Small Molecule Stabilization of non-native Protein-Protein Interactions of SARS-CoV-2 N Protein as a Mechanism of Action against COVID-19

Julián F. Fernández[†,‡] and Martín J Lavecchia[*,¶]

†*Universidad de Buenos Aires, Departamento de Química Orgánica, Facultad de Ciencias Exactas y Naturales, Ciudad Universitaria, Pabellón 2 (1428), Buenos Aires, Argentina.*
‡*CONICET-Universidad de Buenos Aires, Unidad de Microanálisis y Métodos Físicos en Química Orgánica (UMYMFOR), Buenos Aires, Argentina.*
¶*CEQUINOR (UNLP—CONICET, CCT La Plata, associated with CIC PBA). Departamento de Química, Facultad de Ciencias Exactas, Universidad Nacional de la Plata. Blvd 120 1465 (1900) La Plata, Argentina.*

E-mail: lavecchia@quimica.unlp.edu.ar

Phone: +54 0221 445-4393

Abstract

The outbreak of COVID-19, the disease caused by SARS-CoV-2, continues to affect millions of people around the world. The absence of a globally distributed effective treatment makes the exploration of new mechanisms of action a key step to address this situation. Stabilization of non-native Protein-Protein Interactions (PPIs) of the nucleocapsid protein of MERS-CoV has been reported as a valid strategy to inhibit viral replication. In this study, the applicability of this unexplored mechanism of action

against SARS-CoV-2 is analyzed. During our research, we were able to find three inducible interfaces of SARS-CoV-2 N protein NTD, compare them to the previously reported MERS-CoV stabilized dimers, and identify those residues that are responsible for their formation. A drug discovery protocol implemented consisting of docking, molecular dynamics and MM-GBSA enabled us to find several compounds that might be able to exploit this mechanism of action. In addition, a common catechin skeleton was found among many of these molecules, which might be useful for further drug design. We consider that our findings could motivate future research in the fields of drug discovery and design towards the exploitation of this previously unexplored mechanism of action against COVID-19.

Contents

1   Introduction

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), the causative agent of COVID-19, is a positive-sense single-stranded RNA virus that belongs to the β genus of the Coronaviridae family. The SARS-CoV-2 virion consists of at least four (4) structural proteins: Spike (S) protein, membrane (M) protein, envelope (E) protein, and nucleocapsid (N) protein.[1] These proteins, along with other non-structural ones, have been identified as important drug targets,[2] and have been widely investigated since the pandemic started. However, despite the efforts of the global community, only few treatments have shown positive outcomes in their clinical trials, with their worldwide accessibility remaining as a challenge.

Among these main targets, N protein is a major structural protein divided into three domains, namely an N-terminal RNA binding domain (NTD), a poorly structured central Ser/Arg (SR)-rich linker, and a C-terminal dimerization domain (CTD).[3] This protein has been recognized as an important drug target against CoVs since it plays several crucial functions in the viral cycle, such as packaging the RNA genome into ribonucleoproteins (RNP), modulating host cell metabolism and regulating viral RNA synthesis during replication and trascription.[4]

Recent studies showed that the stabilization of non-native Protein-Protein Interactions

(PPIs) between these N terminal domains (NTD) in MERS coronavirus is a valid strategy to halt viral replication.[5–7] However, to date no study has evaluated the applicability of exploiting this mechanism of action against COVID-19.

Modulation of PPIs by small molecules is an emerging and versatile strategy in drug development. The direct stabilization of these type of interactions is conceptually challenging since it requires the simultaneous targeting of more than one protein within the complex. Even though there are numerous examples of small molecules (especially natural products) that relay their biological activity on this type of interactions,[8] very few studies have focused on modeling PPIs for drug design.

In the first part of this work, the stabilization of the reported MERS-CoV non-native nucleocapsid homodimers is analyzed using molecular dynamics and end-point MM-GBSA energy calculations, with results in good correlation with the reported experimental data. Then, a proposed protocol led to the identification of a privileged scaffold that may be used for further rational drug design. In addition, two already effective drugs against SARS-CoV-2, Ruxolitinib and Sorivudine, were also found between the potential active compounds, and therefore this alternative mechanism of action might explain part of their activity. We believe that the contributions made in this research might be useful in future drug design, not only for this particular mechanism of action, but also for any study that aims for PPIs stabilization by small molecules.

2   Methods

A general scheme of the protocol is presented in Figure 1.

2.1   MERS-CoV Interface Preparation

The published MERS-CoV homodimers PDB ID 6KL5 and 6KL6 were rebuilt for the initial analysis. Missing residues were added using 4UD1 as template. The isolated system was
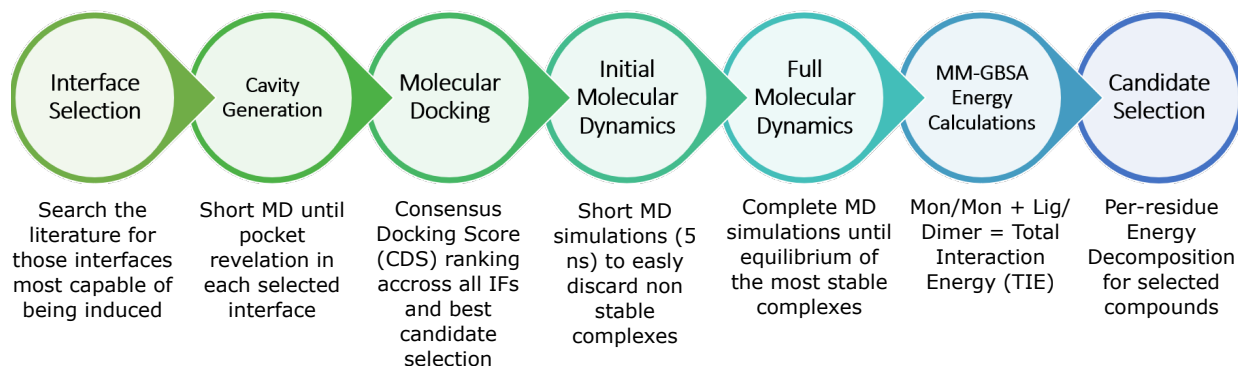
Figure 1: General Scheme of the protocol developed for the present study.

