Comparative study of SARS-CoV-2 infection in different cell types: biophysical-computational approach to the role of potential receptors

Lenin González-Paz[a,b]*, María José Alvarado[a], María Laura Hurtado-León[a], Carla Lossada[c], Joan Vera-Villalobos[d], Marcos Loroño[e], J. L. Paz[f], Laura N. Jeffreys[g]; Ysaias J. Alvarado[c]*

[a] Universidad del Zulia (LUZ). Facultad Experimental de Ciencias (FEC). Departamento de Biología. Laboratorio de Genética y Biología Molecular (LGBM). 4001 Maracaibo, República Bolivariana de Venezuela.

[b] Instituto Venezolano de Investigaciones Científicas (IVIC). Centro de Estudios Botánicos y Agroforestales (CEBA). Laboratorio de Protección Vegetal (LPV). 4001 Maracaibo, República Bolivariana de Venezuela.

[c] Instituto Venezolano de Investigaciones Científicas (IVIC). Centro de Investigación y Tecnología de Materiales (CITeMA). Laboratorio de Caracterización Molecular y Biomolecular. 4001 Maracaibo, República Bolivariana de Venezuela.

[d] Facultad de Ciencias Naturales y Matemáticas, Departamento de Química y Ciencias Ambientales, Laboratorio de Análisis Químico Instrumental (LAQUINS), Escuela Superior Politécnica del Litoral, Guayaquil, Ecuador.

[e] Departamento Académico de Química Analítica e Instrumental, Facultad de Química e Ingeniería Química, Universidad Nacional Mayor de San Marcos. Lima, Perú.

[f] Departamento Académico de Química Inorgánica, Facultad de Química e Ingeniería Química, Universidad Nacional Mayor de San Marcos. Lima, Perú.

[g] Centre for Drugs and Diagnostics, Department of Tropical Disease Biology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK

Correspondence to: Lenin González-Paz, LUZ-FEC. LGBM. Maracaibo 4001 – Zulia, República Bolivariana de Venezuela. IVIC- CEBA- LPV. Maracaibo 4001 – Zulia, República Bolivariana de Venezuela. E-mail: lgonzalezpaz@gmail.com; Ysaias J. Alvarado, Instituto Venezolano de Investigaciones Científicas (IVIC). Centro de Investigación y Tecnología de Materiales (CITeMA). Laboratorio de Caracterización Molecular y Biomolecular. Maracaibo, República Bolivariana de Venezuela. E-mail: alvaradoysaias@gmail.com.

Abstract

Cellular susceptibility to SARS-CoV-2 infection in the respiratory tract has been associated with the ability of the virus to interact with potential receptors on the host membrane. We have modeled viral dynamics by simulating various cellular systems and artificial conditions, including macromolecular crowding, based on experimental and transcriptomic data to infer parameters associated with viral growth and predict cell susceptibility. We have accomplished this based on the type, number and level of expression of the angiotensin-converting enzyme 2 (ACE2), transmembrane serine 2 (TMPRSS2), basigin2 (CD147), FURIN protease, neuropilin 1 (NRP1) or other less studied candidate receptors such as heat shock protein A5 (HSPA5) and angiotensin II receptor type 2 (AGTR2). In parallel, we studied the effect of simulated artificial environments on the accessibility to said proposed receptors. In addition, viral kinetic behavior dependent on the degree of cellular susceptibility was predicted. The latter was observed to be more influenced by the type of proteins and expression level, than by the number of potential proteins associated with the SARS CoV-2 infection. We predict a greater theoretical pro- pensity to susceptibility in cell lines such as NTERA-2, SCLC-21H, HepG2 and Vero6, and a lower theoretical propensity in lines such as CaLu3, RT4, HEK293, A549 and U-251MG. An important relationship was observed between expression levels, protein diffusivity, and thermodynamically favorable interactions between host proteins and the viral spike, suggesting potential sites of early infection other than the lungs. This research is expected to stimulate future quantitative experiments and promote systematic investigation of the effect of crowding presented here.

Introduction

Cellular susceptibility to SARS-CoV-2 infection in the respiratory tract has been associated with the ability of the virus to interact with potential receptors on the host

membrane. We have modeled viral dynamics by simulating various cellular systems and artificial conditions, including macromolecular crowding, based on experimental and transcriptomic data to infer parameters associated with viral growth and predict cell susceptibility. We have accomplished this based on the type, number and level of expression of the angiotensin-converting enzyme 2 (ACE2), transmembrane serine 2 (TMPRSS2), basigin2 (CD147), FURIN protease, neuropilin 1 (NRP1) or other less studied candidate receptors such as heat shock protein A5 (HSPA5) and angiotensin II receptor type 2 (AGTR2). In parallel, we studied the effect of simulated artificial environments on the accessibility to said proposed receptors. In addition, viral kinetic behavior dependent on the degree of cellular susceptibility was predicted. The latter was observed to be more influenced by the type of proteins and expression level, than by the number of potential proteins associated with the SARS CoV-2 infection. We predict a greater theoretical pro- pensity to susceptibility in cell lines such as NTERA-2, SCLC-21H, HepG2 and Vero6, and a lower theoretical propensity in lines such as CaLu3, RT4, HEK293, A549 and U-251MG. An important relationship was observed between expression levels, protein diffusivity, and thermodynamically favorable interactions between host proteins and the viral spike, suggesting potential sites of early infection other than the lungs. This research is expected to stimulate future quantitative experiments and promote systematic investigation of the effect of crowding presented here.

Materials and methods

Search for cell lines and proteins in databases

Cellular susceptibility to SARS-CoV-2 infection in the respiratory tract has been associated with the ability of the virus to interact with potential receptors on the host membrane. We have modeled viral dynamics by simulating various cellular systems and artificial conditions, including macromolecular crowding, based on experimental and

transcriptomic data to infer parameters associated with viral growth and predict cell susceptibility. We have accomplished this based on the type, number and level of expression of the angiotensin-converting enzyme 2 (ACE2), transmembrane serine 2 (TMPRSS2), basigin2 (CD147), FURIN protease, neuropilin 1 (NRP1) or other less studied candidate receptors such as heat shock protein A5 (HSPA5) and angiotensin II receptor type 2 (AGTR2). In parallel, we studied the effect of simulated artificial environments on the accessibility to said proposed receptors. In addition, viral kinetic behavior dependent on the degree of cellular susceptibility was predicted. The latter was observed to be more influenced by the type of proteins and expression level, than by the number of potential proteins associated with the SARS CoV-2 infection. We predict a greater theoretical pro- pensity to susceptibility in cell lines such as NTERA-2, SCLC-21H, HepG2 and Vero6, and a lower theoretical propensity in lines such as CaLu3, RT4, HEK293, A549 and U-251MG. An important relationship was observed between expression levels, protein diffusivity, and thermodynamically favorable interactions between host proteins and the viral spike, suggesting potential sites of early infection other than the lungs. This research is expected to stimulate future quantitative experiments and promote systematic investigation of the effect of crowding presented here.

Design of theoretical expression systems

A set of 25 types of study systems (from 1 to 25) divided into 2 categories, theoretical systems of possible individual expression (from 1 to 13) further subdivided into 7 subtypes (from A-G) was established in a totally random manner; and theoretical systems of possible multiple expression (from 14 to 25) which represent expression systems that simulate the simultaneous participation of ≥ 2 proteins (up to a maximum of 7), and subdivided in turn into 13 types (from A-M). All this resulted in 247 possible expression systems that group the cell lines Vero6, A549, HepG2, CaCo2, HEK293, HeLa, U-251

MG, CaLu3, NTERA-2, RPTEC-TERT1, RT4, HBEC3-KT and SCLC -21H, as well as the possible receptors or facilitators NRP1, CD147, FURIN, TMPRSS2, ACE2, HSPA5 and AGTR2. To delimit the number of simulations of the theoretical expression systems, it was proposed to study individually all the receptors in each cell line, and also due to their relevance, a set of possible "fixed" receptor proteins was proposed assuming a simultaneous and constitutive expression, specifically , the ACE2 and TMPRSS2 pair was assumed as a main and fixed set of proteins expressed simultaneously, as has been suggested [2], so the ACE2-TMPRSS2 pair was combined individually and in conjunction with the second group of proteins most relevant reported (NRP1, CD147 and FURIN) [1,2,3]. The third group of less relevant proteins (HSPA5 and AGTR2) [2] was simulated only jointly and associated with the ACE2-TMPRSS2 pair. The expression of all the proteins was also simulated simultaneously in each cell line considered (a total of 7) as it is a probable event, although single-cell studies have revealed considerable variation from one cell to another, and patterns of very heterogeneous co-expression including cells that do not express all the genes individually, which justifies simulating individual and group expression scenarios (systems) where the susceptibility of cells is mediated by the type, number and level of expression of proteins that they behave as potential recipients [23].

In this sense, we assume that a simulated expression system represents the expression of a specific protein or set of proteins in one of the cell lines considered as the case may be. In the case of the individual expression systems, the cell lines were numbered, and the receptors were assigned letters, as indicated above, for example, the 1-A system, corresponded to the individual expression of the NRP1 (A) receptor in the Vero6 cell line (see Table 1). While, in the expression systems of more than one possible receptor, the nomenclature was reversed, and a number was assigned to the proteins followed by a

letter to identify the cell line, for example, the 14-A system, corresponded to the co-expression of putative possible receptors NRP1, CD147, FURIN, TMPRSS2, ACE2, HSPA5 and AGTR2 in the same Vero6 (A) cell line. The numbers and letters assigned to the designed systems are completely random, as is the order of the proteins and cell lines considered (see Table 1).

Table 1. Profile of the proposed theoretical expression systems.

System	Cell lines	Individual Expression*
1	Vero6	NRP1 (A), CD147 (B), FURIN (C), TMPRSS2 (D), ACE2 (E), HSPA5 (F), AGTR2 (G)
2	A549	
3	HepG2	
4	CaCo2	
5	HEK293	
6	HeLa	
7	U251 MG	
8	CaLu3	
9	NTERA-2	
10	RPTEC TERT1	
11	RT4	
12	HBEC3-KT	
13	SCLC-21H	

System	Multiple Expression*	Cell lines
14	NRP1+CD147+FURIN+TMPRSS2+ACE2+HSPA5+AGTR2	Vero6 (A), A549 (B), HepG2 (C), CaCo2 (D), HEK293 (E), HeLa (F), U251 MG (G), CaLu3 (H), NTERA-2 (I), RPTEC TERT1 (J), RT4 (K), HBEC3-KT (L), SCLC-21H (M)
15	TMPRSS2+ACE2	
16	NRP1+CD147+FURIN+TMPRSS2+ACE2	
17	NRP1+FURIN+TMPRSS2+ACE2	
18	CD147+FURIN+TMPRSS2+ACE2	
19	NRP1+CD147+TMPRSS2+ACE2	
20	FURIN+TMPRSS2+ACE2	
21	NRP1+TMPRSS2+ACE2	
22	CD147+TMPRSS2+ACE2	
23	CD147+TMPRSS2	
24	HSPA5+AGTR2	
25	HSPA5+AGTR2+TMPRSS2+ACE2	

*Angiotensin converting enzyme 2 (ACE2), transmembrane serine 2 (TMPRSS2), basigin2 (CD147), FURIN protease, neuropilin 1 (NRP1), heat shock proteinA5 (HSPA5) and angiotensin II receptor type 2 (AGTR2).

To compare the susceptibility propensity of cell lines to be infected by SARS-CoV-2, we assumed as a reference the expression level of each possible putative receptor using the consensus transcriptomic data used to classify all genes according to their expression specific tissue, single cell, brain region, blood cell or cell line specific reported in *The*

Human Protein Atlas, especially since it has been reported that the expression levels of cellular receptors can confer permissiveness for infections including those associated with SARS-CoV-2 [24,25,26]. In this study, the consensus normalized expression (NX) value was used for each gene, which represents the maximum value of NX in the transcriptomics database in *The Human Protein Atlas*. An NX value ≥1 is indicative of the expression of the protein in at least one type of tissue/region/cell, while an NX <1 represents the absence of expression, however, since the expression levels of some important receptors are very low in some cell lines, it has been suggested to assign minimum expression values as a reference [11]. Therefore, to reduce the error resulting from the limitations that both the databases and the expression level detection methods may present, we assume that an NX = 0.0 would be equivalent to an NX = 0.01 (see Figure 1, see Figure 2). For more details on the normalization of transcriptomic data, as well as its classification and the source of the data, it is recommended to read the "Assays and Annotations section" in *The Human Protein Atlas* [18].

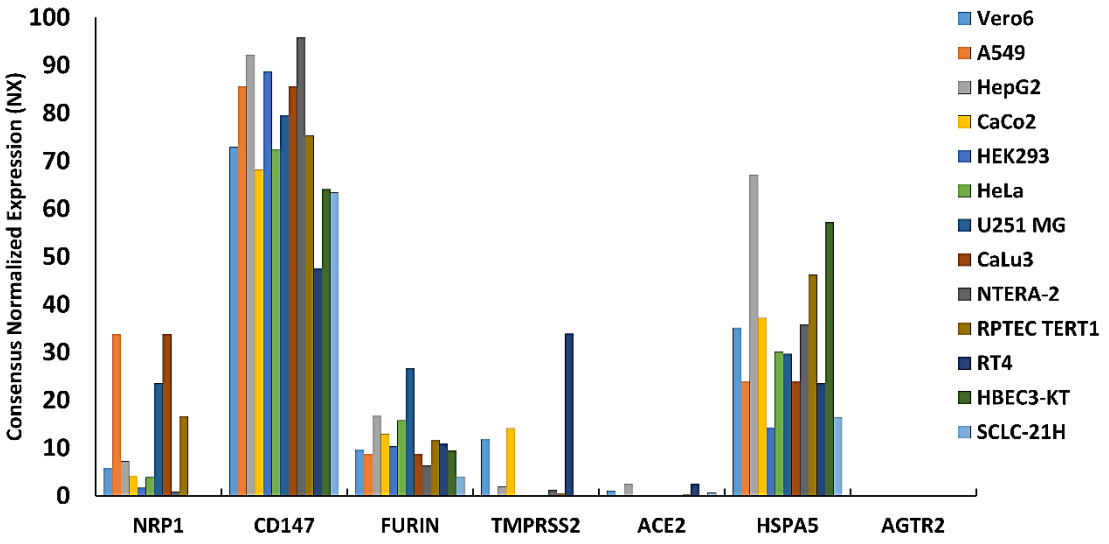

Figure 1. Cell lines and the level of RNA expression of the potential receptors considered in this study in terms of normalized NX values. The expression levels of angiotensin converting enzyme 2 (ACE2), transmembrane serine 2 (TMPRSS2), basigin2 (CD147),

FURIN protease, neuropilin 1 (NRP1), heat shock protein A5 (HSPA5) and angiotensin II receptor type 2 (AGTR2), are shown.

Cell lines	NRP1	CD147	FURIN	TMPRSS2	ACE2	HSPA5	AGTR2
HepG2	7.2	92.2	16.8	1.9	2.5	67.1	0.01
U251 MG	23.5	79.6	26.6	0.01	0.01	29.7	0.01
A549	33.8	85.6	8.8	0.01	0.01	24.0	0.01
CaLu3	33.8	85.6	8.8	0.01	0.01	24.0	0.01
RPTEC TERT1	16.6	75.3	11.7	0.5	0.4	46.2	0.01
NTERA-2	0.8	95.8	6.3	1.2	0.1	35.8	0.01
CaCo2	4.2	68.2	12.9	14.1	0.01	37.2	0.01
Vero6	5.8	72.9	9.7	11.9	1.0	35.2	0.01
HBEC3-KT	0.01	64.1	9.4	0.01	0.1	57.2	0.01
HeLa	3.9	72.4	15.8	0.01	0.01	30.1	0.01
RT4	0.1	47.6	11.0	33.9	2.5	23.5	0.01
HEK293	1.7	88.7	10.3	0.01	0.1	14.3	0.01
SCLC-21H	0.01	63.4	4.0	0.01	0.6	16.4	0.01

Figura 2. Summary of the RNA expression levels of the receptors studied in different cell lines analyzed in the Atlas of Human Proteins. The generated RNA sequencing results are reported as normalized NX values. Numerical values and a pie chart representation of the expression levels of angiotensin converting enzyme 2 (ACE2), transmembrane serine 2 (TMPRSS2), basigin2 (CD147), FURIN protease, neuropilin 1 (NRP1), heat shock protein A5 (HSPA5) and angiotensin II receptor type 2 (AGTR2).

