method Cheatham et al. (1995) for the long-range electrostatic interactions.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

EM, AB, and GR conceived research; MM and LB performed molecular dynamics simulations. MM, LR, and EM performed calculations and computational analysis. All authors analyzed results; all authors wrote and revised the paper.

REFERENCES

- Andersen, J. H., Osbakk, S. A., Vorland, L. H., Traavik, T., and Gutteberg, T. J. (2001). Lactoferrin and cyclic lactoferricin inhibit the entry of human cytomegalovirus into human fibroblasts. *Antiviral. Res.* 51, 141–149. doi:10.1016/s0166-3542(01)00146-2
- Baker, E. N., and Baker, H. M. (2005). Lactoferrin. *Cell. Mol. Life Sci.* 62, 2531. doi:10.1007/s00018-005-5368-9
- Berluti, F., Pantanella, F., Natalizi, T., Frioni, A., Paesano, R., Polimeni, A., et al. (2011). Antiviral properties of lactoferrin—a natural immunity molecule. *Molecules* 16, 6992–7018. doi:10.3390/molecules16086992
- Berman, H. M., Bourne, P. E., Westbrook, J., and Zardecki, C. (2003). “The protein data bank.” in *Protein Structure*, (Boca Raton, FL: CRC Press). 394–410.
- Bianchi, M., Benvenuto, D., Giovanetti, M., Angeletti, S., Ciccozzi, M., and Pascarella, S. (2020). Sars-CoV-2 envelope and membrane proteins: structural differences linked to virus characteristics? *BioMed. Res. Int.* 2020, 1–6. doi:10.1155/2020/4389089
- Bishop, J. R., and Gagneux, P. (2007). Evolution of carbohydrate antigens—microbial forces shaping host glycomes? *Glycobiology* 17, 23R–34R. doi:10.1093/glycob/cwm005
- Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., et al. (2009). CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30, 1545–614. doi:10.1002/jcc.21287
- Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126, 014101. doi:10.1063/1.2408420
- Caccavo, D., Pellegrino, N. M., Altamura, M., Rigon, A., Amati, L., Amoroso, A., et al. (2002). Antimicrobial and immunoregulatory functions of lactoferrin and its potential therapeutic application. *J. Endotoxin. Res.* 8, 403–417. doi:10.1179/096805102125001000
- Campione, E., Cosio, T., Rosa, L., Lanna, C., Di Girolamo, S., Gaziano, R., et al. (2020). Lactoferrin as protective natural barrier of respiratory and intestinal mucosa against coronavirus infection and inflammation. *Ijms* 21, 4903. doi:10.3390/ijms21144903
- Casalino, L., Gaieb, Z., Goldsmith, J. A., Hjorth, C. K., Dommer, A. C., Harbison, A. M., et al. (2020). Beyond shielding: The roles of glycans in the sars-cov-2 spike protein. *ACS Cent. Sci.* 6, 1722–1734. doi:10.1021/acscentsci.0c01056
- Chang, R., Ng, T. B., and Sun, W. Z. (2020). Lactoferrin as potential preventative and adjunct treatment for COVID-19. *Int. J. Antimicrob. Agents* 56, 106118. doi:10.1016/j.ijantimicag.2020.106118
- Cheatham, T. E. I., Miller, J. L., Fox, T., Darden, T. A., and Kollman, P. A. (1995). Molecular dynamics simulations on solvated biomolecular systems: The particle mesh ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am. Chem. Soc.* 117, 4193–4194. doi:10.1021/ja00119a045
- Chen, R., Li, L., and Weng, Z. (2003). ZDOCK: an initial-stage protein-docking algorithm. *Proteins* 52, 80–87. doi:10.1002/prot.10389

- Daberdaku, S., and Ferrari, C. (2019). Antibody interface prediction with 3d zernike descriptors and svm. *Bioinformatics* 35, 1870–1876. doi:10.1093/bioinformatics/bty918
- de Carvalho, C. A. M., da Rocha Matos, A., Caetano, B. C., de Sousa Junior, I. P., da Costa Campos, S. P., Geraldino, B. R., et al. (2020). In vitro inhibition of SARS-CoV-2 infection by bovine lactoferrin [Preprint]. *bioRxiv*. doi:10.1101/2020.05.13.093781