named **IF-MERS-CoV** whereas ligand-bonded dimers were called P1/**IF-MERS-CoV** and P3/**IF-MERS-CoV** (ligands were named as in the Hou's article[5]).

2.2   SARS-CoV-2 Interface Preparation

In order to obtain inducible homodimer interfaces between NTD N protein monomers, a crystal tetramer was selected (PDB ID 6VYO). This structure consists of four monomers stabilized with Zn^{2+} cations. The selected interfaces were covered by the union of chains A/B (**IF-1**) and chains B/C (**IF-2**), which are equivalent to the interface generated by chains C/D and D/A, respectively. Because there are missing residues (UniProt ID P0DTC9), some of them close to the interface, the dimers were rebuilt by superimposing a monomeric experimental structure obtained by solution NMR (PDB ID 6YI3).

Furthermore, an additional interface was generated using the same methodology on 6KL5 dimer structure, an X-ray non-native dimer corresponding to the MERS-CoV N protein.[5] This dimer was called **IF-3** (see Figure 1).

Finally, these homodimers were subjected to a minimization, heating and NPT equilibration process in order to generate small cavities capable of harboring a stabilizing ligand.

2.3   Sequence Alignment and Comparison

MERS-CoV, SARS-CoV and SARS-CoV-2 Nucleocapsid proteins sequences (Uniprot IDs: R9UM87, P59595 and P0DTC9 respectively) were aligned and compared using JalView 2.1.11.0 software.[9]

2.4   Molecular Docking Studies

Compounds from DrugBank 5.1.6 database[10] were screened against the selected dimer interfaces. The protonation state of each ligand at pH 7.4 was based on analysis using FixpKa, included in QUACPAC 2.0.2.2[11] and its conformers were generated using Omega.[12] A docking region in each interface was selected using Make Receptor tool, included in the OEDocking 3.4.0.2 suite, OpenEye.[13] Parameters were kept to their default values. Molecular docking was performed to find and score binding poses by means of FRED software, included also in the OEDocking suite. Parameters were kept to their default value. Based on the results, the Phenol-Explorer database[14] was included to be treated with this same docking protocol.

A Consensus Docking Score (CDS) was calculated with the following equation:

$$\mathrm{CDS}_i = \sum_{n=1}^{3} \frac{S_i(\mathbf{IF\text{-}n})}{\langle S(\mathbf{IF\text{-}n}))\rangle} + \frac{S_i(RNA)}{\langle S(RNA)\rangle}$$

where CDS_i is the consensus docking score of compound i, $S_i()$ and $\langle S()\rangle$ are score of compound i and average score in IF-n (**IF-1**, **IF-2**, **IF-3** interface) and RNA interaction residues docking.

2.5   Molecular Dynamics (MD) Simulations

MD simulations were performed on complexes of the selected ligands and dimers. Since these systems have a positive net charge, chlorides were added as counterions with Leap module to achieve electroneutrality. The neutralized complexes were immersed in a box of TIP3P waters which extended up to 15 Å from the complex. Dimers were described using the Amber14SB force field.[15] Ligand topologies were built using the Generalized Amber Force

Field (GAFF)[16] with charges derived from RESP, which were calculated with Antechamber module and Gaussian03.[17] Leap and Antechamber[18] are included in the package AmberTools 19.0.[19]

All MD simulations were run using the NAMD 2.13 software.[20] The van der Waals interaction cutoff distances were set at 12 Å and long-range electrostatic forces were computed using the particle mesh Ewald summation method with a grid size set to 1.0 Å. The 1-4 contributions were multiplied by a factor of 0.83 to match the AMBER force field requirements. The complexes were submitted to a series of minimization stages (including hydrogens, waters, residue side chains and the complete system) of 100,000 maximum steps each, and heating from 0 to 310K. For equilibration/production simulations, constant temperature (310 K) was maintained using Langevin dynamics with a damping coefficient of 5 ps^{-1}, while pressure was kept constant at 1 atm through the Nosé-Hoover Langevin piston method with a decay period of 200 fs and a damping time constant of 100 fs. A time step of 1 fs was used along molecular mechanics. Bonds involving H atoms of waters were constrained using the SHAKE algorithm. Cationic Dummy Atom method[21] was used when the Zn^{2+} cation is included in the system.

RMSD plots were depicted to determine the convergence and stability of simulations and are available in the Supplementary Information (Tables S3-8).