Viral dynamics data

For illustrative purposes, we consider an *in vitro* data set of SARS-CoV-2 virus infection in cell lines recommended for culturing SARS-CoV-2. In order to observe the theoretical infectivity of the virus as a function of the level of susceptibility mediated by the type, number and level of expression of the proteins associated with virus receptors, we assume constant values of the viral titer with a multiplicity of infection (MOI) equal to 1 as suggested to ensure the same probability of infection [10,27,28,29,30,31,32,33], and especially to consider the recommended viral load to induce cytopathological effects after infection (*pi*) [28]. We also assume an initial cell concentration of ~$1x10^5$ cells/mL as suggested [1,10,11,31]. The time period for measuring the infection was set at 24 hours post infection (hpi) as has also been reported [10,30,32,33,34], especially since it has been described that at 24 hpi in plaque assays of infections with SARS-CoV-2 the infectious

virus titer may spike in several cell lines [30,35]. An effective infection rate (β) equal to 2 was assumed, considering the reported number of basic reproduction, R_0 [8].

To determine the hypothetical concentration of target or susceptible cells based on the type, number and level of expression of the proteins considered, the following was carried out. First, we assumed the same initial total number of cells/mL previously described for each cell line, and in view of the fact that the expression levels of some important receptors are very low in some cell lines, we assigned minimum values of percentage expression to calculate a hypothetical number of cells, and the percentage of receptors expressed, as suggested [11,36]. The percentage expressions of some of the receptors considered very close to the NX values have already been experimentally reported [1], in this sense, we assume that each NX value would represent the percentage value of susceptible cells in the case of the receptors considered individually, and the mean of the expression systems with 2 or more proteins would equally represent the percentage of susceptible cells in the modeled systems.

Mechanistic model for the simulation of viral dynamics

It was adjusted using a limited model of target cells, the mathematical model considered depends on the available data and the hypothesis to be addressed based on the susceptibility of cells as a function of the type, number and level of expression of receptors, for which we sought to estimate rates of infection of susceptible cells and death of infected cells. Therefore, a widely used model for viral infection called the limited target cell model was utilised [11,14]. The model includes three compartments: susceptible cells (U), infected cells (I) and viral titers (V). The applied model is represented in the following differential equations:

$$\frac{dU}{dt} = -\beta UV \qquad \text{(eq.1)}$$

$$\frac{dI}{dt} = \beta UV - \delta I \qquad \text{(eq.2)}$$

The term on the left of the equations represents the change of the variables with respect to time. Parameters β and δ represent the effective infection rates and dead infected cells, respectively. It is considered that the virus (V) infects susceptible cells (U) with a β rate, and that infected cells are eliminated with a δ rate. Based on this model, the basic reproduction number was calculated, R_0, representing the average number of cells that can be infected from a single infected cell at the beginning of infection:

$$Ro = \frac{p\beta U}{\delta(c + \beta U)} \qquad \text{(eq.3)}$$

Where virions are released from infected cells productively at a rate (p) per day, and are eliminated from the circulation at a rate c or are lost when infecting a target cell, for which it was assumed that the rate of elimination of the virus, c, was equal to 10 days^{-1} as recommended [11]. For SARS-CoV-2 it has been suggested that the rate p is equal to ≈22.7 copies/day per cell, and that since the rate p depends on U, it must be estimated by the product $p \times U$ [11].

Protein-protein docking and molecular dynamics (MD) of the SARS-CoV-2 spike protein and the proteins of interest

As it has been shown that the spike protein is responsible for the ability of SARS-CoV-2 to interact with receptors expressed on the membrane of host cells leading to virus entry [1]. We performed a comparative analysis using two popular molecular docking models for a rigorous prediction of the standard free energy (ΔG) of protein-protein complexes made up of the spike protein and each of the proteins that behave as possible facilitators of the interaction between the virus and the target cells. The protein structures

used were neuropilin 1 (NRP1) (PDB: 7JJC), CD147 (PDB: 3I84), FURIN protease (PDB: 6HZB), serine transmembrane protease 2 (TMPRSS2) (PDB: 7MEQ) [1,2,3], as well as angiotensin converting enzyme 2 (ACE2) (PDB: 6VW1) which has been described as the main cellular receptor for interaction with the spike protein [4]. Additionally, the less popular HSPA5 (PDB: 6ZYH) and AGTR2 (PDB: 6JOD) were also investigated [2]. The spike protein has been crystallised in a number of forms including in theclosed conformation (PDB: 6VXX), the open conformaiton (PDB: 6VYB) and in fragments including the S1 region (PDB: 7DEO) and the S2 region (PDB: 7COT). It was sought determine if there is the theoretical possibility of a thermodynamically favorable interaction between any region of the viral spike protein with any of the potential receptors considered in this study and so all the spike protein structures listed were used in these experiments. All structures were obtained in PDB format from the RCSB protein database (https://www.rcsb.org/), and the quality of the crystal structures was validated using MolProbity (http://molprobity.biochem.duke.edu/) [37].

Complexes were constructed using HDOCK (http://hdock.phys.hust.edu.cn/) an algorithm for hybrid protein-protein docking with template-based modeling [38], and HawkDock (http://cadd.zju.edu.cn/hawkdock/) a web server for the prediction and structural analysis of the protein-protein complex that combines the ATTRACT algorithm for global macromolecular docking and the HawkRank algorithm for scoring, in addition to the MM/GBSA method that is used to predict free bond energy [39]. As usual, for all the docking methods all the water molecules were removed and the PDB files were separated into two different files, one containing the spike protein and the other containing the structure of the potential receptor. In the sampling of the probabilistically most feasible and thermodynamically most favorable positions in the complexes, only the three runs with the most favorable berth were considered. This criterion was used to

discriminate the complexes that would be subjected to further analysis, including molecular dynamics [40]. Additionally, to validate the results of the docking, the DockScore algorithm (http://caps.ncbs.res.in/dockscore/) was used to score the protein-protein complexes. This method identifies the optimal interactions between the two associated proteins using several putative interface characteristics such as area, short contacts, conservation, spatial clustering, and the presence of hydrophobic and positively charged residues [41].

To determine the subset of residues at the interface that account for most of the free binding energy in the protein-protein complex, residues called "binding hot spots" were predicted. For which the KFC Server (Knowledge-based FADE and Contacts) (https://mitchell-web.ornl.gov/KFC_Server/index.php) was used, which provides a web tool to predict protein binding hot spots based on machine learning approaches. For each residue within the link interface, the KFC2 server characterizes its local framework and compares it to known experimentally determined hot spot environments. Specifically, the server analyzes various chemical and physical characteristics surrounding an interface residue and predicts the residue's classification using a model trained on previous experimental data [42].

Molecular dynamics (MD) simulations for favorable docking were performed with two purposes: 1) to study the relative stability of the bond; and 2) obtain the minimum energy conformation of the complexes. For a protein-protein complex, the MD system was first relaxed through a series of minimization procedures that include three phases: relaxation, equilibrium, and sampling, as recommended [40]. The MD simulation of the crystalline structures was carried out in an explicit water system. Specifically, the solvation of the system was carried out in an 8.0 Å solvation box. Our MD system also consisted of a copy of the coupled chain region of each protein system. An Amber99SB-

ILDN force field was applied to the complex, with TIP3P water model. The whole system was neutralized. Water molecules were treated as rigid bodies in all models, allowing a simulation time interval of 2 fs. Periodic boundary conditions were applied and the Berendsen algorithm was adopted for the docking of temperature and pressure. After a first steeper descent to 5000 steps and a conjugate gradient to 5000 steps energy minimizations with positional constraints on the solute, an initial simulation of 100 ps was performed with the positions of the solute atoms constrained by a constant of force of 10 kcal/(mol*Å^2) to allow water to diffuse around the molecule and for equilibrium. The PME method was used to calculate the electrostatic contribution to unbound interactions with a limit of 14.0 Å and a time interval of 1 fs. The cutoff distance of the van der Waals interaction was 14.0 Å. After this equilibration run, the NVT production run (number of particles, volume, temperature) at 300 K was performed with the cell size remaining the same. The SHAKE algorithm was applied to the system and the time interval was set at 2 fs. Ten structures were obtained every 10 ns as target structures extracted from a total path of 100 ns. For calculations of mean squared deviations (RMSD), the equation,

$$RMSD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\delta_i^2} \qquad \text{(eq.4)}$$

where δi is the distance between atom i and a reference structure or the average position of the n equivalent atoms. All MD simulations and additional adjustments were performed using COSGENE/myPresto [40,43].

Additionally, to predict conformational changes in each protein-protein complexes after its interaction, the NMSim software (https://cpclab.uni-duesseldorf.de/nmsim/main.php) was used, which is a computational technique that uses a three-step including coarse-graining (CG), normal mode analysis (NMA), and elastic

network model (ENM) to provide realistic conformations in reasonable simulation time. For this, the minimum energy structure obtained with mypresto at 100 ns was used to also calculate the root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF), using the radius-guided movements (ROG) approach of the NMSim Server. The ROG-guided NMSim simulation is a method to search for a conformation of a given protein-protein complex, and allows to describe the compactness of a protein. In an ROG-guided NMSim simulation, the trajectory is tailored towards the bound structure by selecting the pathway that leads to a decrease in Radius of gyration (Rg), and conformations are generated by structure distortion along directions of random linear combinations of low-frequency normal modes [40].

Prediction of the theoretical diffusivity of the proteins of interest associated with cellular susceptibility in SARS-CoV-2 infection

It is known that the mobility of biological molecules depends not only on the concentration of the crowding macromolecular solutions, which is generally related to the viscosity η of the crowding solution, but also on the relative size between the crowded and the biological molecule of interest. The diffusion coefficients of the molecules of interest were measured as a function of the concentrations of binders (crowding solution), for which the viscosity of model cell lines used in this study was used as A549 (adenocarcinomic human alveolar basal epithelial cells with a viscosity ~$1.4x10^{3}$ Pa/s) and the HeLa negative control model used (with a cytoplasmic viscosity of ~$4.4x10^{-2}$ Pa/s) [44]. Normal Swiss 3T3 cells (with a viscosity of ~$2.4x10^{-2}$ Pa/s) were also considered in order to model the mobility of biological molecules considering solutions of high and normal macromolecular crowding [44,45]. Additionally, the cytoplasmic viscosity reported for other cell lines such as ASTC-a-1 (human lung adenocarcinoma

with a viscosity of ~1.6x10^{-3} Pa/s) [46] and H1299 (lung carcinoma with a viscosity of ~1.1x10^{2} Pa/s) [47].

The translational diffusion coefficient (D_t) was determined according to the Stokes-Einstein equation (equation 5) as recommended [44,45,48,49,50]:

$$D^t = \frac{k_{B^T}}{6\pi\eta R} \qquad (\text{eq. } 5)$$

In this model, the only source of dissipative effects (internal friction) is the η of the solvent (cytoplasm) [45]. Where k_B is Boltzmann's constant, T is absolute temperature, η is viscosity and R is the radius of the particle. The results were compared with the relative viscosity of the water calculated considering the friction ratio (f/f_o) [51] of the minimum energy structures and the molecular weight, both parameters calculated with the HullRad server (https://hullrad.wordpress.com/) [52] in each case, and using the module *Protein Research* for diffusion calculation from *Cytiva Life Sciences* (https://www.cytivalifesciences.com/en/us/solutions/protein-research/products-and-technologies/diffusion-coefficient-calculator) [53].

On the other hand, when taking the weighted average over the distances in the equation described by Phillips et al., 2009 [54], which describes how a group of molecules located at a point extend over time (t), and taking into account several dimensions (using the Pythagorean theorem), we obtain according to Schavemaker et al., 2018 [55]:

$$t = \frac{d^2}{6nD} \qquad (\text{eq. } 6)$$

Here d is the distance, n is the number of dimensions considered and D, the diffusion coefficient with translational contribution.

Likewise, a preliminary study was carried out to determine the potential theoretical flux of the proteins considered in this study through the cell membrane using according to Fick's law:

$$J = -D\left(\frac{\Delta C}{\Delta X}\right) \qquad \text{(eq. 7)}$$

where J (mol m^{-2} s^{-1}) is the flow, D (m^2 s^{-1}) the diffusion coefficient, and $\Delta C/\Delta X$ (mol m^{-4}) is the concentration gradient. According to Fick's law [56,57,58,59,60], the flux (J) is directly proportional to the difference in concentration (ΔC) to the diffusion coefficient (D) and inversely proportional to the thickness of the membrane (ΔX). For illustrative purposes only, a ΔC with a diffusion of one third of the hypothetical initial concentration of ≈1 mol m^{-4} was assumed for transmembrane flux, and dimensions reported for the plasma membrane (~ 5 nm) were considered for ΔX [61].

Results and Discussion

Theoretical propensity for cellular susceptibility depending on the type, number and level of expression of proteins associated with SARS-CoV-2 infection

No significant correlation was found between the consensus normalized expression (NX) levels between any of the proteins associated with possible receptors to mediate SARS-CoV-2 infection regardless of the type of cell line, except for a moderate positive correlation between the ACE2 receptor and TMPRSS2 protease (r = 0.60) as expected, and interestingly a moderate negative correlation between CD147 protein and TMPRSS2 (r = 0.70) (see Figure 3).

A **B**

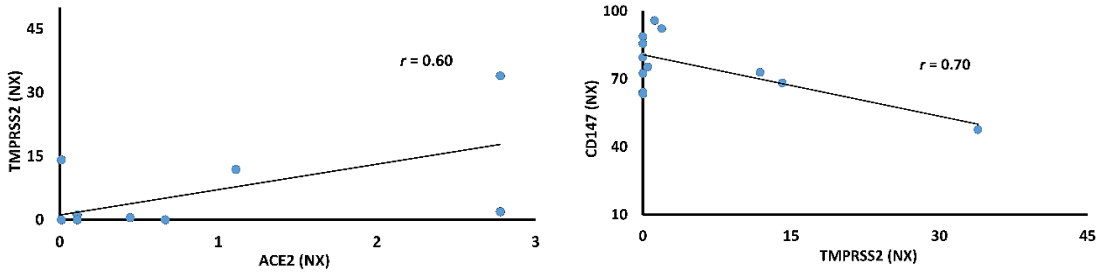

Figura 3. Proteins associated with SARS-CoV-2 infection with correlation between their levels of consensus normalized expression (NX). A) moderate positive correlation between the receptor ACE2 and TMPRSS2 protease (r = 0.60), and B) moderate negative correlation between CD147 protein and TMPRSS2 (r = 0.70).

On the other hand, under the conditions of this study, it was predicted that 85.42% (211/247) of the theorized expression systems could exhibit a susceptibility (U) of up to 95.8% cells/mL in the case of systems proposed for the expression of proteins individually, with a range between $8x10^2$ - $1x10^5$ cells/mL potentially susceptible. While in multiple expression systems for 2 or more proteins, a susceptibility of up to 48.5% cells/mL was predicted, with a range of $6x10^2$ - $5x10^4$ cells/mL. In contrast, 14.57% (36/247) of the simulated expression systems were predicted to exhibit a theoretical susceptibility $\leq 2x10^2$ cells/mL. This susceptibility was represented by 11.33% (28/247) of the theoretical individual expression systems, and 3.23% (8/247) in the modeled systems for the expression of 2 or more proteins. These results show that when considering the NX expression values for each protein as a factor of cellular susceptibility, the individual expression systems would present 49.4% more susceptibility compared to the expression systems of ≥ 2 proteins (Table 1-2, Figure 1-2, Figure 4).

In the theoretical individual expression systems, CD147 was the protein with the highest expression in all the cell lines considered (13/13), promoting theoretical susceptibility by an average of 76.26% cells/mL, being the protein with the highest expression among all the simulated individual systems; followed by the HSPA5 protein

with a high expression in 76.92% (10/13) of cell lines, mediating a theoretical susceptibility of 36.92% cells/mL. The proteins with the lowest expression in the individual systems were AGTR2, ACE2 and TMPRSS2 with a frequency of 100% (13/13), 61.54% (8/13) and 53.85% (7/13) in the cell lines studied, respectively, mediating a low theoretical susceptibility of between 0.01% - 0.04% cells/mL. The cell line in which the highest mean individual protein expression was predicted and therefore the highest theoretical susceptibility was HepG2, with a mean NX = 26.82, mediating a mean theoretical susceptibility of ~$3x10^4$ cells/mL, followed by U-251 MG with a mean NX = 22.8, with a mean theoretical susceptibility of ~$2x10^4$ cells/mL. The cell line in which the lowest expression level was observed was SCLC-21H, with a mean NX = 12.1 and a theoretical susceptibility of ~$1x10^4$ cells/mL (see Tables 1-2, see Figures 1-2, see Figure 4).