- Di Rienzo, L., Milanetti, E., Alba, J., and D'Abramo, M. (2020). Quantitative characterization of binding pockets and binding complementarity by means of zernike descriptors. *J. Chem. Inf. Model.* 60, 1390–1398. doi:10.1021/acs.jcim.9b01066
- Di Rienzo, L., Milanetti, E., Lepore, R., Olimpieri, P. P., and Tramontano, A. (2017). Superposition-free comparison and clustering of antibody binding sites: implications for the prediction of the nature of their antigen. *Sci. Rep.* 7, 45053. doi:10.1038/srep45053
- Drobni, P., Näslund, J., and Evander, M. (2004). Lactoferrin inhibits human papillomavirus binding and uptake in vitro. *Antiviral. Res.* 64, 63–68. doi:10.1016/j.antiviral.2004.05.005
- Drosten, C., Günther, S., Preiser, W., Van Der Werf, S., Brodt, H. R., Becker, S., et al. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* 348, 1967–1976. doi:10.1056/NEJMoa030747
- Gainza, P., Sverrisson, F., Monti, F., Rodolà, E., Boscaini, D., Bronstein, M., et al. (2020). Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nat. Methods.* 17, 184–192. doi:10.1038/s41592-019-0666-6
- Giansanti, F., Panella, G., Leboffe, L., and Antonini, G. (2016). Lactoferrin from milk: Nutraceutical and pharmacological properties. *Pharmaceuticals (Basel)* 9, 61. doi:10.3390/ph9040061
- Graham, R. L., and Baric, R. S. (2010). Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J. Virol.* 84, 3134–46. doi:10.1128/JVI.01394-09
- Groves, M. L. (1960). The isolation of a red protein from milk. *J. Am. Chem. Soc.* 82, 3345–3350. doi:10.1021/ja01498a029
- Harmsen, M. C., Swart, P. J., de Béthune, M.-P., De Clercq, E., The, T. H., et al. (1995). Antiviral effects of plasma and milk proteins: Lactoferrin shows potent activity against both human immunodeficiency virus and human cytomegalovirus replication in vitro. *J. Infect. Dis.* 172, 380–388. doi:10.1093/infdis/172.2.380
- Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997). LINC: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18, 1463–1472. doi:10.1002/(sici)1096-987x(199709)18:12<1463::aid-jcc4>3.0.co;2-h
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *Lancet* 395, 497–506. doi:10.1016/s0140-6736(20)30183-5
- Hulswit, R. J. G., Lang, Y., Bakkers, M. J. G., Li, W., Li, Z., Schouten, A., et al. (2019). Human coronaviruses oc43 and hku1 bind to 9-o-acetylated sialic acids via a conserved receptor-binding site in spike protein domain a. *Proc. Natl. Acad. Sci. USA* 116, 2681–2690. doi:10.1073/pnas.1809667116
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi:10.1063/1.445869
- Kihara, D., Sael, L., Chikhi, R., and Esquivel-Rodriguez, J. (2011). Molecular surface representation using 3d zernike descriptors for protein shape comparison and docking. *Curr. Protein Pept. Sci.* 12, 520–530. doi:10.2174/138920311796957612
- Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., et al. (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* 348, 1953–66. doi:10.1056/NEJMoa030781
- Kuo, L., Godeke, G. J., Raamsman, M. J., Masters, P. S., and Rottier, P. J. (2000). Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: crossing the host cell species barrier. *J. Virol.* 74, 1393–406. doi:10.1128/jvi.74.3.1393-1406.2000
- Lang, J., Yang, N., Deng, J., Liu, K., Yang, P., Zhang, G., et al. (2011). Inhibition of SARS pseudovirus cell entry by lactoferrin binding to heparan sulfate proteoglycans. *PLoS One* 6, e23710. doi:10.1371/journal.pone.0023710
- Langford-Smith, A., Day, A. J., Bishop, P. N., and Clark, S. J. (2015). Complementing the sugar code: Role of GAGs and sialic acid in complement regulation. *Front Immunol* 6, 25. doi:10.3389/fimmu.2015.00025
- Li, F. (2008). Structural analysis of major species barriers between humans and palm civets for severe acute respiratory syndrome coronavirus infections. *J. Virol.* 82, 6984–6991. doi:10.1128/JVI.00442-08