2.6   Total Interaction Energy Calculations

In order to quantify the dimer stabilization exerted by the ligand, two stabilization components were considered: ligand-dimer and monomer-monomer interactions. The sum of these two components was called Total Interaction Energy (TIE). These energies were calculated using free energy decomposition analysis, which was performed using a pairwise energy decomposition scheme (idecomp option 3) with the MMPBSA.py module.[22]

E(Lig/Dimer) = $\sum_i$ E(Lig/Res$_i$) + $\sum_j$ E(Lig/Res$_j$)

TIE(Lig) = E(Lig/Dimer) + $\sum_{i,j}$ E(Res$_j$/Res$_j$)
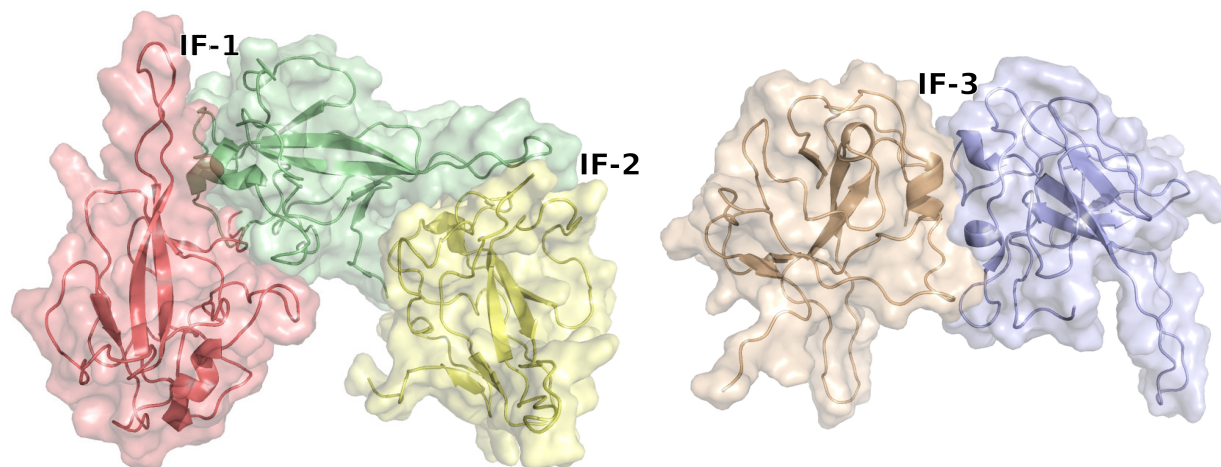
Figure 2: Selected homodimer interfaces to be stabilized: **IF-1** and **IF-2** built from 6VYO, and **IF-3** from 6KL6 structure. Image created using Pymol.

In this scheme, interactions are decomposed by specific residue pairs by including only those interactions in which one atom from each of the analyzed residues is participating, following the work of Gohlke et al.[23] A single trajectory based on 50 snapshots was taken from the last 5 ns portion (100 ps interval) of the MD production trajectories. The entropic change of the complexes was assumed constant to reduce computational cost. The solvation free energy (ΔG_{solv}) was separated into polar and non-polar contributions. The polar contribution to ΔG_{solv} was calculated using the generalized Born (GB) model implemented in MMPBSA.py module: igb=2/mbondi2 radii sets and igb=8/mbondi3 radii sets as selected models for ligand/protein and monomer/monomer interactions, respectively. Further, this decomposition scheme allowed us to assign the percentage of the ligand-dimer interaction to each monomer, and thus ensure that there is a balanced distribution. CPPTRAJ[24] and ParmEd programs, included in AmberTools, were used for trajectory handling.

3   Results

3.1   Interface Selection and Analysis

3.1.1   Analysis of the Reported MERS-CoV Interface

In order to understand, model and compare the behavior of PPIs between the nucleocapsid NTD of SARS-CoV-2, a computational analysis of the previously reported[5] non-native MERS-CoV Nucleocapsid NTD dimers was conducted.

This was accomplished by performing molecular dynamics simulations in the reported ligand-bonded dimers (P1/**IF-MERS-CoV** and P3/**IF-MERS-CoV**) and the isolated interface (**IF-MERS-CoV**). All of them proved to be stable systems, since equilibrium was reached in all cases. TIE calculations correlated with the activities previously reported (P3, $EC50 = 32 \ \mu M \ and \ P1, \ EC50 > 100 \mu M$) and also the P3/**IF-MERS-CoV** complex proved to be more stable than the isolated dimer (see Table 2). Moreover, in the case of P3 ligand, per-residue decomposition analysis showed also excellent correlation between the reported ligand-dimer interacting residues and the ones with the strongest interaction in our simulations. However, this was not accomplished in the case of P1, since during the trajectory the pose of the ligand changes (Figure S4). This can be understood considering the lower number of interactions reported in comparison with P3, which also explains the differences in the EC50 values.

Per-residue decomposition analysis for the ligand-free **IF-MERS-CoV** showed some key interactions between both monomers which contribute to the system stability (Figure 3). A:Gly105/B:Asn68 and A:Gly103/B:Thr70 are the two strongest pair interactions and Gly103, Thr104, Asn138 of chain A and Thr134 of chain B presented significant interactions with more than one residue of the other monomer. It is important to mention that the added highly mobile amino acid sequence introduced some uncertainty to these energy calculations, and their residues were excluded in these analyses.

Sequence comparison between the MERS-CoV nucleocapsid protein and its analog in

SARS-CoV-2 revealed that there is only 48,42% of identity between them (see Figure S1), suggesting that the active ligands in MERS-CoV were unlikely to be effectively repurposed in SARS-CoV-2.