Specifically, in terms of the simulated models of individual expression, it was predicted that the 9-B system, represented by the individual expression of the CD147 protein in the NTERA-2 line, was the system with the highest theoretical susceptibility, mediating a susceptibility of up to $1x10^5$ cells/mL, followed by the 3-B system, represented by the CD147 protein in the HepG2 line, with a theoretical susceptibility of up to $9x10^4$ cells/mL. On the contrary, the individual systems 1, 3, 10 and 11 (all of type G); 4 and 9 (types E and G), 2, 5-8 and 12 (D, E and G); and 13 (D and G) presented the lowest theoretical susceptibility with a range between 10 - 100 cells/mL (see Table 1-2, see Figure 1-2, see Figure 4).

All systems designated with the letters D, E and G, were represented by individual expression systems of the TMPRSS2, ACE2 and AGTR2 proteins, respectively. It is important to note that in all theoretical individual expression systems at least two cell lines predicted a theoretical mean susceptibility ≥ $1x10^4$ cells/mL. Regarding the HeLa

cell line used as a susceptibility control, it was predicted as one of the lines with the lowest susceptibility propensity, generally located and with respect to the rest of 12 cell lines below the first 5 cell lines with the best propensity to susceptibility as the case may be, being located in some cases between positions 7-10 (Tables 1-2, Figures 1-2, Figure 4).

In the specific case of the theoretical models of multiple expression, it was observed that one of the systems with the highest theoretical expression of proteins simultaneously (7 proteins) was represented by the 14-C system, in which the highest propensity was simulated to cellular susceptibility with ~$3x10^4$ cells/mL, this system was represented by the HepG2 cell line, hypothetically expressing all the proteins simultaneously (NRP1, CD147, FURIN, TMPRSS2, ACE2, HSPA5 and AGTR2). However, at the global level, the simultaneous expression system that the highest cellular susceptibility according to the protein arrangement was type 23, represented by the theoretical expression of only 2 proteins (CD147 and TMPRSS2) with a global mean of $4x10^4$ cells/mL. and a susceptibility range between $3x10^4$ - $5x10^4$ cells/mL depending on the cell type. Specifically, the 23-I system (NTERA-2 cell line) was the one that exhibited the highest propensity to cell susceptibility, and the 23-M system (SCLC-21H cell line) was the one with the lowest susceptibility within the group (Tables 1-2, Figures 1-2, Figure 4).

While the systems in which the lowest propensity for cell susceptibility was predicted were type 15 (joint expression of TMPRSS2 and ACE2) with a global mean of $3x10^3$ cells/mL, and a susceptibility range of between 10 - $2x10^4$ cells/mL depending on the type of cell line. Specifically, in system 15, the 15-K subtype (cell line RT4) exhibited the highest propensity to cell susceptibility, unlike systems 15-B (cell line A549), 15-F (cell line HeLa), 15-G (U-251 MG cell line) and 15-H (CaLu3 cell line) in which the lowest cell susceptibility was predicted (~10 cells/mL) (Tables 1-2, Figures 1-2, Figure 4).

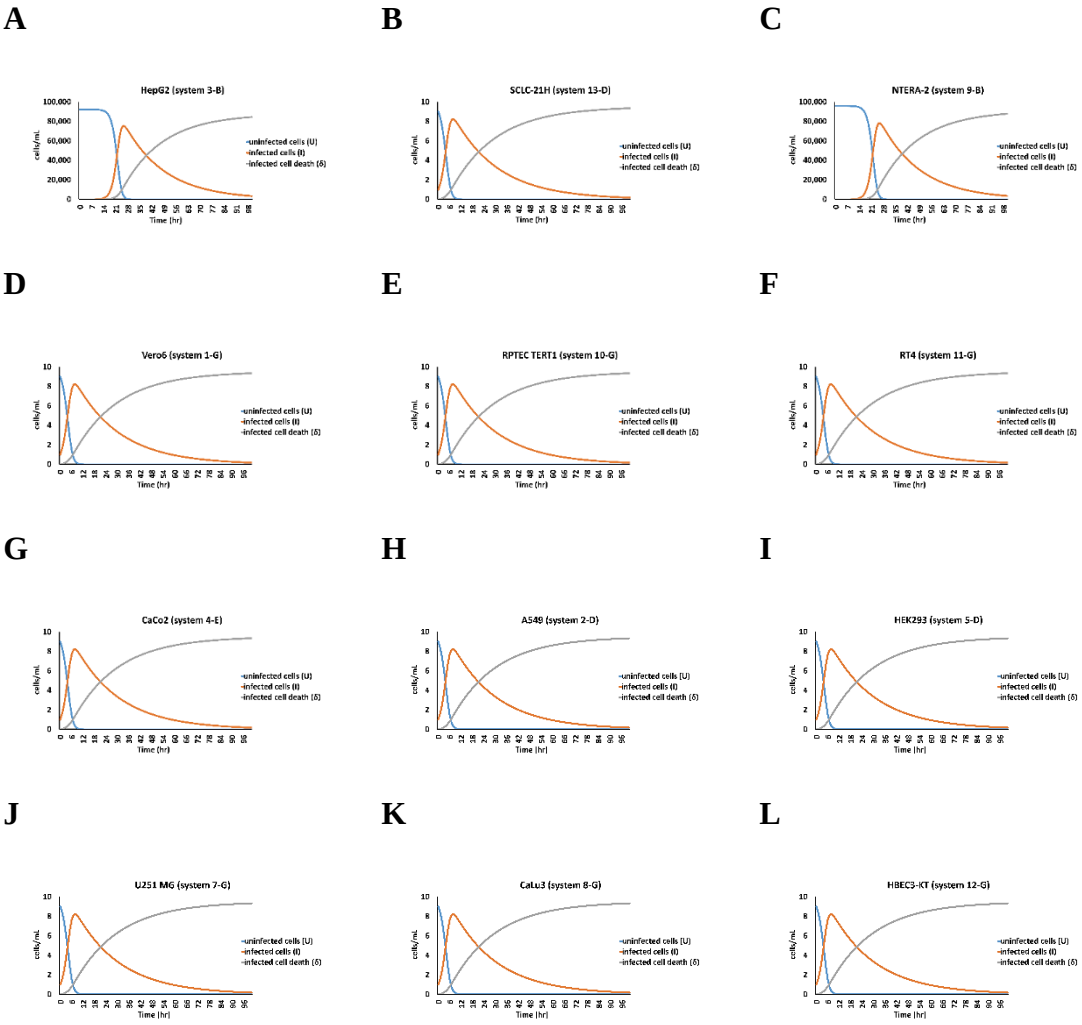

Figure 4. Mechanistic model of the estimates of the rates of uninfected cells (*U*), infected cells (*I*) and death of infected cells (δ). For illustrative purposes, only a representative graph of the individual expression systems cited in the text is shown for each cell line studied using the limited target cell model. Additional graphics are shown in supplementary material. A) Individual expression of CD147 in the HepG2 cell line (3-B system), B) individual expression of FURIN in the SCLC-21H cell line (13-D system), C) individual expression of CD147 in the NTERA-2 cell line (9-B system), D) individual expression of AGTR2 in the Vero6 cell line (1-G system), E) Individual expression of AGTR2 in the RPTEC TERT1 cell line (10-G system), F) individual expression of AGTR2 in cell line RT4 (11-G system), G) Individual expression of ACE2 in the CaCo2 cell line (4-E system), H) individual expression of TMPRSS2 in the A549 cell line (2-D system), I) individual expression of TMPRSS2 in the HEK293 cell line (5-D system), J) individual expression of AGTR2 in the U251 MG cell line (7-G system), K) individual expression of AGTR2 in the CaLu3 cell line (8-G system), L) individual expression of AGTR2 in the HBEC3-KT cell line (12-G system). The rest of the graphs are shown in supplementary material.

The susceptibility profiles predicted in a strict sense, for all cell lines of the systems type 23 (CD147 + TMPRSS2), 19 (NRP1 + CD147 + TMPRSS2 + ACE2), 18 (CD147 + FURIN + TMPRSS2 + ACE2), 16 (NRP1 + CD147 + FURIN + TMPRSS2 + ACE2) and 14 (NRP1 + CD147 + FURIN + TMPRSS2 + ACE2 + HSPA5 + AGTR2) presented a theoretical mean susceptibility $\geq 1x10^4$ cells/mL in 100% of the cell lines considered, respectively; while in system 24 (HSPA5 + AGTR2) it was 85%, in system 22 (CD147 + TMPRSS2 + ACE2) it was 77%, in system 17 (NRP1 + FURIN + TMPRSS2 + ACE2) it was 31%, and in systems 21 (NRP1 + TMPRSS2 + ACE2) was 23%. These results show that ≈ 66% (6/9) of the most prevalent expression profiles had the presence of CD147 + TMPRSS2 proteins (Tables 1-2, Figures 1-2).

Regarding the amount of virions that could theoretically be released from susceptible, productively infected cells, the type 23 system was predicted to have a rate (*p*) with a mean of 9.2 $x10^5$ copies/day per cell, with minimum values of 7.2 $x10^5$ and maximums of 1.1 $x10^6$ copies/day per cell, represented by the cell lines NTERA-2 (system 23-I) and SCLC-21H (system 23-M), respectively, which corresponds to the previous susceptibility values described. Globally, it was predicted that systems 14, 16, 18, 19, 22-24 could theoretically release the largest number of virions from infected susceptible cells (classified from AM) at a rate (*p*) of between 1.6 $x10^5$ - 7.6 $x10^5$ copies/day per cell. While in system 21 the lowest *p* rate was predicted with a mean of 1.2 $x10^5$ copies/day per cell, and minimum and maximum values of 9.1 $x10^2$ and 2.8 $x10^5$ copies/day per cell corresponding to the RT4 cell lines (system 21- L) and HBEC3-KT (21-K system), respectively (Tablew 1-2, Figurew 1-2).

The theoretical number of dead infected cells (δ) was also predicted and it was calculated that systems 18 and 19 could generate the highest number of dead cells with a mean of 2.5 $x10^4$ cells/mL with minimum and maximum values of 1.9 $x10^4$ and 3.0 $x10^4$

cells/mL, in SCLC-21H cell lines (18-M system) and (19-M system) respectively. In general terms, systems 14, 16-19 and 22-24 (sub-classified from A-M) could generate the highest number of dead cells, specifically between 1.4 $x10^4$ - 3.9 $x10^4$ cells/mL. On the contrary, with system 21 the lowest number of dead infected cells was predicted with a mean of 3.0 $x10^3$ cells/mL, and minimum and maximum values of 2.7 $x10^1$ and 6.9 $x10^3$ cells/mL corresponding in the same way to the RT4 cell lines (21-L system) and HBEC3-KT (21-K system), respectively (Tables 1-2, Figures 1-2).

To validate the theoretical susceptibility of the cell lines considered under the conditions of this study, the basic reproduction number, R_0, which represents the average number of susceptible cells infected from a single cell infected at the beginning of the infection, was also calculated predicting in systems 18, 22 and 23 an $R_0 \geq 21$. Specifically, in system 18 all cell lines, except in systems 18-L (HBEC3-KT) and 18-M (SCLC-21H), an $R_0 \geq 21$ was predicted, as in 22, with the exception of system 22-M (SCLC-21H), while in system 23 the 23-C (HepG2) and 23-I (NTERA-2) models presented the highest basic reproduction value with an $R_0 \geq 22$. Unlike systems 2, 6, 7, 8 and 15 in which the lowest basic reproduction was predicted with an $R_0 \leq 16$, specifically the 2-D, 2-E and 2-G systems, as in the case of systems 6, 7 and 8 D, E and G, which are represented by the cell lines CaCo2, HEK293 and U251 MG, respectively (see Table 1-2, see Figure 1-2).

It is important to note that the type 2, 6, 7 and 8 systems subdivided into D, E and G correspond to the simulation of individual expression of the receptors/mediators ACE2, TMPRSS2 and AGTR2, respectively, as do the multiple expression systems of type 15-B, 15-F, 15-G and 15-H, constituted by the cell lines A549, HeLa, U251 MG and CaLu3 under each hypothetical condition of joint expression of TMPRSS2 and ACE2. The difference between the predicted R_0 values was statistically significant ($p < 0.01$) under the conditions of this study and, it seems to show a differential viral kinetic behavior

dependent on the degree of cellular susceptibility, which is influenced to a greater extent by the type and expression level of potentially receptor/mediator proteins associated with SARS-CoV-2 infection, and to a lesser extent by the number of potential receptor/mediator proteins expressed (Tables 1-2, Figures 1-2). All estimates of U, δ, p and R_0 are shown in detail in the supplementary material (Table S1).

Table 2. Estimates of the number of susceptible cells (U), dead infected cells (δ), productive release rate of virions from cells (p) and basic reproduction number (R_0).

System(s)	U (cell/mL) (mean [min-max])	p (copies/cell*day) (mean [min-max])	δ (cells/mL) (mean [min-max])	R_0 (mean ± SD)*
1 (A-G)	1.9×10^4 [1.0×10^1 - 7.3×10^4]	4.4×10^5 [2.3×10^2 - 1.7×10^6]	2.1×10^4 [1.4×10^1 - 7.6×10^4]	19.1 ± 3.8
2 (A-G)	2.2×10^4 [1.0×10^1 - 8.6×10^4]	4.9×10^5 [2.3×10^2 - 1.9×10^6]	2.3×10^4 [1.4×10^1 - 8.9×10^4]	16.7 ± 5.5
3 (A-G)	2.7×10^4 [1.0×10^1 - 9.2×10^4]	6.1×10^5 [2.3×10^2 - 2.1×10^6]	2.8×10^4 [1.4×10^1 - 9.5×10^4]	19.2 ± 3.8
4 (A-G)	2.0×10^4 [1.0×10^1 - 6.8×10^4]	4.4×10^5 [2.3×10^2 - 1.5×10^6]	2.1×10^4 [1.4×10^1 - 7.1×10^4]	18.0 ± 4.9
5 (A-G)	1.7×10^4 [1.0×10^1 - 8.9×10^4]	3.8×10^5 [2.3×10^2 - 2.0×10^6]	1.7×10^4 [1.4×10^1 - 9.2×10^4]	17.6 ± 4.7
6 (A-G)	1.7×10^4 [1.0×10^1 - 7.2×10^4]	4.0×10^5 [2.3×10^2 - 1.6×10^6]	1.8×10^4 [1.4×10^1 - 7.5×10^4]	16.6 ± 5.4
7 (A-G)	2.3×10^4 [1.0×10^1 - 8.0×10^4]	5.2×10^5 [2.3×10^2 - 1.8×10^6]	2.4×10^4 [1.4×10^1 - 8.3×10^4]	16.8 ± 5.6
8 (A-G)	2.2×10^4 [1.0×10^1 - 8.6×10^4]	4.9×10^5 [2.3×10^2 - 1.9×10^6]	2.3×10^4 [1.4×10^1 - 8.9×10^4]	16.7 ± 5.5
9 (A-G)	2.0×10^4 [1.0×10^1 - 9.6×10^4]	4.6×10^5 [2.3×10^2 - 2.2×10^6]	2.1×10^4 [1.4×10^1 - 9.9×10^4]	18.7 ± 3.7
10 (A-G)	2.2×10^4 [1.0×10^1 - 7.5×10^4]	4.9×10^5 [2.3×10^2 - 1.7×10^6]	2.3×10^4 [1.4×10^1 - 7.8×10^4]	18.9 ± 3.8
11 (A-G)	1.7×10^4 [1.0×10^1 - 4.8×10^4]	3.9×10^5 [2.3×10^2 - 1.1×10^6]	1.8×10^4 [1.4×10^1 - 5.0×10^4]	19.1 ± 3.8
12 (A-G)	1.9×10^4 [1.0×10^1 - 6.4×10^4]	4.3×10^5 [2.3×10^2 - 1.5×10^6]	2.0×10^4 [1.4×10^1 - 6.7×10^4]	18.6 ± 5.7
13 (A-G)	1.2×10^4 [1.0×10^1 - 6.3×10^4]	2.7×10^5 [2.3×10^2 - 1.4×10^6]	1.3×10^4 [1.4×10^1 - 6.6×10^4]	18.3 ± 5.6
14 (A-M)	2.0×10^4 [1.2×10^4 - 2.7×10^4]	4.5×10^5 [2.7×10^5 - 6.1×10^5]	2.1×10^4 [1.3×10^4 - 2.9×10^4]	20.9 ± 0.1
15 (A-M)	2.9×10^3 [1.0×10^1 - 1.8×10^4]	6.5×10^4 [2.3×10^2 - 4.1×10^5]	3.2×10^3 [1.4×10^1 - 2.0×10^4]	16.7 ± 4.1
16 (A-M)	2.1×10^4 [1.4×10^4 - 2.6×10^4]	4.7×10^5 [3.1×10^5 - 5.9×10^5]	2.3×10^4 [1.5×10^4 - 2.8×10^4]	20.9 ± 0.1
17 (A-M)	6.8×10^3 [1.2×10^3 - 1.3×10^4]	1.5×10^5 [2.4×10^4 - 2.8×10^5]	7.6×10^3 [1.4×10^3 - 1.4×10^4]	20.0 ± 0.5
18 (A-M)	2.3×10^4 [1.7×10^4 - 2.8×10^4]	5.3×10^5 [3.9×10^5 - 6.4×10^5]	2.5×10^4 [1.9×10^4 - 3.0×10^4]	21.0 ± 0.1
19 (A-M)	2.3×10^4 [1.6×10^4 - 3.0×10^4]	5.2×10^5 [3.6×10^5 - 6.8×10^5]	2.5×10^4 [1.9×10^4 - 3.0×10^4]	20.6 ± 2.8
20 (A-M)	5.7×10^3 [1.5×10^3 - 1.6×10^4]	1.3×10^5 [3.5×10^4 - 3.6×10^5]	3.3×10^3 [9.2×10^2 - 8.8×10^3]	19.0 ± 0.8
21 (A-M)	5.2×10^3 [4.0×10^1 - 1.2×10^4]	1.2×10^5 [9.1×10^2 - 2.8×10^5]	3.0×10^3 [2.7×10^1 - 6.9×10^3]	17.9 ± 2.9
22 (A-M)	2.7×10^4 [2.1×10^4 - 3.2×10^4]	6.2×10^5 [4.8×10^5 - 7.3×10^5]	1.5×10^4 [1.2×10^4 - 1.8×10^4]	21.2 ± 0.2
23 (A-M)	4.1×10^4 [3.2×10^4 - 4.9×10^4]	9.2×10^5 [7.2×10^5 - 1.1×10^6]	2.2×10^4 [1.7×10^4 - 2.6×10^4]	21.4 ± 1.5
24 (A-M)	1.7×10^4 [7.2×10^3 - 3.4×10^4]	3.8×10^5 [1.6×10^5 - 7.6×10^5]	9.4×10^3 [4.1×10^3 - 1.8×10^4]	20.5 ± 0.6
25 (A-M)	9.8×10^3 [3.6×10^3 - 1.8×10^4]	2.2×10^4 [8.2×10^4 - 4.1×10^5]	5.6×10^4 [2.1×10^3 - 1.0×10^4]	19.8 ± 0.6