- Li, F., Li, W., Farzan, M., and Harrison, S. C. (2005a). Structure of sars coronavirus spike receptor-binding domain complexed with receptor. *Science* 309, 1864–1868. doi:10.1126/science.1116480
- Li, W., Zhang, C., Sui, J., Kuhn, J. H., Moore, M. J., Luo, S., et al. (2005b). Receptor and viral determinants of sars-coronavirus adaptation to human ace2. *EMBO. J.* 24, 1634–1643. doi:10.1038/sj.emboj.7600640
- Liu, L., Chopra, P., Li, X., Wolfert, M. A., Tompkins, S. M., and Boons, G.-J. (2020). SARS-CoV-2 spike protein binds heparan sulfate in a length- and sequence-dependent manner [Preprint]. *bioRxiv*. doi:10.1101/2020.05.10.087288
- Marchetti, M., Trybala, E., Superti, F., Johansson, M., and Bergström, T. (2004). Inhibition of herpes simplex virus infection by lactoferrin is dependent on interference with the virus binding to glycosaminoglycans. *Virology* 318, 405–13. doi:10.1016/j.virol.2003.09.029
- McCann, K., Lee, A., Wan, J., Roginski, H., and Coventry, M. (2003). The effect of bovine lactoferrin and lactoferricin b on the ability of feline calicivirus (a norovirus surrogate) and poliovirus to infect cell cultures. *J. Appl. Microbiol.* 95, 1026–1033. doi:10.1046/j.1365-2672.2003.02071.x
- Milanetti, E., Miotto, M., Di Rienzo, L., Monti, M., Gosti, G., and Ruocco, G. (2020a). In-silico evidence for two receptors based strategy of sars-cov-2. *arXiv, arXiv:2003*. doi:10.1101/2020.03.24.006197
- Milanetti, E., Miotto, M., Rienzo, L. D., Monti, M., Gosti, G., and Ruocco, G. (2020b). 2d zernike polynomial expansion: finding the protein-protein binding regions. *Comput. Struct. Biotechnol. J.* 19, 29–36. doi:10.1016/j.csbj.2020.11.051
- Miotto, M., Rienzo, L. D., Gosti, G., Milanetti, E., and Ruocco, G. (2020). Does blood type affect the covid-19 infection pattern? *arXiv:2007.06296*
- Niaz, B., Saeed, F., Ahmed, A., Imran, M., Maan, A. A., Khan, M. K. I., et al. (2019). Lactoferrin (LF): a natural antimicrobial protein. *International Journal of Food Properties* 22, 1626–1641. doi:10.1080/10942912.2019.1666137
- Park, Y.-J., Walls, A. C., Wang, Z., Sauer, M. M., Li, W., Tortorici, M. A., et al. (2019). Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. *Nat. Struct. Mol. Biol.* 26, 1151–1157. doi:10.1038/s41594-019-0334-7
- Parrinello, M., and Rahman, A. (1980). Crystal structure and pair potentials: A molecular-dynamics study. *Phys. Rev. Lett.* 45, 1196–1199. doi:10.1103/physrevlett.45.1196
- Prompetchara, E., Ketloy, C., and Palaga, T. (2020). Immune responses in covid-19 and potential vaccines: Lessons learned from sars and mers epidemic. *Asian Pac. J. Allergy Immunol.* 38, 1–9. doi:10.12932/AP-200220-0772
- Puddu, P., Borghi, P., Gessani, S., Valenti, P., Belardelli, F., and Seganti, L. (1998). Antiviral effect of bovine lactoferrin saturated with metal ions on early steps of human immunodeficiency virus type 1 infection. *Int. J. Biochem. Cell Biol.* 30, 1055–1062. doi:10.1016/s1357-2725(98)00066-1
- Raman, R., Tharakaraman, K., Sasisekharan, V., and Sasisekharan, R. (2016). Glycan-protein interactions in viral pathogenesis. *Curr. Opin. Struct. Biol.* 40, 153–162. doi:10.1016/j.sbi.2016.10.003
- Redwan, E. M., Uversky, V. N., El-Fakharany, E. M., and Al-Mehdar, H. (2014). Potential lactoferrin activity against pathogenic viruses. *C R Biol.* 337, 581–595. doi:10.1016/j.crv.2014.08.003
- Richards, F. M. (1977). Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* 6, 151–176. doi:10.1146/annurev.bb.06.060177.001055
- Robson, B. (2020). Bioinformatics studies on a function of the sars-cov-2 spike glycoprotein as the binding of host sialic acid glycans. *Comput. Biol. Med.* 122, 103849. doi:10.1016/j.compbiomed.2020.103849
- Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi:10.1038/nprot.2010.5