3.1.2   SARS-CoV-2 Ligand-Free Interfaces Analysis

The stabilization of PPIs with small molecules is a challenging task that begins with the accurate selection of the interfaces that might be stabilized. A numerous amount of interfaces can be built with two monomers of the same protein, but not all of them are capable of being induced by small molecules. As mentioned in the methodology section, three interfaces were built relaying on experimental information: **IF-1** and **IF-2**, which were reported in a non-native tetramer crystal structure including a Zn^{2+} cation, and **IF-3**, which was built by homology modelling of the previously analyzed MERS-CoV interface.

Molecular dynamics on the three ligand-free interfaces were performed. These simulations showed that the only stable system was **IF-1**, with a TIE of -91 kcal/mol which is a similar value to the one obtained for the **IF-MERS-CoV** interface. On the other hand, the TIE values of **IF-2** and **IF-3** were weaker than the previous ones, as expected.

Per-residue energy decomposition for these interfaces showed that, when comparing the two stable systems with the unstable ones, the first ones present a higher number of strong interactions (Figure 3). In the case of **IF-1** three important pair interactions can be observed, these are B:Asp101/A:Lys196, B:Asp38/A:Lys263 and B:Thr33 with A:Ser270 and A:Gly272. The contribution to the stability provided by these residues should be considered in future drug design, since their blockage may act against dimer stabilization.

3.1.3   IF-1 and IF-2 and their interaction with Zn^{2+}

With the intention of understanding the role played by Zn^{2+} cation in the crystal structure, it was incorporated to **IF-1** and **IF-2** dimers. Molecular dynamics simulations and TIE calculations showed that this cation is able to stabilize both **IF-1** and **IF-2** dimers (see
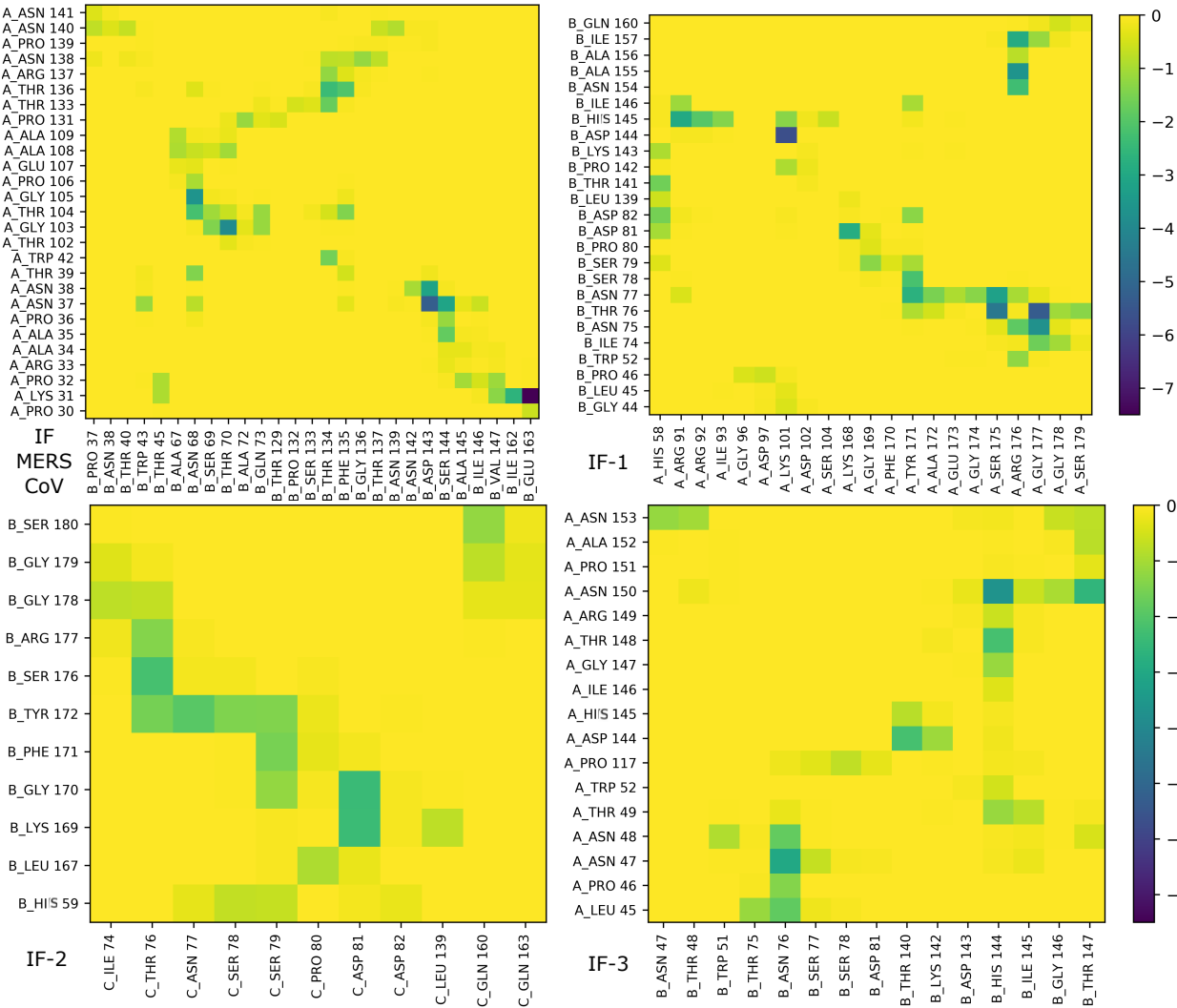
Figure 3: Per-residue decomposition analyses of the ligand-free interfaces.