The mean, minimum and maximum values of the simulated expression systems are shown, *, The mean and standard deviation values are displayed.

The ACE2 receptor is, so far, the best-known host factor for the entry of SARS-CoV-2, but various studies have identified other essential elements. Therefore, the correlation found between the level of expression of the ACE2 protein and TMPRSS2 was expected because a positive correlation between ACE2 and TMPRSS2 has already been reported in most organs where they have been detected [62]. Indeed, other coronaviruses also utilize ACE2 and TMPRSS for viral entry, for example SARS-CoV-1 [62]. TMPRSS2 facilitates viral entry by cleaving the SARS-CoV-2 spike protein between the S1 and S2 subunits causing a dramatic structural change leading to endocytosis [63]. However, it is important to note that, although a moderate positive correlation was found between the expression of ACE2 and TMPRSS2 under the conditions of this study, an absence of correlation has been reported between the expressions of the ACE2, TMPRSS2 and the FURIN protein (also considered in this study) in healthy volunteers with various clinical characteristics [63].

Although we found a negative correlation between TMPRSS2 and CD147, it has been reported that some cell types are more likely to be infected with SARS-CoV-2 through the CD147 receptor and the TMPRSS2 protease than through ACE2 [64], because the expression of ACE2 is one of the lowest while CD147 has a higher level of expression in various cell lines and tissues as observed in this study, while TMPRSS2 has a constitutive expression in many tissues and cells [64]. In fact, it was found that receptors such as ACE2, TMPRSS2 and AGTR2 have the lowest levels of expression according to the consensus transcriptomics data of *The Human Protein Atlas*, which corresponds to what has been reported in other studies [64, 65].

Our observations show that the higher probability of an infection could be mediated by the individual expression of certain candidate receptors than by the simultaneous expression of several of them. This corresponds to what has been reported by some

authors who point out that the expression of several proteins simultaneously is an event with considerable variation from one cell to another, and heterogeneous co-expression patterns have been observed including cells that do not express all genes even individually [23]. Similarly, to what has been described above [64] some studies have proposed as unlikely that differences in gene expression in cells associated, for example, with the respiratory and epithelial tracts of ACE2, TMPRSS2 and FURIN, confer an increased risk of COVID-19 [63].

In fact, it has been reported that the expression of ACE2 in some tissues can be approximately 1/3 of the level of expression of TMPRSS2 in the same tissues, which indicates a higher level of expression of TMPRSS2 than that observed for ACE2 [66,67]. Furthermore, it has been reported that ACE2 and TMPRSS2 may be upregulated in cells associated with the airways, while ACE2 may become downregulated in nasal epithelial cells simultaneously with TMPRSS2 may be upregulated in bronchial cells. Additionally, it has been described that some respiratory epithelia can be negative for the simultaneous expression of the ACE2 and TMPRSS2 protein, while at the same time they can be positive for CD147 and FURIN [68].

Although more attention has been paid to the ACE2 receptor and TMPRSS2 as well as to cells associated with the respiratory tract which showed important basic reproduction numbers in this study. Our pre- dictions show very low levels of expression in all the types of cells studied indicating a possible greater theoretical propensity to suscepti- bility in cell lines such as NTERA-2, SCLC-21H, HepG2 and Vero6, and a lower propensity to infectivity in lines such as CaLu3, RT4, HEK293, A549 and U-251 MG. This corresponds to the RNA sequence profiles for our proteins of interest that showed high levels of expression in epithelial tissues outside the respiratory tract, indicative that cells of another nature may be particularly susceptible to infection by

SARS- CoV-2 [44,45]. In the case of the CD147 protein, a higher level of expression has been observed in brain cell lines and tissues, as was observed in this study [42].

It has also been shown that ACE2 and TMPRSS2, in addition to being expressed in cells associated with the lung and esophagus, are expressed at very high levels in human colorectal, stomach and liver-like cell, indicating that the gastrointestinal system could also be a possible route of infection by SARS-CoV-2 [44,45,47]. Interestingly, the ACE2 gene has been shown to be negatively regulated in some cells during SARS-CoV-1 infection [45]. These variations in the susceptibility of cells associated with the respiratory system have already been described in other studies that have demonstrated topographic differences in the replication of SARS-CoV-2 in the respiratory tract and almost undetectable levels of subgenomic viral material at the bronchoalveolar level [48].

The susceptibility profiles predicted as the most prevalent and capable of theoretically mediating the greater propensity to infection by SARS-CoV-2 based on the type, number and level of expression of the potential receptors coincided in their entirety with the presence of the CD147 and TMPRSS2 proteins, with or without the presence of ACE2 and the rest of the receptor candidates. This corresponds to what has been previously described in brain cell lines where it has been proposed that there may be a greater probability of infection with SARS- The viral spike protein has also been proposed to bind to the extracellular protease CD147 to mediate cell invasion by SARS-Cov-2 [49]. However, this is debated as purified recombinant spike protein is unable to interact with CD147 [5]. Of note, it has been experimentally confirmed that single nucleotide polymorphisms (SNPs) affect expression with a level of alteration that is often enough to induce phenotypic changes [49,50]. Therefore, although our results are based on experimental and tran- scriptomic kinetic data, they must be confirmed experimentally due to the inherent genetic variability.

Theoretical affinity of the spike protein and the proteins of interest associated with cellular susceptibility in SARS-CoV-2 infection

Thermodynamically favorable dockings were predicted between the spike protein (for all its conformations and regions) and the target proteins considered and described as potential facilitators of SARS-CoV-2 infection. As expected, the HDOCK algorithm predicted a thermodynamically favorable docking between the spike protein and the ACE2 and TMPRSS2 receptors described as the most important with a mean of ΔG = -294.21 kcal/mol and ΔG = -300.79 kcal/mol, respectively. The most favorable docking was predicted between the AGTR2 protein and the spike in all its conformations and tested regions, with a thermodynamic mean ΔG = -380.42 kcal/mol, being only surpassed by ≈-1 kcal/mol by the docking between ACE2 and the S1 spike region. Specifically, the most favorable docking of the AGTR2 protein with the spike regions was predicted for the S2 portion, followed by the binding with the TMPRSS2 (-262.47 kcal/mol) for the same S2 region, while for ACE2 it was with the S1 region (-352.45 kcal/mol) (see Table 3).

Table 3. Comparative analysis of the relative macromolecular binding affinity of the spike and each of its domains (S1 and S2) with each of the proteins of interest using different docking methods and scoring functions.

Protein	**PDB**	**HDOCK (kcal/mol)**				**MM/GBSA (kcal/mol)***			
		Spike (Open)	***Spike (Close)***	***S1***	***S2***	***Spike (Open)***	***Spike (Close)***	***S1***	***S2***
NRP1	7JJC	-274.84	-277.27	-269.17	-244.32	-61.81	-37.79	-53.60	-35.93
CD147	3I84	-218.99	-244.20	-210.55	-183.22	-46.47	-45.17	-46.57	-27.40
FURIN	6HZB	-290.38	-289.39	-292.35	-251.20	-49.54	-39.84	-44.43	-37.58
TMPRSS2	7MEQ	-341.26	-331.70	-267.71	-262.47	-31.83	-36.36	-49.94	-47.19
ACE2	6VW1	-304.94	-269.98	-352.45	-249.45	-21.90	-24.18	-49.53	-28.02
HSPA5	6ZYH	-266.58	-270.19	-233.86	-208.61	-40.04	-35.80	-29.75	-25.32
AGTR2	6JOD	-446.61	-366.70	-351.16	-357.21	-63.71	-60.49	-39.26	-68.29

For the spike protein, the open (PDB: 6VYB) and closed (PDB: 6VXX) conformations of the structure are considered; *, For the energy calculation with the molecular mechanics approach combined with generalized Born calculations and continuous surface area solvation (MM/GBSA) calculations to predict the binding free energy of the protein-protein complex per residue, the HawkDock web server was used.

Table 4. Comparative analysis of the residues called hot spots in the relative macromolecular junction of the spike domains (S1 and S2) with each of the proteins of interest.

PDB	**Spike**	***Residues determined as Binding Hot Spots****						
		NRP1 (*n*=30/24)	CD147 (*n*=27/18)	FURIN (*n*=21/27)	TMPRSS2 (*n*=28/31)	ACE2 (*n*=30/35)	HSPA5 (*n*=25/28)	AGTR2 (*n*=18/23)
7DEO	S1	Glu16(A), Glu21(A), Glu24(A), Glu43(A), Glu51(A), Asp128(A) / Glu340(B), Glu471(B), Glu484(B)	Glu340(A), Glu516(A) / Glu49(B), Glu64(B), Glu84(B), Glu92(B)	Asp23(A), Asp284(A), Asp322(A) / Asp405(B), Glu406(B)	Glu289(B), Glu329(B), Asp359(B), Asp491(B) / Asp405(A), Glu406(A), Asp420(A), Glu484(A)	Asp118(A), Glu122(A), Glu132(A), Asp139(A), Glu142(A), Asp274(A), Asp277(A), Asp281(A), Asp285(A), Asp317(A) / Glu340(B), Asp442(B), Glu465(B), Asp467(B), Glu471(B), GLU484(B)	Glu44(A), Glu46(A), Glu116(A), Glu117(A), Glu128(A), Glu131(A), Asp159(A) / Glu471(B), Glu484(B)	Glu340(B)
		NRP1 (*n*=24/16)	CD147 (*n*=16/24)	FURIN (*n*=19/17)	TMPRSS2 (*n*=21/31)	ACE2 (*n*=18/28)	HSPA5 (*n*=21/24)	AGTR2 (*n*=34/36)
7COT	S2	Glu58(A), Asp59(A), Glu113(A), Glu151(A) / Glu106(B)	Glu119(A), Asp123(A) / Asp32(B), Glu84(B), Glu92(B)	Asp430(B)	Glu406(B), Asp482(B) / Asp27(A), Glu106(A)	Asp195(A), Glu417(A), Glu509(A), Glu518(A), Glu553(A), Glu571(A) / Asp70(B)	Glu89(B), Glu121(B), Asp131(B), Asp281(B), Glu314(B) / Asp70(A), Asp92(A)	Glu257(A) / Asp92(B), Glu106(B)

*Only the residues present at the interface are shown that represent most of the binding free energy (*hot spots*) in the protein-protein complex according to the KFC (Knowledge-based FADE and Contacts) model. *n* (total number of residues at the interface (number of S domain residues/number of receptor type residues), including "*no hot spots*" residues).

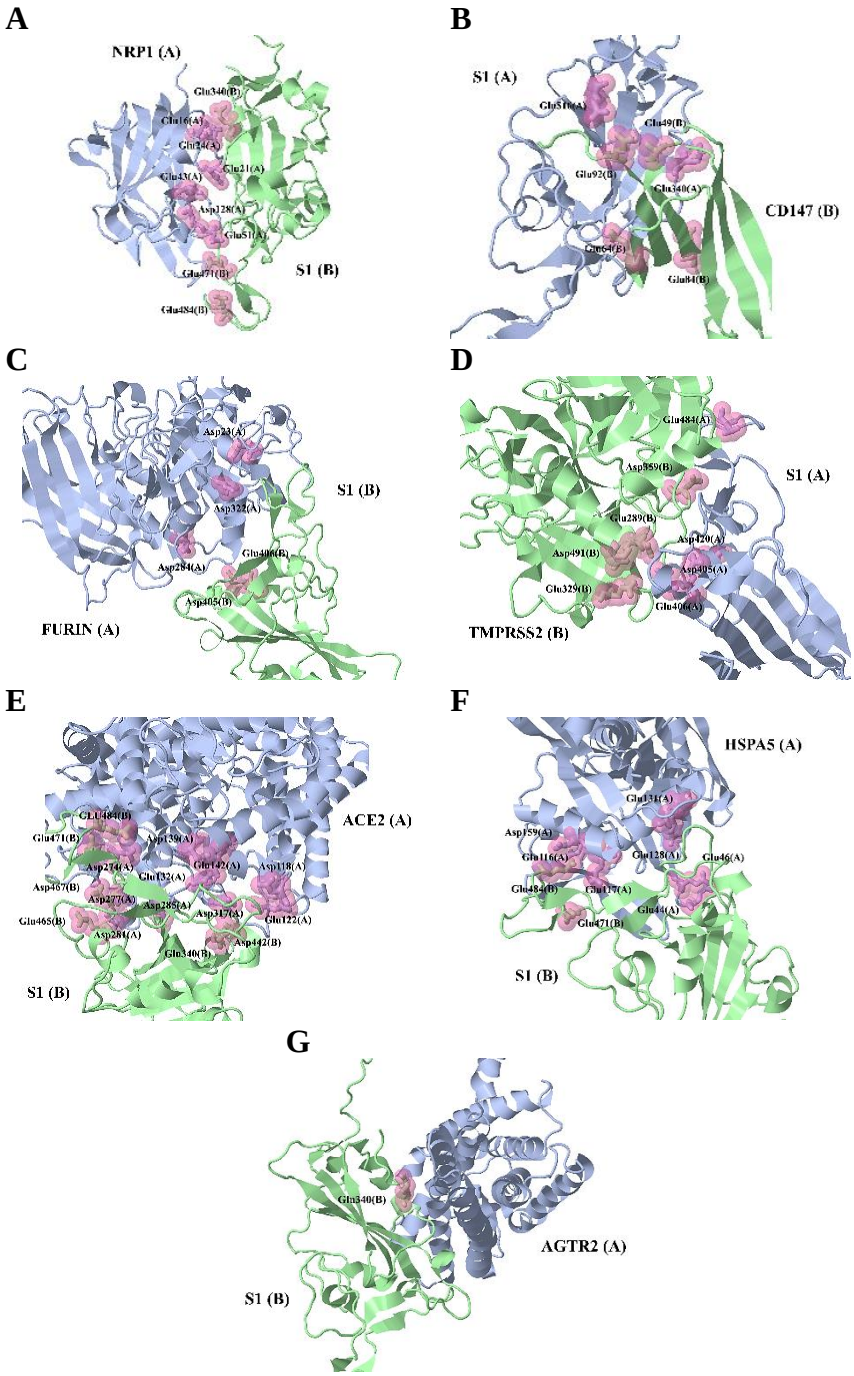

Figure 5. Graphic representation of the interface between the residues called hot spots in the relative macromolecular junction of the spike S1 domain with each of the proteins of interest. A) Neuropilin 1 (NRP1), B) Basigin2 (CD147), C) FURIN protease, D) Transmembrane serine 2 (TMPRSS2), E) Angiotensin converting enzyme 2 (ACE2), F) Heat shock protein A5 (HSPA5) and G) angiotensin II receptor type 2 (AGTR2). Only the residues present at the interface are shown that represent most of the binding free energy in the protein-protein complex according to the KFC (Knowledge-based FADE and Contacts) model.