- Schwegmann-Weßels, C., and Herler, G. (2006). Sialic acids as receptor determinants for coronaviruses. *Glycoconj J* 23, 51–58. doi:10.1007/s10719-006-5437-9
- Seah, I., Su, X., and Lingam, G. (2020). Revisiting the dangers of the coronavirus in the ophthalmology practice. *Eye (Lond)* 34, 1155–1157. doi:10.1038/s41433-020-0790-7
- Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., et al. (2020). Improved protein structure prediction using potentials from deep learning. *Nature* 577, 706–710. doi:10.1038/s41586-019-1923-7

- Sorensen, M., and Sorensen, S. (1940). The proteins in whey. *Compte rendu des Travaux du Laboratoire de Carlsberg, Ser. Chim.* 23, 55–99. doi:10.3168/jds.s0022-0302(40)95543-6
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26, 1701–1718. doi:10.1002/jcc.20291
- Su, S., Wong, G., Shi, W., Liu, J., Lai, A. C., Zhou, J., et al. (2016). Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* 24, 490–502. doi:10.1016/j.tim.2016.03.003
- Superti, F., Siciliano, R., Rega, B., Giansanti, F., Valenti, P., and Antonini, G. (2001). Involvement of bovine lactoferrin metal saturation, sialic acid and protein fragments in the inhibition of rotavirus infection. *Biochim. Biophys. Acta.* 1528, 107–115. doi:10.1016/s0304-4165(01)00178-7
- Thomas, S. (2020). The structure of the membrane protein of SARS-CoV-2 resembles the sugar transporter SemiSWEET. *Pathog Immun* 5, 342–363. doi:10.20411/pai.v5i1.377
- Tortorici, M. A., Walls, A. C., Lang, Y., Wang, C., Li, Z., Koerhuis, D., et al. (2019). Structural basis for human coronavirus attachment to sialic acid receptors. *Nat. Struct. Mol. Biol.* 26, 481–489. doi:10.1038/s41594-019-0233-y
- van der Strate, B., Beljaars, L., Molema, G., Harmsen, M., and Meijer, D. (2001). Antiviral activities of lactoferrin. *Antiviral Res* 52, 225–239. doi:10.1016/s0166-3542(01)00195-4
- Vandelli, A., Monti, M., Milanetti, E., Ponti, R. D., and Tartaglia, G. G. (2020). Structural analysis of sars-cov-2 and prediction of the human interactome. *Nucleic Acids Res.* 48 (20), 11270–11283. doi:10.1093/nar/gkaa864
- Varki, A. (2016). Biological roles of glycans. *Glycobiology* 27, 3–49. doi:10.1093/glycob/cww086
- Venkatraman, V., Yang, Y. D., Sael, L., and Kihara, D. (2009). Protein-protein docking using region-based 3d zernike descriptors. *BMC bioinformatics* 10, 407. doi:10.1186/1471-2105-10-407
- Waarts, B. L., Aneke, O. J., Smit, J. M., Kimata, K., Bittman, R., Meijer, D. K., et al. (2005). Antiviral activity of human lactoferrin: inhibition of alphavirus interaction with heparan sulfate. *Virology* 333, 284–292. doi:10.1016/j.virol.2005.01.010
- Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., and Zhou, Q. (2020). Structural basis for the recognition of the sars-cov-2 by full-length human ace2. *Science* 367, 1444–1448. doi:10.1126/science.abb2762
- Yi, M., Kaneko, S., Yu, D. Y., and Murakami, S. (1997). Hepatitis C virus envelope proteins bind lactoferrin. *J. Virol.* 71, 5997–6002. doi:10.1128/JVI.71.8.5997-6002.1997
- Zaki, A. M., Van Boheemen, S., Bestebroer, TM, Osterhaus, AD, and Fouchier, RA (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* 367, 1814–1820. doi:10.1056/NEJMoa1211721
- Zhang, Y., and Skolnick, J (2005). Tm-align: a protein structure alignment algorithm based on the tm-score. *Nucleic Acids Res* 33, 2302–2309. doi:10.1093/nar/gki524
- Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 1–4. doi:10.1038/s41586-020-2012-7
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 382, 727–783. doi:10.1056/NEJMoa2001017

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Miotto, Di Rienzo, Bò, Boffi, Ruocco and Milanetti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.