Table 1), verifying that these two interfaces can be effectively induced. This discovery might contribute to the understanding of Zn^{2+} potential therapeutic action against COVID-19, since it has aroused interest during the pandemic due to its medicinal properties[25] and has been proposed in many clinical trials.

Table 1: Decomposed Total Interaction Energy for the studied interfaces calculated with MM-GBSA method. Ligand/dimer and monomer/monomer interaction energies were modelled using igb=2/mbondi2 radii sets and igb=8/mbondi3 radii sets as selected models, respectively. RMSD plots are available in tables S3-5. Standard deviations are shown between parentheses. n.a.: not applicable

Interface	Ligand/Dimer Interaction (kcal/mol)	Mon/Mon Interaction (kcal/mol)	TIE (kcal/mol)
IF-MERS-CoV	n.a.	-96	-96
P1/**IF-MERS-CoV**	-31 (2)	-32 (4)	-63 (4)
P3/**IF-MERS-CoV**	-41 (2)	-65 (3)	-106 (4)
IF-1	n.a.	-91 (5)	-91 (5)
IF-2	n.a.	-33 (5)	-33 (5)
IF-3	n.a.	-45 (4)	-45 (4)
Zn^{2+}/**IF-1**	-101 (4)	-57 (4)	-158 (6)
Zn^{2+}/**IF-2**	-101 (4)	-41 (6)	-142 (7)

3.2   Repurposed drugs as stabilizers

Once the interfaces were selected and the pockets generated, molecular docking calculations with both investigational and approved DrugBank's databases on all the interfaces were performed. At first, the high mobility of these systems seemed incompatible with the rigid nature of standard docking calculations. To address this situation, the compounds were ranked according to a *consensus* docking score (CDS), constructed from the relative docking position of each compound in each interface. Another important issue was that each system was being considered as equally probable to be induced, which was not necessarily true. Since RNA assembly is the main function of this protein, a good strategy to solve this problem would be to prioritize those ligands capable to interact with the residues involved in this process, which had already been described in literature.[26] With this in mind, a fourth system centered on those residues was added to the previous *consensus* docking score ranking (Table S1, Column "Score RNA and Pos. RNA").

Table 2: Information on the selected compounds. An extended table is found in Supplementary information. RMSD plots are available in tables S6-8. Standard deviations are shown between parentheses. *Consensus* Docking position for compounds of DrugBank / DrugBank + Explorer-Phenol databases. n.i.: not included.

Name/Abbrev.	Type	*Consensus* Docking Position*	TIE **IF-1** (kcal/mol)	TIE **IF-2** (kcal/mol)	TIE **IF-3** (kcal/mol)
Masoprocol	Approved, Polyphenol	17 / 23	Dimer Separation	-76(5)	-185 (6)
Cianidanol	Approved, Catechin	38 / 10	-142 (6)	-84 (6)	-240 (8)
(+)-Gallocatechin Gallate / (+)GCG	Catechin	n.i. / 11	-134 (9)	-166 (8)	-129 (6)
(-)-Gallocatechin Gallate / (-)GCG	Catechin	n.i. / 55	-144 (7)	-133 (8)	Out of binding site
(-)-Epicatechine / (-)EC	Investigational, Catechin	18 / 27	-154 (8)	-150 (6)	-110 (6)
(+)-Gallocatechin / (+)GC	Catechin	n.i. / 8	-139 (7)	-120 (5)	-113 (6)
(-)-Epigallocatechin / (-)EGC	Catechin	3 / 7	-97 (6)	-103 (7)	-185 (6)
Quercetin	Flavonoid	27 / 59	-127 (6)	-88 (5)	Dimer Separation
(-)-Catechin Gallate / (-)CG	Catechin	n.i. / 32	-130 (6)	-162 (8)	-113(5)
(-)-Catechin / (-)C	Investigational, Catechin	n.i / 46	-127 (5)	-100 (7)	-125 (5)
(-)-Epigallocatechin Gallate / (-)EGCG	Investigational, Catechin	22 / 79	-118 (7)	-150 (6)	-116 (5)
Ruxolitinib	Approved	26 / 18	-88 (5)	-114 (5)	Out of binding site
Sorivudine	Approved	219 / 289	-95 (7)	-78 (5)	-89 (6)

Rational manual selection of the best candidates was made. For this task, the best approved ligands were given priority over the investigational ones, and those compounds that had a top score in any of the interfaces were also considered, even if their *consensus* ranking was not the best. With this strategy, 13 initial compounds were selected (Table S1) for molecular dynamics.

Stabilization was accomplished if ligand, interface-forming residues and protein backbone RMSD reached a plateau, and also if the compound showed balanced interactions with residues from both monomers, characterized with MM-GBSA energy decomposition.

These simulations showed that several compounds were able to effectively stabilize one or more of the studied interfaces (see Table S9, and Figure S3). Moreover, we observed that those compounds selected because of their higher CDS behaved better as stabilizers rather than the ones selected because of their good performance in only one of the interfaces. However, we noticed that positively charged ligands were not able to stabilize any of the interfaces, which is in accordance with the highly positive charge of both monomers reported in the literature.[26]

The lack of experimental data makes it very difficult to establish a threshold and select the most promising compounds from this list. Therefore, we believe that it is very well justified to use the values observed in P3/**IF-MERS-CoV** as reference, even though they belong to a different system.