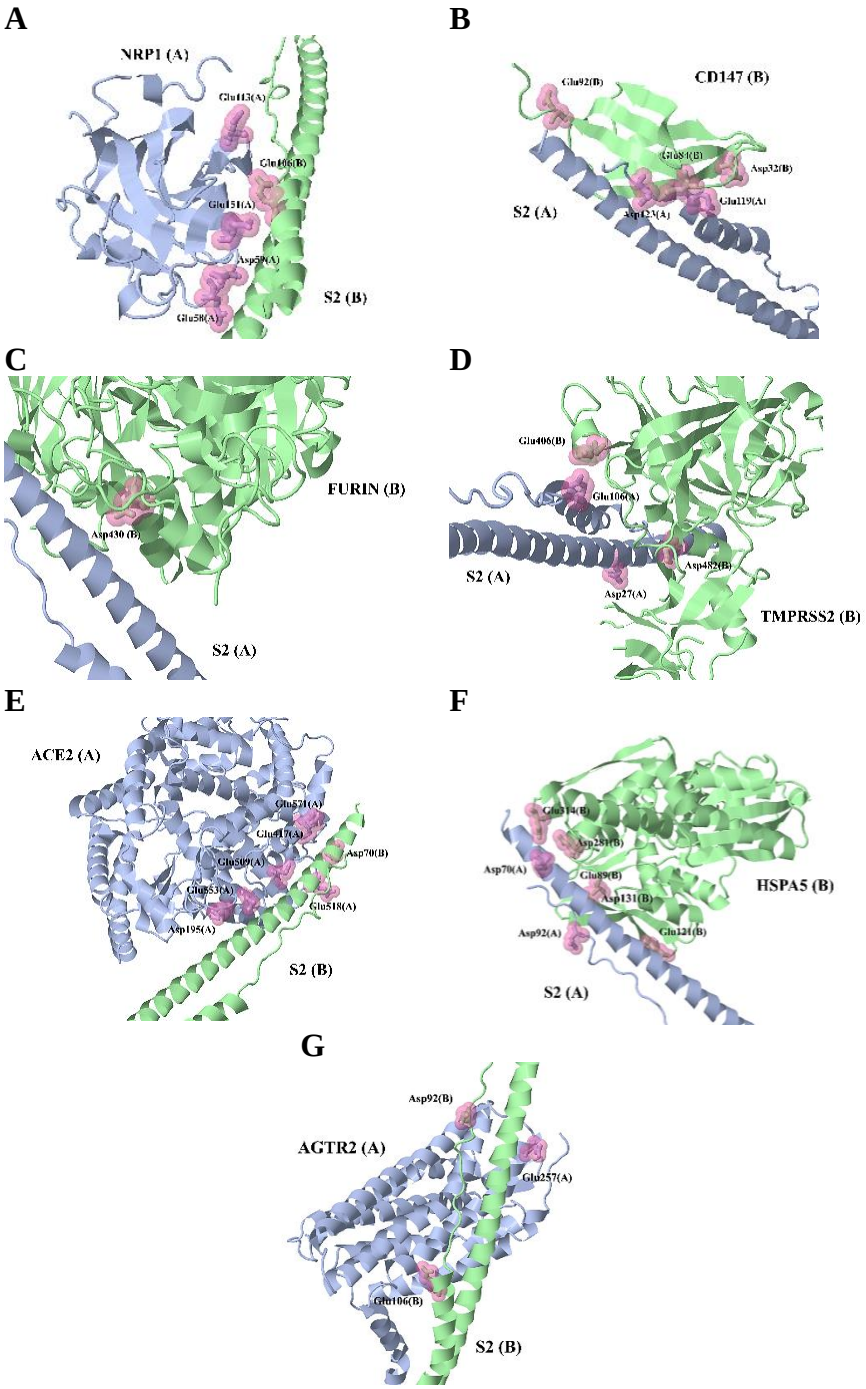

Figure 6. Graphic representation of the interface between the residues called hot spots in the relative macromolecular junction of the spike S2 domain with each of the proteins of interest. A) Neuropilin 1 (NRP1), B) Basigin2 (CD147), C) FURIN protease, D) Transmembrane serine 2 (TMPRSS2), E) Angiotensin converting enzyme 2 (ACE2), F) Heat shock protein A5 (HSPA5) and G) angiotensin II receptor type 2 (AGTR2). Only the residues present at the interface are shown that represent most of the binding free energy in the protein-protein complex according to the KFC (Knowledge-based FADE and Contacts) model.

It is important to highlight that according to the HDOCK algorithm, the S1 portion was the most favored spike region to establish thermodynamically feasible interactions with all the potential receptors considered, compared to the S2 region with a statistically significant difference ($p < 0.05$) at the energy level. These results correspond to the MM/GBSA measurements using the HawkDock method where it was predicted that for the S1 region a favorable docking was observed in a similar way with TMPRSS2 (-49.94 kcal/mol) and ACE2 (-49.53 kcal/mol). While for the S2 region the most thermodynamically favorable docking with the MM/GBSA approach was predicted with AGTR2 (-68.29 kcal/mol) followed by TMPRSS2 (-47.19 kcal/mol). In relation to the open and closed conformation of the spike protein, the most favorable interaction at the thermodynamic level was predicted with the AGTR2 protein with MM/GBSA values of -63.71 kcal/mol and -60.49 kcal/mol, respectively. It is important to note that the AGTR2 protein presented the most favorable docking with all the conformations and tested regions of the spike, except for the affinity for the S1 region and in comparison with the rest of the potential receptors, exhibiting a thermodynamic mean $\Delta G = -57.94$ kcal/mol (Table 3).

To study the contribution of the residues in the predicted interactions, the subset of residues designated as binding hot spots was determined, which constitute a small fraction of the interface residues in the protein-protein complexes that are determinants in the binding affinity calculated. In this sense, after calculating the binding hot spots considering characteristics such as interface solvation, atomic density and plasticity, the highest number of hot spots between the spike protein and the ACE2 receptor was predicted, which corresponds to the energy values of binding previously calculated, specifically more than 20% of the residues present in the interface correspond to residues located at a distance <4 Å (Table 4, Figures 5-6).

The highest number of residues with important interactions were predicted between the ACE2 receptor and the S1 region of the spike, contributing to the interface with ≈28% (10/35) and ≈20% (6/30) of residues with the most interaction probable, respectively, unlike the interaction between the ACE2 receptor and the S2 region in which ≈21% (6/28) and ≈5% (1/18) of binding hot spots were predicted, respectively. These results correspond to the highest affinity calculated between ACE2 and S1, compared to S2. The largest contribution in the interaction between the potential NRP1 receptor and the S1 region was predicted with ≈25% (6/24) and ≈10% (3/30) of binding hot spots, respectively. What is important considering that according to the Hawkdock algorithm the interaction between the receptor potential NRP1 and the S1 region was the most favorable energy level (ΔG = -292.35 kcal/mol) according to the MM/GBSA approach. 71% (5/7) of the dockings with the highest number of residues with important interactions were predicted in the S1 region, compared to ≈29% (2/7) in the S2 region (Table 4; Figures 5-6).

The only interfaces in which a larger subset of residues designated as binding hot spots was predicted in the S2 region were betweenthe potential CD147 receptor and the S2 region with ≈8% (2/24) and ≈19% (3/16) of binding hot spots, correspondingly, and between HSPA5 and the S2 region with ≈21% (5/24) and ≈10% (2/21) of binding hot spots, correspondingly. Similarly, ≈10% (2/21) of important residues in the S2 region were predicted with the TMPRSS2 protein. The interfaces with the least favorable interactions were predicted between the FURIN protease and the S2 region, and between AGTR2 and the S1 region in which, under the conditions of this study, it was not possible to determine more than 6 relevant binding hot spots (Table 4; Figures 5-6).

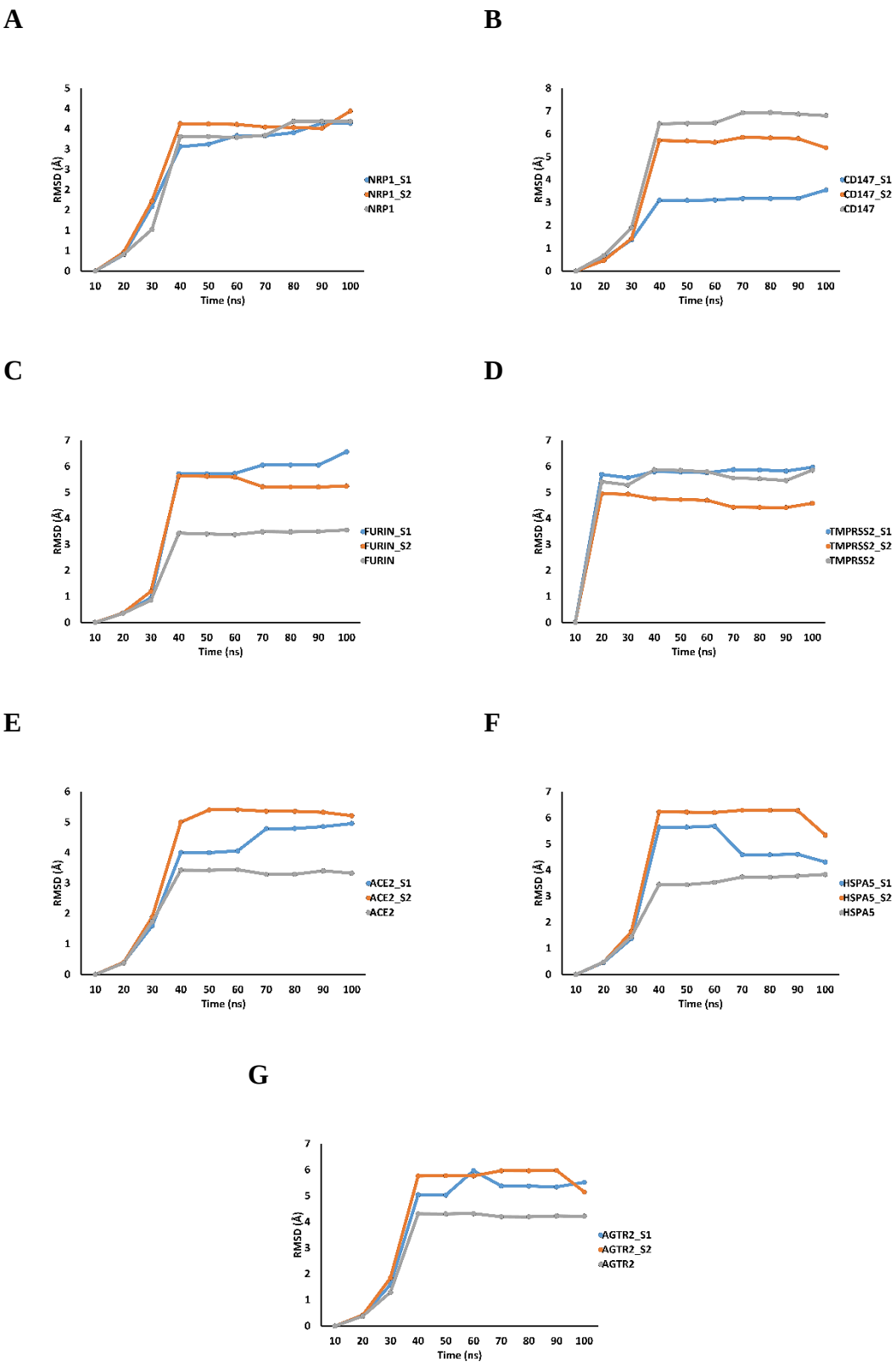

Figure 7. The mean squared deviations (RMSD) of Cα during each 10ns of MD simulation for each complex (total time 100ns). A) NRP1, B) CD147, C) FURIN, D) TMPRSS2, E) ACE2, F) HSPA5 and G) AGTR2. The gray lines represent the free receptor proteins, and the blue and orange lines the complexes with the S1 and S2 region, respectively.

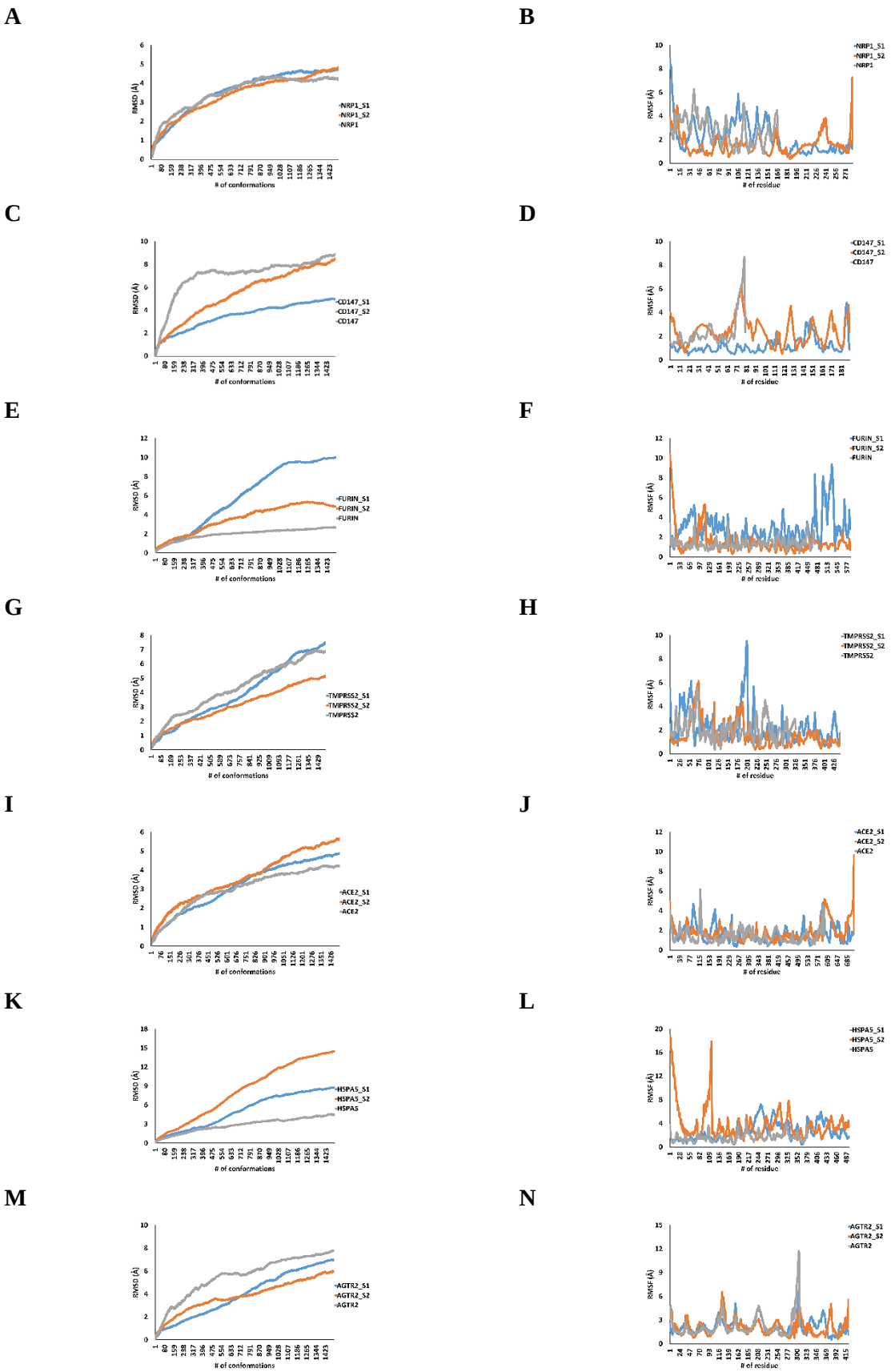

Figure 8. The root mean square deviations (RMSD) of Cα and root mean square fluctuations (RMSF) of individual amino acid residues relative to the starting frame during 100ns MD simulation of each complex using NMSim. NRP1 (A and B), CD147 (C and D), FURIN (E and F), TMPRSS2 (G and H), ACE2 (I and J), HSPA5 (K and L), AGTR2 (M and N). The gray lines represent the free receptor proteins, and the blue and orange lines the complexes with the S1 and S2 region, respectively.