From this list, Ruxolitinib is already being tested in several clinical trials against COVID-19, and some results proving its effectiveness have been published.[27,28] In our simulations, this molecule was able to stabilize **IF-1** and **IF-2**, with a TIE of -81 and -114 kcal/mol respectively. As can be observed in Table 2, this last energy value is 12 kcal/mol more stable compared to P3/**IF-MERS-CoV**, the active ligand in MERS-CoV. Although it is believed that this molecule's effectiveness relays on JAK inhibition, our results suggest that part of its potency may be explained by this other mechanism of action. Another molecule included in this list that is also in clinical trials against COVID-19 is Quercetin, with results still

pending.

Meanwhile, during the course of our research several approved drugs where tested *in vitro* against SARS-CoV-2, and 100 compounds were identified as inhibitors of viral replication.[29] Comparing this list with our docking results, we noticed that Sorivudine had one of the top 20 lowest docking scores in **IF-1**, and therefore it was considered to run molecular dynamics simulations. These calculations showed that this compound is able to stabilize all three interfaces, but with a lower TIE than the one obtained for P3/**IF-MERS-CoV** in all cases (Table 1). These results suggest that even though this mechanism of action is unlikely to be the main reason for Sorivudine's effectiveness, its potency may also be affected by these type of stabilization.

3.3   Catechins as Promising Candidates

Our previous docking results showed that most of the molecules with the lowest score were polyphenols, in particular flavonoids (Table S1). Moreover, molecular dynamics simulations and energy calculations showed that all of these molecules were able to stabilize at least two interfaces. This is particularly interesting since polyphenolic compounds are phytochemicals widely spread throughout plants and fruits, and are known to have medicinal and chemopreventive activities, for example they are good natural antioxidants. Moreover, the flavonoid skeleton has been reported as a privileged substructure for PPIs modulation.[30] Interestingly, a previous study has reported that some polyphenols inhibit nucleocapsid phosphoprotein's function in SARS-CoV,[31] which has a 90.52% of identity with SARS-CoV-2 analog (Figure S2). In this work, Roh compares the binding affinity of an engineered RNA oligonucleotide with several polyphenols that block its interaction with the protein, and therefore inhibit its functionality. Even though many of these molecules are considered PAINS due to their catechol substructure and are usually excluded in virtual screening, we believe that these mentioned literature together with the high availability of these compounds in nature makes them relevant to further study their potential applications in the present context. Conse-
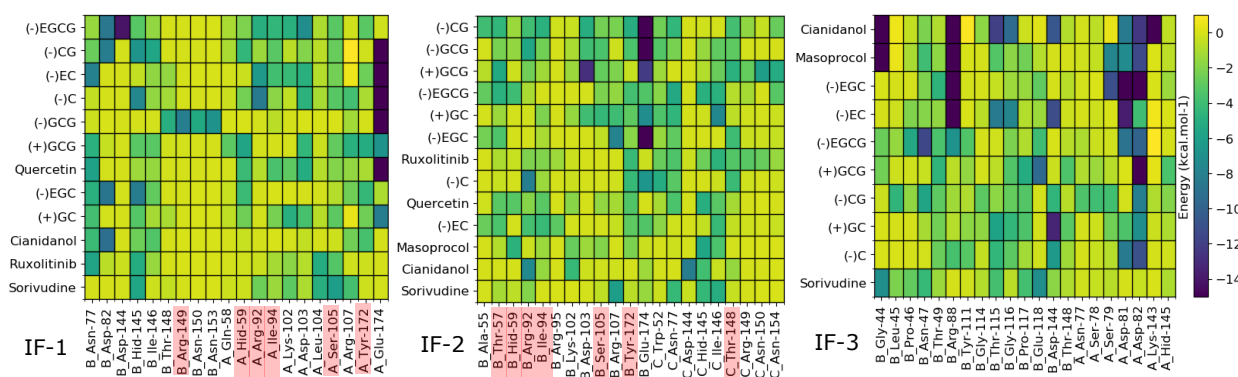
Figure 4: Per-residue decomposition analyses for selected compound of ligand-dimer energy component. The rows are ordered according to the total ligand-dimer contribution, decreasing downwards. In addition, the residues that interact with RNA, according to Kang et al.,[26] are highlighted. Compounds full name are showed in Table 2.

quently, a polyphenol-specific database was added to further study this possibility.

Molecular docking calculations of these added compounds showed that, among the 378 polyphenolic molecules in the new database, flavonoids had the lowest *consensus* docking scores. Moreover, a good differentiation between the compounds tested as active and inactive in Roh's work was found,[31] particularly considering the **IF-2** docking score (Table S2). These results motivated the addition of (+)-Gallocatechin Gallate, (+)-Gallocatechin, (-)-Catechin Gallate for further molecular dynamics simulations. Also, (-)-Catechin and (-)-Gallocatechin Gallate were taken into account for comparison, since they had been tested in Roh's work. Finally, despite not being a catechin itself, Gallic acid 3-O-gallate was also considered since it is a synthetic precursor of many of the molecules selected, and it had the second best *consensus* docking score of all.

Molecular dynamics simulations were in line with our previous calculations, since all catechins tested stabilized at least two interfaces (see Table S9). However, this was not the case for Gallic acid 3-O-gallate, which left the binding site in **IF-2** and **IF-3**. This result highlights the importance of the common catechin heterocyclic motif for this type of stabilization.

3.4 Total Interaction Energy Calculations and Per-Residue Decomposition Analysis

The final step in this protocol consists of TIE calculations from MM-GBSA results, shown in Table 3.

Correlation was found between previously known active/inactive compounds in Roh's[31] work and the calculated interaction energies in **IF-1** and **IF-2**. This was not the case for **IF-3**, since this interface does not include the residues that are known to interact with RNA, which were specifically tested in the cited article. Another important aspect to notice in this comparison is that dimer stabilization was not considered in the performed biological assay, which focuses only in the inhibition of the protein function.

Some stable complexes where the monomer/monomer interaction term was significantly stronger when compared with the isolated dimer were observed in the case of **IF-2**. Per-residue decomposition showed that many new interaction pairs were induced when a ligand was present (Figure S5), contributing for the complex stability. Even though few patterns can be observed between the ligands, most interactions involve Arg177 (chain B) and Asn76 (chain C) or residues close to them. These results suggest that the ligand's ability to induce strong interactions between the interface-forming residues is an important aspect to take into account when aiming for PPIs stabilization.

Energy decomposition is presented in Figure 4. In these graphics, the most important ligand/residue interactions for each interface can be easily observed. Ligands with the strongest ligand/dimer interactions can be visualized in Figures S6-13. A possible strategy to analyze these results in **IF-1** and **IF-2** is to focus on those residues that are known to interact with RNA.[26]

Focusing on these residues in **IF-1**, it can be seen that His59, Arg92, Ile94, Ser105 and Tyr172 from chain A, and Arg149 from chain B, are part of this binding site (Figure 4). Even though few patterns can be observed, (-)EGCG, (-)CG, (-)EC and (-)C showed strong interactions with many of these key residues, whereas the other ligands did not. Finally,

most compounds showed a remarkable interaction with Glu174 (chain B), which proved to be an important residue for this dimer stabilization.

On the other hand, the number of RNA-binding residues in **IF-2** is slightly higher. In its pocket, Thr57, His59, Arg92, Ile94, Ser105, Arg107, Tyr172 from chain B showed significant interaction with many of these compounds, especially with those that had a lower TIE value. Moreover, when comparing these results with the activities measured in Roh's work, it can be observed that the active compounds (-)GCG and (-)CG interact significantly better with these residues in comparison with the inactive molecules, (-)C and Quercetin. This result could explain the differences between their reported biological activities. Finally, similar to **IF-1**, Glu174 shows a strong interaction with those ligands with better performance, highlighting its crucial role in the stabilization of both complexes.

Lastly, **IF-3** proved to be a completely different interface than the previous two. No RNA-binding residues are found in its pocket, but others turned out to be important for its stabilization. These are: Gly44, Arg88, Asp144 from chain A and Asp81, Asp82 from chain B. In addition, some of the best stabilizing ligands in this system, such as Cianidanol and Masoprocol, showed very weak ligand/dimer interaction in the previous two interfaces, which can be explained due to the differences between the interface-forming residues.

4   Conclusions

A computational analysis of the stabilization of non-native PPIs of SARS-CoV-2 N protein as a potential mechanism of action against COVID-19 is provided. A first comparison with MERS-CoV enabled us to explore its applicability towards SARS-CoV-2, where three inducible interfaces were built based on experimental data. Consequently, a drug discovery protocol was developed and applied with a drug repurposing approach, which allowed us to identify the main residues that build each interface. In addition, potential candidates that might take advantage of this mechanism were found, many of them with a common

catechin skeleton that might be useful for further drug design. Even though further *in vitro* testing is needed to confirm these findings, we believe they provide valuable insights into this mechanism of action that may promote future research in the fields of drug discovery and medicinal chemistry.

Acknowledgement

JFF thanks the University of Buenos Aires for a doctoral fellowship. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research and OpenEye for providing an academic license to use their suite. Also, we thank Dr. J. A. Palermo and Dra. V. Ferraresi Curotto for their invaluable help during this research.

This work was supported by CONICET and UNLP. MJL is member of the Carrera del Investigador CONICET.

Supporting Information Available

- supplementary_material.pdf: Additional tables and figures that were referenced in the main text.

- IF_structures.zip: Contain .pdb files of the protein dimer structures used in MD with AMBER format.

- selected_compounds_smiles.csv: ligand smiles.

References

(1) Li, X.; Geng, M.; Peng, Y.; Meng, L.; Lu, S. Molecular immune pathogenesis and diagnosis of COVID-19. *Journal of Pharmaceutical Analysis* **2020**,

(2) Dong, S.; Sun, J.; Mao, Z.; Wang, L.; Lu, Y.-L.; Li, J. A guideline for homology modeling of the proteins from newly discovered betacoronavirus, 2019 novel coronavirus (2019-nCoV). *Journal of medical virology* **2020**,

(3) Chang, C.-k.; Chen, C.-M. M.; Chiang, M.-h.; Hsu, Y.-l.; Huang, T.-h. Transient oligomerization of the SARS-CoV N protein–implication for virus ribonucleoprotein packaging. *PloS one* **2013**, *8*, e65045.

(4) Chang, C.-k.; Hou, M.-H.; Chang, C.-F.; Hsiao, C.-D.; Huang, T.-h. The SARS coronavirus nucleocapsid protein–forms and functions. *Antiviral research* **2014**, *103*, 39–50.