The calculation of RMSD and RMSF allowed predicting that the complexes constructed were stable during the simulation period, exhibiting fluctuations ≤ 2.5 Å in all cases and with respect to the proteins proposed as possible free receptors of the S1 or S2 portions of the viral spike, except for the RMSD values of the CD147 + S1, FURIN + S1 and HSPA5 + S2 complexes. Specifically, the systems that presented the lowest fluctuations were the complexes formed by the S1 region of the spike and the TMPRSS2, NRP1 and ACE2 proteins with an RMSD and RMSF ≤ 0.5 Å, as well as the predicted RMSD and RMSF values for the complex made up of the S2 region and the ACE2 protein (RMSD and RMSF ≤ -0.6 Å). The complexes of the two regions of the spike with the proteins AGTR2, TMPRSS2, CD147, and of the S2 region with NRP1 were the systems in which the lowest fluctuations were predicted, with RMSD and RMSF values lower than those calculated from the free proteins (Figure 7; Figure 8).

In the case of the CD147 + S1 and CD147 + S2 complexes, the lowest fluctuations were predicted compared to those observed for their corresponding free protein with a difference of RMSD ≈ 3.5 Å and RMSD ≈ 1.5 Å, respectively, and a difference in terms of RMSF ≈ 1 Å and RMSF ≈ -0.01 Å, correspondingly. These results show that the docking with the portions of the viral spike protein studied can stabilize the trajectories of the complexes formed with the CD147 protein, unlike the free protein, which shows high flexibility and potential conformational deformability. On the contrary, the complexes formed by FURIN and HSPA5 with each of the regions of the viral spike showed a greater fluctuation with respect to the free proteins both in terms of RMSD and RMSF, specifically a difference was observed between the FURIN + S1 complex and the FURIN + S2 complex with respect to the control represented by values of RMSD ≈ -4 Å and RMSD ≈ -1 Å, respectively, and values of RMSF ≈ 2 Å and RMSF ≈ -0.1 Å, correspondingly. While the complex HSPA5 + S1 and HSPA5 + S2 with respect to the

control showed an RMSD ≈ -2 Å and RMSD ≈ -5 Å, respectively, as well as values of RMSF ≈ -1 Å and RMSF ≈ -2 Å, correspondingly. Although the trajectories calculated for the complexes built between the spike and the FURIN or HSPA5 proteins show greater fluctuation and, therefore, less stability, the RMSD and RMSF values show slightly favorable dockings throughout the simulation, especially between FURIN and the S2 region, and HSPA5 with the S1 region (Figure 7; Figure 8).

It is important to note that the ACE2 protein with which a stable and favorable binding energy was predicted, as well as the greater number of determining residues in the binding affinity in the complexes formed with any of the regions of the viral spike protein, is possibly expressed under the conditions of this study in a concentration of between 1-3% of the cells and only in ≈23% (3/13) of the cell models studied, specifically in the HepG2, RT4 and Vero6 lines, so their theoretical affinity and its level of expression could favor the infection of the cell types associated with the urinary and hepatic systems.

While due to its thermodynamically stable and favorable affinity, as well as due to a significant number of determining residues in the binding affinity for spike, the TMPRSS2 protein could be expressed depending on the cell type between 1-33% of the cells, and in the ≈ 38% (5/13) of the cell models considered, specifically in the CaCo2, RT4, Vero6, HepG2 and NTERA-2 lines, which would possibly also contribute to a better infection of the cell types associated with the urinary and liver systems, including to the gastrointestinal system. The NRP1 protein expressed in ≈31% (4/13) of the cell lines tested in a concentration of between 1-34% of the cells, depending on the cell type, presented a significant number of key residues for binding, as well as a stable and thermodynamically favorable interaction with the S1 region, which could facilitate infection in all cell lines considered except for NTERA-2, RPTEC-TERT1, RT4, HBEC3-KT and SCLC-21H, in this sense, these predictions would translate into a

susceptibility mediated by this receptor in only 50% and 25% of the cell types studied associated with the pulmonary and urinary systems, respectively.

Although HSPA5 was one of the candidates with the highest level of theoretical expression in all cell models considered (between 14-67% of cells) and presented one of the largest subsets of key residues for binding with a thermodynamically favorable affinity, the molecular dynamics was represented by junctions with many fluctuations in time for the S2 region, which could compromise the stability of said cross, although conditions for a stable affinity with the S1 portion were predicted. Although the FURIN protease presented a low number of residues considered as keys for the interaction under the conditions of this study, which corresponded to one of the least stable predicted junctions, it presented thermodynamically favorable junctions and important levels in all cell lines (between 4-27% depending on the cell type). Interestingly, although the AGTR2 protein presented the best values in terms of the relative affinity energy mean, as well as a stable binding throughout the simulation time, it established complexes with both the spike regions with the least number of hot spots of relevant binding, and presented the lowest level of detectable expression reported in all cell lines considered in this study, which could limit their participation as a determinant in the medication of cell susceptibility to infection.

The results of the predicted interactions correspond to those reported by other investigations that have shown that the S1/S2 complex of the viral spike protein adheres to the host cell membrane through the ACE2 enzyme and its interaction with the spike prior to priming of the transmembrane protease TMPRSS2, which cuts the S2 region to expose the viral fusion peptide that allows entry of the SARS-CoV2 virus into the cell [70,72,73] an affinity of ACE2 for the S1 domain and of TMPRSS2 for the S2 region also predicted in energy terms in this study. Simulations of the protein-protein interaction

based on the 3D structure have also revealed that AGTR2 shows a higher binding affinity with the spike protein of SARS-CoV-2 than ACE2, as observed in this study [74]. Furthermore, experimental and clinical findings also suggest that CD147 could act as a receptor that mediates the entry of SARS-CoV-2 through its interaction with the protein S of the viral spike [64,70]. Additionally, the predicted thermodynamically favorable dockings between the spike protein and FURIN have been associated with the presence of a unique motif in SARS-CoV-2-S that is easily recognized and hydrolyzed by FURIN, therefore, FURIN has been proposed as a therapeutic target. [75,76].

In the case of the controversial protein CD147 for which its participation in the infection by SARS-CoV-2 is still debated, the interaction between the candidate receptor CD147 and the peak protein of SARS-CoV-2 has been reported by surface plasmon resonance (SPR), enzyme-linked immunosorbent assays (ELISA), co-immunoprecipitation (Co-IP) and optimized negative staining electron microscopy (OpNS-EM). It has been noted that loss or blockade of CD147 in cell lines (such as those considered in this study) can inhibit the amplification of SARS-CoV-2, even proposing it as a new route of entry of virus independent of ACE2 [77]. These observations correspond to our results in which we predicted a thermodynamically favorable interaction between spike protein and CD147. A later study reported that it could not find evidence to support the role of CD147 as a spike-binding receptor, using specialized assays designed to detect even weak interactions for the proposed CD147-spike binding using various conformations of the CD147 protein [21]. However, a recent study that used cell lines similar to those considered in this study, demonstrated that the silencing of CD147 could reduce viral entry into cells, directly or indirectly through the reduction of ACE2 expression levels and therefore affected even spike protein levels, although apparently at the post-translational level [1]. And as with ACE2, it was demonstrated that at the RNA

level, the SARS-CoV-2 viral infection reduces both CD147 and ACE2, a phenomenon of negative regulation that, although not considered in this study, would further affect susceptibility cell mediated by this type of receptors as has been reported [1,67].

Similarly, the favorable energetic docking between the spike protein and the HSPA5 protein has been reported because this protein is considered a possible route for the adhesion and entry of SARS-CoV-2. Indeed, it has been described that the binding site in the SARS-CoV-2 peak is found in the C-terminal domain of the S1 region, as predicted in this study as its most favorable docking, and although they are required even experimental validations, other authors have pointed out that this receptor potential is key in endogenous coronaviruses and, due to its conservation, could mediate both the infectivity of SARS and a certain degree of cross-immunity against SARS-CoV-2 [78,79,80].

Therefore, as observed in this study and as suggested by other authors, binding free energy may support that the conformation and domain type of the SARS-CoV-2 spike protein leads to a stronger or weaker binding . However, a stronger receptor binding still cannot fully explain the high infectivity of this particular type of coronavirus. Therefore, it has been proposed that part of this infectivity may be mediated by the binding of the virus to other predominant receptors in other organs and that the lung may not be the earliest site of infection, as observed previously [75].

Theoretical diffusivity of the proteins of interest associated with cellular susceptibility in SARS-CoV-2 infection

The ability of the proteins considered to move in artificial cellular environments was studied simulating various types of cytoplasmic congestion, with the interest of seeing the impact of macromolecular congestion on the diffusion of proteins and finding out

whether there is any theoretical relationship with their levels of expression and relative affinity described above for the spike protein. In this sense, and after studying 5 conditions of cell-type cytoplasmic congestion and an aqueous model, a minimum translational diffusion coefficient (D_t) of $7.9x10^{-27}$ cm^2 s^{-1} and a maximum of $6.7x10^{-21}$ cm^2 s^{-1} was predicted for the cellular models, in contrast to a minimum D_t of $4.6x10^{-11}$ cm^2 s^{-1} and a maximum of $6.5x10^{-11}$ cm^2 s^{-1} for the aqueous model. Specifically, although all the diffusion values between the proteins studied were very close, it was predicted that under conditions of cytoplasmic congestion the NRP1 protein presented the best diffusion coefficient (D_t = $1.9x10^{-21}$ cm^2 s^{-1}) followed by CD147 (D_t = $1.8x10^{-21}$ cm^2 s^{-1}). Interestingly, the ACE2 protein (D_t = $1.2x10^{-21}$ cm^2 s^{-1}) presented the lowest diffusion coefficient in the cell models tested. In cell models, a mean diffusivity time of $\approx 3x10^{-10}$ sec was predicted, with a minimum time of $\approx 2x10^{-10}$ sec for TMPRSS2 and a maximum time of $\approx 5x10^{-10}$ sec for ACE2 (see Table 5).

In relation to the diffusion model in aqueous medium, the results were close and also showed CD147 as one of the proteins with the best diffusion coefficient (D_t = $8.3x10^{-11}$ cm^2 s^{-1}) followed by TMPRSS2 (D_t = $7.5x10^{-11}$ cm^2 s^{-1}) and FURIN (D_t = $7.1x10^{-11}$ cm^2 s^{-1}), keeping ACE2 as the protein with the lowest diffusivity (D_t = $4.6x10^{-11}$ cm^2 s^{-1}). Diffusivity in aqueous medium allowed predicting that a mean diffusion time of $\approx$13 sec, with a minimum and maximum time of $\approx$10 sec and $\approx$18 sec, respectively, diffusion rates that correspond to the CD147 and ACE2 proteins, respectively. The cell lines considered with experimental viscosity data in which the lowest and highest diffusion coefficients were predicted were A549 (D_t = $7.9x10^{-27}$ cm^2 s^{-1}) and ASTC-a-1 (D_t = $6.7x10^{-21}$ cm^2 s^{-1}) (the latter not considered in this study for the susceptibility analyzes), respectively. The HeLa cell line used in this study as a negative control for SARS-CoV-2 infectivity assays

presented a better diffusion coefficient (D_t = 2.5x10^{-22} cm^2 s^{-1}) than that predicted in A549 (line also considered in this study) (Table 5).

After making a preliminary approximation of the transmembrane flux (J) of these proteins under the various types of cytoplasmic congestion tested, similar results were predicted, with a mean of J = 9.9x10^{-16} mol cm^2 s^{-1} where the value minimum was predicted for ACE2 (J = 8.0x10-16 mol cm2 s-1) and the maximum for NRP1 (J = 1.3x10^{-15} mol cm^2 s^{-1}) followed by the predicted values for CD147 (J = 1.2x10^{-15} mol cm^2 s^{-1}). The cell lines considered in which the lowest and highest transmembrane flux was predicted were therefore A549 (J = 5.2x10-21 mol cm2 s-1) and ASTC-a-1 (J = 4.5x10^{-15} mol cm^2 s^{-1}), respectively. The HeLa control presented a better transmembrane flux (J = 1.6x10^{-16} mol cm^2 s^{-1}) than that calculated for A549 (see Table 6). Together with these results, a high correlation was found between the proteins studied with the lowest expression levels that presented the thermodynamically more favorable binding energy towards the viral spike protein with the lowest diffusion times and coefficients, as well as with the lowest hypothetical flux transmembrane (r = 0.80), these proteins being ACE2, AGTR2 and TMPRSS2, so the proteins described as potential receptors that presented the least thermodynamically favorable affinity and that are described with the highest expression levels had a high correlation with the best times and diffusion coefficients.

Table 5. Comparative analysis of the theoretical translational diffusion coefficient and the diffusion time of the proteins of interest associated with cellular susceptibility in SARS-CoV-2 infection.

Protein	**MW (kDa)**	***r* (Å)**	D_t **($cm^2\ s^{-1}$)**						***t* (seg)**
			HeLa	*3T3*	*H1299*	*A549*	*ASTC-a-1*	*Water*	
NRP1	79.20	15.33	$3.2x10^{-22}$	$5.8x10^{-22}$	$1.3x10^{-25}$	$1.0x10^{-26}$	$8.6x10^{-21}$	$6.3x10^{-11}$	13.2
CD147	21.15	16.06	$3.0x10^{-22}$	$5.6x10^{-22}$	$1.2x10^{-25}$	$9.6x10^{-27}$	$8.2x10^{-21}$	$8.3x10^{-11}$	10.0
FURIN	53.70	21.78	$2.2x10^{-22}$	$4.1x10^{-22}$	$8.9x10^{-26}$	$7.1x10^{-27}$	$6.0x10^{-21}$	$7.1x10^{-11}$	11.7
TMPRSS2	44.50	20.28	$2.4x10^{-22}$	$4.4x10^{-22}$	$9.6x10^{-26}$	$7.6x10^{-27}$	$6.5x10^{-21}$	$7.5x10^{-11}$	11.2
ACE2	192.75	24.29	$2.0x10^{-22}$	$3.7x10^{-22}$	$8.0x10^{-26}$	$6.4x10^{-27}$	$5.4x10^{-21}$	$4.6x10^{-11}$	18.0
HSPA5	84.75	21.38	$2.3x10^{-22}$	$4.2x10^{-22}$	$9.1x10^{-26}$	$7.2x10^{-27}$	$6.2x10^{-21}$	$6.0x10^{-11}$	13.8
AGTR2	92.57	21.94	$2.2x10^{-22}$	$4.1x10^{-22}$	$8.9x10^{-26}$	$7.0x10^{-27}$	$6.0x10^{-21}$	$5.6x10^{-11}$	14.8

D_t, translational diffusion coefficient. Neuropilin 1 (NRP1), Basigin2 (CD147), FURIN protease, Transmembrane Serine 2 (TMPRSS2), Angiotensin Converting Enzyme 2 (ACE2), Protein of Thermal shock A5 (HSPA5), Angiotensin II receptor type 2 (AGTR2).

Table 6. Comparative analysis of the theoretical transmembrane flux of the proteins of interest associated with cellular susceptibility in SARS-CoV-2 infection.