(5) Lin, S.-M.; Lin, S.-C.; Hsu, J.-N.; Chang, C.-k.; Chien, C.-M.; Wang, Y.-S.; Wu, H.-Y.; Jeng, U.-S.; Kehn-Hall, K.; Hou, M.-H. Structure-based stabilization of non-native protein–protein interactions of coronavirus nucleocapsid proteins in antiviral drug design. *Journal of medicinal chemistry* **2020**, *63*, 3131–3141.

(6) Amin, S. A.; Jha, T. Fight against novel coronavirus: A perspective of medicinal chemists. *European Journal of Medicinal Chemistry* **2020**, 112559.

(7) Ghosh, A. K.; Brindisi, M.; Shahabi, D.; Chapman, M. E.; Mesecar, A. D. Drug Development and Medicinal Chemistry Efforts toward SARS-Coronavirus and Covid-19 Therapeutics. *ChemMedChem* **2020**,

(8) Andrei, S. A.; Sijbesma, E.; Hann, M.; Davis, J.; O'Mahony, G.; Perry, M. W.; Karawajczyk, A.; Eickhoff, J.; Brunsveld, L.; Doveston, R. G., et al. Stabilization of protein-protein interactions in drug discovery. *Expert Opinion on Drug Discovery* **2017**, *12*, 925–940.

(9) Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191.

(10) Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research* **2006**, *34*, D668–D672.

(11) OpenEye Scientific Software, Santa Fe, NM., QUACPAC. http://www.eyesopen.com.

(12) Hawkins, P. C.; Nicholls, A. Conformer generation with OMEGA: learning from the data set and the analysis of failures. *Journal of chemical information and modeling* **2012**, *52*, 2919–2936.

(13) McGann, M. FRED and HYBRID docking performance on standardized datasets. *J Comput Aid Mol Des* **2012**, *26*, 897–906.

(14) Neveu, V.; Perez-Jiménez, J.; Vos, F.; Crespy, V.; du Chaffaut, L.; Mennen, L.; Knox, C.; Eisner, R.; Cruz, J.; Wishart, D., et al. Phenol-Explorer: an online comprehensive database on polyphenol contents in foods. *Database* **2010**, *2010*.

(15) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation* **2015**, *11*, 3696–3713.

(16) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *Journal of computational chemistry* **2004**, *25*, 1157–1174.

(17) Frisch, M. J. et al. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004.

(18) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond

type perception in molecular mechanical calculations. *Journal of molecular graphics and modelling* **2006**, *25*, 247–260.

(19) Case, D., et al. AmberTools19. 2019; `http://trolltech.com/products/qt/`.

(20) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *Journal of computational chemistry* **2005**, *26*, 1781–1802.

(21) Pang, Y.-P. Successful molecular dynamics simulation of two zinc complexes bridged by a hydroxide in phosphotriesterase using the cationic dummy atom method. *Proteins: Structure, Function, and Bioinformatics* **2001**, *45*, 183–189.

(22) Miller III, B. R.; McGee Jr, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: an efficient program for end-state free energy calculations. *Journal of chemical theory and computation* **2012**, *8*, 3314–3321.

(23) Gohlke, H.; Kiel, C.; Case, D. A. Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the Ras–Raf and Ras–RalGDS complexes. *Journal of molecular biology* **2003**, *330*, 891–913.

(24) Roe, D. R.; Cheatham III, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *Journal of chemical theory and computation* **2013**, *9*, 3084–3095.

(25) Derwand, R.; Scholz, M. Does zinc supplementation enhance the clinical efficacy of chloroquine/hydroxychloroquine to win todays battle against COVID-19? *Medical Hypotheses* **2020**, 109815.

(26) Kang, S.; Yang, M.; Hong, Z.; Zhang, L.; Huang, Z.; Chen, X.; He, S.; Zhou, Z.; Zhou, Z.; Chen, Q., et al. Crystal structure of SARS-CoV-2 nucleocapsid protein

RNA binding domain reveals potential unique drug targeting sites. *Acta Pharmaceutica Sinica B* **2020**,

(27) Koschmieder, S.; Jost, E.; Cornelissen, C.; Müller, T.; Schulze-Hagen, M.; Bickenbach, J.; Marx, G.; Kleines, M.; Marx, N.; Brümmendorf, T. H., et al. Favorable COVID-19 course despite significant comorbidities in a ruxolitinib-treated patient with primary myelofibrosis. *European journal of haematology* **2020**,

(28) La Rosée, F.; Bremer, H.; Gehrke, I.; Kehr, A.; Hochhaus, A.; Birndt, S.; Fellhauer, M.; Henkes, M.; Kumle, B.; Russo, S., et al. The Janus kinase 1/2 inhibitor ruxolitinib in COVID-19 with severe systemic hyperinflammation. *Leukemia* **2020**, 1–11.

(29) Riva, L.; Yuan, S.; Yin, X.; Martin-Sancho, L.; Matsunaga, N.; Pache, L.; Burgstaller-Muehlbacher, S.; De Jesus, P. D.; Teriete, P.; Hull, M. V., et al. Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature* **2020**, 1–11.

(30) Bosc, N.; Kuenemann, M. A.; Becot, J.; Vavrusa, M.; Cerdan, A. H.; Sperandio, O. Privileged Substructures to Modulate Protein–Protein Interactions. *Journal of chemical information and modeling* **2017**, *57*, 2448–2462.

(31) Roh, C. A facile inhibitor screening of SARS coronavirus N protein using nanoparticle-based RNA oligonucleotide. *Int J Nanomedicine* **2012**, *7*, 2173.

Graphical TOC Entry