Protein	***J* (mol m^{-2} s^{-1})**				
	HeLa	*3T3*	*H1299*	*A549*	*ASTC-a-1*
NRP1	$2.1x10^{-16}$	$3.9x10^{-16}$	$8.5x10^{-20}$	$6.7x10^{-21}$	$5.7x10^{-15}$
CD147	$2.0x10^{-16}$	$3.7x10^{-16}$	$8.1x10^{-20}$	$6.4x10^{-21}$	$5.5x10^{-15}$
FURIN	$1.5x10^{-16}$	$2.7x10^{-16}$	$6.0x10^{-20}$	$4.7x10^{-21}$	$4.0x10^{-15}$
TMPRSS2	$1.6x10^{-16}$	$2.9x10^{-16}$	$6.4x10^{-20}$	$5.1x10^{-21}$	$4.3x10^{-15}$
ACE2	$1.3x10^{-16}$	$2.5x10^{-16}$	$5.3x10^{-20}$	$4.2x10^{-21}$	$3.6x10^{-15}$
HSPA5	$1.5x10^{-16}$	$2.8x10^{-16}$	$6.1x10^{-20}$	$4.8x10^{-21}$	$4.1x10^{-15}$
AGTR2	$1.5x10^{-16}$	$2.7x10^{-16}$	$5.9x10^{-20}$	$4.7x10^{-21}$	$4.0x10^{-15}$

J, diffusion flow. Neuropilin 1 (NRP1), Basigin2 (CD147), Protease FURIN, Transmembrane Serine 2 (TMPRSS2), Angiotensin Converting Enzyme 2 (ACE2), Protein of Thermal shock A5 (HSPA5), Angiotensin II receptor type 2 (AGTR2).

Analysis to establish diffusivity parameters such as the diffusion coefficient of structures associated with SARS-CoV-2 infection have focused on isolated regions of the spike protein [81,82], as well as of the viral particle [83]. Interestingly, there are few reports that consider the comparison between diffusivity parameters such as diffusion coefficient and speed, as well as the transmembrane flux of candidate receptors for SARS-CoV-2 infection, despite the fact that it has been described that membrane proteins such as the receptors studied can be affected by crowding effects similar to those of the solution phase [84,85]. Although the predicted diffusivity reports for these viral structures allow us to infer in general terms that all the simulated structures in this study in congested environments show a comparatively slow diffusion as a result of confined environments [86], it is important to note that transcription tends to be significantly improved by macromolecular crowding as a result of increased effective concentrations of enzymes and biomolecular reagents associated with transcription. In this sense, there may be a trade-off between the diffusion coefficient of the transcript and the transcription efficiency in congested environments [87,88]. However, this phenomenon is complex, because it has also been reported that increasing the viscosity of solutions can drastically reduce the diffusion coefficients of protein-type biomolecules by factors of up to 10 times, as predicted in this study by comparing the dynamics in aqueous medium and cytoplasmic medium [87].

Therefore, it is recommended to experimentally measure the effects of macromolecular crowding on the performance of transcription and diffusivity of the transcripts described as candidate receptors for SARS-CoV-2 infection, as has been suggested for other protein systems. For example, using strategies such as the construction of *de novo* designed *in vitro* protein expression systems, in which the macromolecular concentration could be varied by adding compatible solutes, or inert

polymers. Alternatively, one could take a diluted cytoplasmic cell extract and study the importance of crowding under physiological conditions by adding congestive agents as has been proposed to study the effect of crowding on protein signaling. The challenge of *in vivo* assays is represented in part by controlling the intracellular crowding concentration by manipulating extracellular conditions such as osmotic pressure with compatible solutes [85,89,90]. Additionally, it is recommended to validate the predictions made in this study with available scRNA-seq data from healthy humans, including a greater number of cell types and using other alternative databases such as *Gene Expression Omnibus* (GEO) and the *Tissue Stability Cell Atla* [36], as well as considering other factors or external stimuli [36,91].

CONCLUSIONS

Our predictions based on experimental and transcriptomic data indicate differential viral kinetic behavior dependent on the degree of cellular susceptibility, which is influenced to a greater extent by the type and level of expression of potentially receptor/mediator proteins associated with SARS-CoV-2 infection, and to a lesser extent by the number of potential receptor/mediator proteins expressed. In addition, a relationship was evidenced between lower expression levels and less favorable diffusivity parameters (diffusion coefficient, diffusion speed and transmembrane flux) with the thermodynamically more favorable interactions observed between the studied proteins and the viral spike, and vice versa, and we can conclude that part of the explanation for cellular susceptibility to infections caused by SARS-CoV-2 could be mediated by the ability of virus to stably bind to low-expression receptors in lung, which would suggest other potential sites for the earliest infection events.

Since this comparative research used previously published experimental data to be able to design expression systems and calculations of the effect of molecular crowding, it

is expected that it will stimulate quantitative analysis in future experiments and promote systematic investigation of the effect of phenomena such as crowding presented here. Especially as cancer cell lines naturally present a high degree of cytoplasmic congestion, which could affect the expression levels of the receptors considered. Additionally, the cells would affect the diffusivity of said receptors and the infection mediated by them, due to the increase in cytoplasmic viscosity, as already mentioned. To these limitations, we need to consider other factors such as healthy cells, polymorphisms or external stimuli (ethnic groups, age, sex, pathologies, exposure to pollutants, among others) in order to incorporate more factors that could alter the expression of the receptors studied and thus impact on susceptibility to SARS-CoV-2 infections, a viral mechanism that should continue to be further explored for the optimization of therapeutic approaches.

BIBLIOGRAPHIC REFERENCES

[1] Fenizia, C., Galbiati, S., Vanetti, C., Vago, R., Clerici, M., Tacchetti, C., Daniele, T. SARS-CoV-2 Entry: At the Crossroads of CD147 and ACE2. Cells 10 (1434) (2021), https://doi.org/10.3390/cells10061434.

[2] Kyrou, I., Randeva, H.S., Spandidos, D.A. et al. Not only ACE2—the quest for additional host cell mediators of SARS-CoV-2 infection: Neuropilin-1 (NRP1) as a novel SARS-CoV-2 host cell entry mediator implicated in COVID-19. Sig Transduct Target Ther 6 (21) (2021), https://doi.org/10.1038/s41392-020-00460-9.

[3] Dobrindt, K., Hoagland, D. A., Seah, C., Kassim, B., O'Shea, C. P., Murphy, A., ... Brennand, K. J. Common Genetic Variation in Humans Impacts In Vitro Susceptibility to SARS-CoV-2 Infection, Stem Cell Reports, 16 (3) (2021) 505-518, https://doi.org/10.1016/j.stemcr.2021.02.010.

[4] Jafary, F., Jafari, S., Ganjalikhany, M.R. In silico investigation of critical binding pattern in SARS-CoV-2 spike protein with angiotensin-converting enzyme 2. Sci Rep 11 (6927) (2021), https://doi.org/10.1038/s41598-021-86380-2.

[5] Hikmet, F., Méar, L., Edvinsson, Å., Micke, P., Uhlén, M., Lindskog, C. The protein expression profile of ACE2 in human tissues. Molecular systems biology, 16 (7) (2020) e9610, https://doi.org/10.15252/msb.20209610.

[6] Campbell, R. A., Boilard, E., Rondina, M. T. Is there a role for the ACE2 receptor in SARS-CoV-2 interactions with platelets?. Journal of Thrombosis and Haemostasis, 19 (1) (2021) 46-50, https://doi.org/10.1111/jth.15156.

[7] Keyaerts, E., Vijgen, L., Maes, P., Neyts, J., Van Ranst, M. Growth kinetics of SARS-coronavirus in Vero E6 cells. Biochemical and biophysical research communications, 329 (3) (2005) 1147-1151, https://doi.org/10.1016/j.bbrc.2005.02.085.

[8] Bar-On, Y. M., Flamholz, A., Phillips, R., Milo, R. Science Forum: SARS-CoV-2 (COVID-19) by the numbers. elife, 9 (2020) e57309, https://doi.org/10.7554/eLife.57309.

[9] Liu, F., Han, K., Blair, R., Kenst, K., Qin, Z., Upcin, B., ... Qin, X. SARS-CoV-2 Infects Endothelial Cells In Vivo and In Vitro. Frontiers in Cellular and Infection Microbiology, 11 (2021), https://doi.org/10.3389/fcimb.2021.701278.

[10] Kumar, N. D., Ter Ellen, B., Bouma, E. M., Troost, B., van de Pol, D. P., van der Ende-Metselaar, H. H., ... Smit, J. Moxidectin and ivermectin inhibit SARS-CoV-2 replication in Vero E6 cells but not in human primary airway epithelium cells. Antimicrobial Agents and Chemotherapy (2021), AAC-01543, https://doi.org/10.1128/AAC.01543-21.

[11] Gonçalves, A., Bertrand, J., Ke, R., Comets, E., De Lamballerie, X., Malvy, D., ... Guedj, J. Timing of antiviral treatment initiation is critical to reduce SARS-CoV-2 viral load. CPT: pharmacometrics & systems pharmacology, *9* (9) (2020) 509-514, https://doi.org/10.1002/psp4.12543.

[12] Nguyen, V. K., Binder, S. C., Boianelli, A., Meyer-Hermann, M., Hernandez-Vargas, E. A. Ebola virus infection modeling and identifiability problems. Frontiers in microbiology, 6 (2015) 257, https://doi.org/10.3389/fmicb.2015.00257.

[13] Baccam, P., Beauchemin, C., Macken, C. A., Hayden, F. G., Perelson, A. S. Kinetics of influenza A virus infection in humans. Journal of virology, 80 (15) (2006) 7590-7599, https://doi.org/10.1128/JVI.01623-05.

[14] Hernandez-Vargas, E. A. Parameter estimation in mathematical models of viral infections using R. Influenza Virus, 1836 (2018) 531-549, https://doi.org/10.1007/978-1-4939-8678-1_25.

[15] Abernathy, Z., Abernathy, K., Stevens, J. A mathematical model for tumor growth and treatment using virotherapy. AIMS Math, 5 (5) (2020) 4136-50, http://dx.doi.org/10.3934/math.2020265.

[16] Digre, A., Lindskog, C. The human protein atlas—spatial localization of the human proteome in health and disease. Protein Science, 30 (1) (2021) 218-233, https://doi.org/10.1002/pro.3987.

[17] Cuervo, N. Z., Grandvaux, N. ACE2: Evidence of role as entry receptor for SARS-CoV-2 and implications in comorbidities. Elife, 9 (2020) e61390, https://doi.org/10.7554/eLife.61390.

[18] Bairoch, A. The cellosaurus, a cell-line knowledge resource. Journal of biomolecular techniques: JBT, 29 (2) (2018) 25-38, https://dx.doi.org/10.7171%2Fjbt.18-2902-002.

[19] Caly, L., Druce, J. D., Catton, M. G., Jans, D. A., Wagstaff, K. M. The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 in vitro. Antiviral research, 178 (2020) 104787, https://doi.org/10.1016/j.antiviral.2020.104787.

[20] Jeffreys, L., Pennington, S. H., Duggan, J., Breen, A., Jinks, J., Ardrey, A., ... Biagini, G. Remdesivir-Ivermectin combination displays synergistic interaction with improved in vitro antiviral activity against SARS-CoV-2. BioRxiv (2020), https://doi.org/10.1101/2020.12.23.424232.

[21] Shilts, J., Crozier, T. W., Greenwood, E. J., Lehner, P. J., Wright, G. J. No evidence for basigin/CD147 as a direct SARS-CoV-2 spike binding receptor. Scientific reports, 11 (1) (2021) 1-10, https://doi.org/10.1038/s41598-020-80464-1.

[22] Chu, H., Chan, J. F. W., Yuen, T. T. T., Shuai, H., Yuan, S., Wang, Y., ... Yuen, K. Y. Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. The Lancet Microbe, 1 (1) (2020) e14-e23, https://doi.org/10.1016/S2666-5247(20)30004-5.

[23] Peixoto, A., Monteiro, M., Rocha, B., Veiga-Fernandes, H. Quantification of multiple gene expression in individual cells. Genome research, 14 (10a) (2004) 1938-1947, https://doi.org/10.1101/gr.2890204.

[24] Koutsoudakis, G., Herrmann, E., Kallis, S., Bartenschlager, R., Pietschmann, T. The level of CD81 cell surface expression is a key determinant for productive entry of hepatitis C virus into host cells. Journal of virology, 81 (2) (2007) 588-598, https://doi.org/10.1128/JVI.01534-06.

[25] Daud, I. I., Scott, M. E., Ma, Y. Shiboski, S. Farhat, S. Moscicki, A. B. Association between toll-like receptor expression and human papillomavirus type 16 persistence. International journal of cancer, 128 (4) (2011) 879-886, https://doi.org/10.1002/ijc.25400.

[26] Lecarpentier, Y. Vallée, A. The key role of the level of ACE2 gene expression in SARS-CoV-2 infection. Aging, 13 (11) (2021) 14552–14556, https://dx.doi.org/10.18632%2Faging.203181.

[27] Matsuyama, S., Kawase, M., Nao, N., Shirato, K., Ujike, M., Kamitani, W., ... Fukushi, S. The inhaled steroid ciclesonide blocks SARS-CoV-2 RNA replication by targeting the viral replication-transcription complex in cultured cells. Journal of virology, 95 (1) (2020) e01648-20, https://doi.org/10.1128/JVI.01648-20.

[28] Modrof, J., Kerschbaum, A., Farcet, M. R., Niemeyer, D., Corman, V. M., Kreil, T. R. SARS-CoV-2 and the safety margins of cell-based biological medicinal products. Biologicals, 68 (2020) 122-124, https://doi.org/10.1016/j.biologicals.2020.08.010.

[29] Schroeder, S., Pott, F., Niemeyer, D., Veith, T., Richter, A., Muth, D., ... Drosten, C. Interferon antagonism by SARS-CoV-2: a functional study using reverse genetics. The Lancet Microbe, 2 (5) (2021) e210-e218, https://doi.org/10.1016/S2666-5247(21)00027-6.

[30] Appelberg, S., Gupta, S., Svensson Akusjärvi, S., Ambikan, A. T., Mikaeloff, F., Saccon, E., ... Neogi, U. Dysregulation in Akt/mTOR/HIF-1 signaling identified by proteo-transcriptomics of SARS-CoV-2 infected cells. Emerging microbes & infections, 9 (1) (2020) 1748-1760, https://doi.org/10.1080/22221751.2020.1799723.

[31] ter Ellen, B. M., Kumar, N. D., Bouma, E. M., Troost, B., van de Pol, D. P., van der Ende-Metselaar, H. H., ... Smit, J. M. Resveratrol and pterostilbene potently inhibit SARS-CoV-2 replication in vitro. BioRxiv, (2021), https://doi.org/10.1101/2020.09.24.285940.

[32] Kishimoto, M., Uemura, K., Sanaki, T., Sato, A., Hall, W. W., Kariwa, H., ... Sasaki, M. TMPRSS11D and TMPRSS13 activate the SARS-CoV-2 spike protein. Viruses, 13 (3) (2021) 384, https://doi.org/10.3390/v13030384.

[33] Pontelli, M. C., Castro, I. A., Martins, R. B., Veras, F. P., La Serra, L., Nascimento, D. C., ... Arruda, E. Infection of human lymphomononuclear cells by SARS-CoV-2. BioRxiv, (2020), https://doi.org/10.1101/2020.07.28.225912.

[34] Fiege, J. K., Thiede, J. M., Nanda, H. A., Matchett, W. E., Moore, P. J., Montanari, N. R., ... Langlois, R. A. Single cell resolution of SARS-CoV-2 tropism, antiviral responses, and susceptibility to therapies in primary human airway epithelium. PLoS pathogens, 17 (1) (2021) e1009292, https://doi.org/10.1371/journal.ppat.1009292.

[35] Zhang, B. Z., Chu, H., Han, S., Shuai, H., Deng, J., Hu, Y. F., ... Huang, J. D. SARS-CoV-2 infects human neural progenitor cells and brain organoids. Cell research, 30 (10) (2020) 928-931, https://doi.org/10.1038/s41422-020-0390-x.

[36] Qi, J., Zhou, Y., Hua, J., Zhang, L., Bian, J., Liu, B., ... Jin, S. The scRNA-seq expression profiling of the receptor ACE2 and the cellular protease TMPRSS2 reveals human organs susceptible to SARS-CoV-2 infection. International Journal of Environmental Research and Public Health, 18 (1) (2021) 284, https://doi.org/10.3390/ijerph18010284.

[37] Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., ... Richardson, D. C. MolProbity: More and better reference data for improved all-

atom structure validation. Protein Science 27 (1) (2018) 293-315, https://doi.org/10.1002/pro.3330.

[38] Yan Y, Tao H, He J, Huang S-Y. The HDOCK server for integrated protein-protein docking. Nature Protocols, 15 (2020) 1829–1852, https://doi.org/10.1038/s41596-020-0312-x.

[39] Weng GQ, Wang EC, Wang Z, Liu H, Li D, Zhu F, Hou TJ. HawkDock: a web server to predict and analyze the structures of protein-protein complexes based on computational docking and MM/GBSA. Nucleic Acids Research, 47 (W1) (2019) W322-W330, https://doi.org/10.1093/nar/gkz397.

[40] González-Paz, L., Hurtado-León, M. L., Lossada, C., Fernández-Materán, F. V., Vera-Villalobos, J., Loroño, M., ... Alvarado, Y. J. Comparative study of the interaction of ivermectin with proteins of interest associated with SARS-CoV-2: A computational and biophysical approach. Biophysical chemistry, 278 (2021) 106677, https://doi.org/10.1016/j.bpc.2021.106677.

[41] Malhotra S, Mathew O, Sowdhamini R. DOCKSCORE: a webserver for ranking protein-protein docked poses. BMC Bioinformatics, 16 (127) (2015), https://doi.org/10.1186/s12859-015-0572-6.

[42] Zhu, X. Mitchell, J. C. KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. Proteins: Structure, Function, and Bioinformatics, 79 (9) (2011) 2671-2683, https://doi.org/10.1002/prot.23094.

[43] Kasahara, K. Terazawa, H. Itaya, H., Goto, S., Nakamura, H., Takahashi, T., Higo, J. myPresto/omegagene 2020: a molecular dynamics simulation engine for virtual-system

coupled sampling, Biophys. Physicobiol. 17 (2020) 140–146, https://doi.org/10.2142/biophysico.BSJ-2020013.

[44] Kalwarczyk, T., Ziebacz, N., Bielejewska, A., Zaboklicka, E., Koynov, K., Szymanski, J., ... Hołyst, R. Comparative analysis of viscosity of complex liquids and cytoplasm of mammalian cells at the nanoscale. Nano letters, 11 (5) (2011) 2157-2163, https://doi.org/10.1021/nl2008218.

[45] Regan, K., Dotterweich, R., Ricketts, S. Robertson-Anderson, R. M. Diffusion and conformational dynamics of single DNA molecules crowded by cytoskeletal proteins. Journal of Undergraduate Reports in Physics, 28 (1) (2018) 100005, https://doi.org/10.1063/1.5109559.

[46] Liang, L., Wang, X., Da, X., Chen, T. Chen, W. R. Noninvasive determination of cell nucleoplasmic viscosity by fluorescence correlation spectroscopy. Journal of biomedical optics, 14 (2) (2009) 024013, https://doi.org/10.1117/1.3088141.

[47] Wang, K., Sun, X. H., Zhang, Y., Zhang, T., Zheng, Y., Wei, Y. C., ... Chen, J. Characterization of cytoplasmic viscosity of hundreds of single tumour cells based on micropipette aspiration. Royal Society open science, 6 (3) (2019) 181707, https://doi.org/10.1098/rsos.181707.

[48] Cunningham, J., Estrella, V., Lloyd, M., Gillies, R., Frieden, B. R., Gatenby, R. Intracellular electric field and pH optimize protein localization and movement. PloS one, 7 (5) (2012) e36894, https://doi.org/10.1371/journal.pone.0036894.

[49] Junker, N. O., Vaghefikia, F., Albarghash, A., Höfig, H., Kempe, D., Walter, J., ... Fitter, J. Impact of molecular crowding on translational mobility and conformational properties of biological macromolecules. The Journal of Physical Chemistry B, 123 (21) (2019) 4477-4486, https://doi.org/10.1021/acs.jpcb.9b01239.

[50] von Bülow, S., Siggel, M., Linke, M., Hummer, G. Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. Proceedings of the National Academy of Sciences, 116 (20) (2019) 9843-9852, https://doi.org/10.1073/pnas.1817564116.

[51] Garbett, N. C., Mekmaysy, C. S., Chaires, J. B. Sedimentation velocity ultracentrifugation analysis for hydrodynamic characterization of G-quadruplex structures. G-Quadruplex DNA. Humana Press, (2010) 97-120, https://doi.org/10.1007/978-1-59745-363-9_7.

[52] Fleming, P. J., Fleming, K. G. HullRad: Fast calculations of folded and disordered protein and nucleic acid hydrodynamic properties. Biophysical journal, 114 (4) (2018) 856-869, https://doi.org/10.1016/j.bpj.2018.01.002.

[53] Yehl, C. J., Zydney, A. L. Characterization of dextran transport and molecular weight cutoff (MWCO) of large pore size hollow fiber ultrafiltration membranes. Journal of Membrane Science, 622 (2021) 119025, https://doi.org/10.1016/j.memsci.2020.119025.

[54] Phillips, R., Kondev, J., Theriot, J. Physical Biology of the Cell. Garland Science, (2009), ISBN: 9780815344506.

[55] Schavemaker, P. E., Boersma, A. J., Poolman, B. How important is protein diffusion in prokaryotes?. Frontiers in molecular biosciences, 5 (2018) 93, https://doi.org/10.3389/fmolb.2018.00093.

[56] Brand, J., Voigt, K., Zochowski, B., Kulozik, U. Lysozyme fractionation from egg white at pilot scale by means of tangential flow membrane adsorbers: investigation of the flow conditions. Journal of Chromatography A, 1438 (2016) 143-149, https://doi.org/10.1016/j.chroma.2016.02.023.

[57] Lasave, L. C., Urteaga, R., Koropecki, R. R., Gonzalez, V. D., Arce, R. D. Real-time study of protein adsorption kinetics in porous silicon. Colloids and Surfaces B: Biointerfaces, 111 (2013) 354-359, https://doi.org/10.1016/j.colsurfb.2013.06.024.

[58] Lin, Y. H., Hung, G. Y., Wu, L. C., Chen, S. W., Lin, L. Y., Horng, J. L. Anion exchanger 1b in stereocilia is required for the functioning of mechanotransducer channels in lateral-line hair cells of zebrafish. PloS one, 10 (2) (2015) e0117041, https://doi.org/10.1371/journal.pone.0117041.

[59] Yan, S., Jiao, C., McLamore, E. S., Wang, N., Yao, H., Shen, Y. Insect herbivory of leaves affects the auxin flux along root apices in Arabidopsis thaliana. Journal of Plant Growth Regulation, 36 (4) (2017) 846-854, https://doi.org/10.1007/s00344-017-9688-4.

[60] Xu, Z., Gao, L., Chen, P., Yan, L. T. Diffusive transport of nanoscale objects through cell membranes: a computational perspective. Soft Matter, 16 (16) (2020) 3869-3881, https://doi.org/10.1039/C9SM02338K .

[61] Leonard, A. N., Wang, E., Monje-Galvan, V., Klauda, J. B. Developing and testing of lipid force fields with applications to modeling cellular membranes. Chemical reviews, 119 (9) (2019) 6227-6269, https://doi.org/10.1021/acs.chemrev.8b00384.

[62] Song, J., Li, Y., Huang, X., Chen, Z., Li, Y., Liu, C., ... Duan, X. Systematic analysis of ACE2 and TMPRSS2 expression in salivary glands reveals underlying transmission mechanism caused by SARS-CoV-2. Journal of medical virology, 92 (11) (2020) 2556-2566, https://doi.org/10.1002/jmv.26045.

[63] Bradding, P., Richardson, M., Hinks, T. S., Howarth, P. H., Choy, D. F., Arron, J. R., ... Siddiqui, S. ACE2, TMPRSS2, and furin gene expression in the airways of people with asthma—implications for COVID-19. The Journal of allergy and clinical immunology, 146 (1) (2020) 208 –211, https://dx.doi.org/10.1016%2Fj.jaci.2020.05.013.

[64] Qiao, J., Li, W., Bao, J., Peng, Q., Wen, D., Wang, J., Sun, B. The expression of SARS-CoV-2 receptor ACE2 and CD147, and protease TMPRSS2 in human and mouse brain cells and mouse brain tissues. Biochemical and Biophysical Research Communications, 533 (4) (2020) 867-871, https://doi.org/10.1016/j.bbrc.2020.09.042.

[65] Leonardi, A., Rosani, U., Brun, P. Ocular surface expression of SARS-CoV-2 receptors. Ocular immunology and inflammation, 28 (5) (2020) 735-738, https://doi.org/10.1080/09273948.2020.1772314.

[66] Liu, C., Wang, K., Zhang, M., Hu, X., Hu, T., Liu, Y., ... Yue, J. High expression of ACE2 and TMPRSS2 and clinical characteristics of COVID-19 in colorectal cancer patients. NPJ precision oncology, 5 (1) (2021) 1-7, https://doi.org/10.1038/s41698-020-00139-y.

[67] Nersisyan, S., Shkurnikov, M., Turchinovich, A., Knyazev, E., Tonevitsky, A. Integrative analysis of miRNA and mRNA sequencing data reveals potential regulatory mechanisms of ACE2 and TMPRSS2. PLoS One, 15 (7) (2020) e0235987, https://doi.org/10.1371/journal.pone.0235987.

[68] Matusiak, M., Schürch, C. M. Expression of SARS-CoV-2 entry receptors in the respiratory tract of healthy individuals, smokers and asthmatics. Respiratory research, 21 (1) (2020) 1-6, https://doi.org/10.1186/s12931-020-01521-x.

[69] Sulaiman, I., Chung, M., Angel, L., Koralov, S., Wu, B., Yeung, S., ... Segal, L. Microbial signatures in the lower airways of mechanically ventilated COVID19 patients associated with poor clinical outcome. (2021) https://doi.org/10.21203/rs.3.rs-266050/v1

[70] Chiappelli, F. CoViD-19 Susceptibility. Bioinformation, 16 (7) (2020) 501-504, https://dx.doi.org/10.6026%2F97320630016501.

[71] Piva, F., Sabanovic, B., Cecati, M., Giulietti, M. Expression and co-expression analyses of TMPRSS2, a key element in COVID-19. European Journal of Clinical Microbiology & Infectious Diseases, 40 (2) (2021) 451-455, https://doi.org/10.1007/s10096-020-04089-y.

[72] Hussain, M., Jabeen, N., Amanullah, A., Baig, A. A., Aziz, B., Shabbir, S., ... Uddin, N. Molecular docking between human TMPRSS2 and SARS-CoV-2 spike protein: conformation and intermolecular interactions. AIMS microbiology, 6 (3) (2020) 350-360, https://dx.doi.org/10.3934%2Fmicrobiol.2020021.

[73] Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., ... Pöhlmann, S. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. cell, 181 (2) (2020) 271-280, https://doi.org/10.1016/j.cell.2020.02.052.

[74] Cui, C., Huang, C., Zhou, W., Ji, X., Zhang, F., Wang, L., ... Cui, Q. AGTR2, one possible novel key gene for the entry of 2019-nCoV into human cells. Preprints (2020) 2020020194, https://doi.org/10.20944/preprints202002.0194.v1.

[75] Wu, C., Zheng, M., Yang, Y., Gu, X., Yang, K., Li, M., ... Li, H. (2020). Furin: a potential therapeutic target for COVID-19. Iscience, 23 (10) (2020) 101642, https://doi.org/10.1016/j.isci.2020.101642.

[76] Cheng, Y. W., Chao, T. L., Li, C. L., Chiu, M. F., Kao, H. C., Wang, S. H., ... Yeh, S. H. Furin inhibitors block SARS-CoV-2 spike protein cleavage to suppress virus production and cytopathic effects. Cell reports, 33 (2) (2020) 108254, https://doi.org/10.1016/j.celrep.2020.108254.

[77] Wang, K., Chen, W., Zhang, Z., Deng, Y., Lian, J. Q., Du, P., ... Chen, Z. N. CD147-spike protein is a novel route for SARS-CoV-2 infection to host cells. Signal transduction and targeted therapy, 5 (1) (2020) 1-10, https://doi.org/10.1038/s41392-020-00426-x.

[78] Elfiky, A. A. SARS-CoV-2 spike-heat shock protein A5 (GRP78) recognition may be related to the immersed human coronaviruses. Frontiers in pharmacology, 11 (2020) 1997, https://doi.org/10.3389/fphar.2020.577467.

[79] Allam, L., Ghrifi, F., Mohammed, H., El Hafidi, N., El Jaoudi, R., El Harti, J., ... Ibrahimi, A. Targeting the GRP78-dependant SARS-CoV-2 cell entry by peptides and small molecules. Bioinformatics and Biology Insights, 14 (2020) 1177932220965505, https://doi.org/10.1177%2F1177932220965505.

[80] Ibrahim, I. M., Elfiky, A. A., Elgohary, A. M. Recognition through GRP78 is enhanced in the UK, South African, and Brazilian variants of SARS-CoV-2; An in-silico perspective. Biochemical and Biophysical Research Communications, 562 (2021) 89-93, https://doi.org/10.1016/j.bbrc.2021.05.058.

[81] Das, A., Vishvakarma, V., Dey, A., Dey, S., Gupta, A., Das, M., ... Maiti, S. Biophysical properties of the isolated spike protein binding helix of human ACE2. Biophysical journal, 120 (14) (2021) 2785-2792, https://doi.org/10.1016/j.bpj.2021.06.017.

[82] Liu, Z., Xia, M., Chai, Z., Wang, D. Tracing the driving forces responsible for the remarkable infectivity of 2019-nCoV: 1. Receptor binding domain in its bound and unbound states. Physical Chemistry Chemical Physics, 22 (48) (2020) 28277-28285, https://doi.org/10.1039/D0CP04435K.

[83] Shang, Z., Chan, S. Y., Liu, W. J., Li, P., Huang, W. Recent Insights into Emerging Coronavirus: SARS-CoV-2. ACS Infectious Diseases, 7 (6) (2021) 1369–1388, https://doi.org/10.1021/acsinfecdis.0c00646.

[84] Zhou, H. X. Crowding effects of membrane proteins. The journal of physical chemistry B, 113 (23) (2009) 7995-8005, https://doi.org/10.1021/jp8107446.

[85] Schlame, M. Protein crowding in the inner mitochondrial membrane. Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1862 (1) (2021) 148305, https://doi.org/10.1016/j.bbabio.2020.148305.

[86] Gorshkov, K., Susumu, K., Chen, J., Xu, M., Pradhan, M., Zhu, W., ... Oh, E. Quantum dot-conjugated SARS-CoV-2 spike pseudo-virions enable tracking of angiotensin converting enzyme 2 binding and endocytosis. ACS nano, 14 (9) (2020) 12234-12247, https://doi.org/10.1021/acsnano.0c05975.

[87] Ge, X., Luo, D., Xu, J. Cell-free protein expression under macromolecular crowding conditions. PloS one, 6 (12) (2011) e28707, https://doi.org/10.1371/journal.pone.0028707.

[88] Morelli, M. J., Allen, R. J., Ten Wolde, P. R. Effects of macromolecular crowding on genetic networks. Biophysical journal, 101 (12) (2011) 2882-2891, https://doi.org/10.1016/j.bpj.2011.10.053.

[89] Dupuis, N. F., Holmstrom, E. D., Nesbitt, D. J. Molecular-crowding effects on single-molecule RNA folding/unfolding thermodynamics and kinetics. Proceedings of the National Academy of Sciences, 111 (23) (2014) 8464-8469, https://doi.org/10.1073/pnas.1316039111.

[90] Jack, A., Ferro, L. S., Trnka, M. J., Wehri, E., Nadgir, A., Nguyenla, X., ... Yildiz, A. (2021). SARS-CoV-2 nucleocapsid protein forms condensates with viral genomic RNA. PLOS Biology, 19 (10) (2021) e3001425, https://doi.org/10.1371/journal.pbio.3001425.

[91] Chen, J., Jiang, Q., Xia, X., Liu, K., Yu, Z., Tao, W., ... Han, J. D. J. Individual variation of the SARS-CoV-2 receptor ACE2 gene expression and regulation. Aging cell, 19 (7) (2020) e13168, https://doi.org/10.1111/acel.13168.