Co-evolved ligands to ORF8. Could they reduce SARS-COV-2-excesive inflammation?

short title: ORF8 co-evolutionary docking
Bello-Perez, M+. and Coll, J*.
+Unidad Enfermedades Infecciosas. Hospital Universitario. Elche, Spain
*Department of Biotechnology. Centro Nacional INIA-CSIC. Madrid, Spain.
+Melissa Bello-Perez, orcid: 0000-0002-9212-083X
Email: melissa.bello@alu.umh.es (MB)
*Julio Coll, orcid: 0000-0001-8496-3493
Email: juliocollm@gmail.com (JC)
* Corresponding author

Abstract

ORF8 is an asymmetric-homodimer SARS-COV-2 accessory protein implicated in excesive human inflammation causing numerous deaths. There is no approved drug targeting ORF8, nor it is known whether any anti-ORF8 drugs could reduce human excesive inflammation. Computationally combining ligand co-evolution of parent molecules with affinity-consensus docking, children candidates for docking to ORF8 cavities were generated. Targeting the homodimer interface with the highest affinity children scaffolds, hundreds of grandchildren predicting nanoMolar affinities, unique scaffolds, high specificities and low toxicity risks were generated. Although remaining hypothetical without experimental confirmation, this constitute a new methodological attempt to search for drug-like candidates to interfere with SARS-COV-2-dependent excessive inflammation.

Keywords: co-evolutionary docking; consensus docking; ORF8, SARS-CoV-2

Introduction

The ORF8 is one of the six viral accessory proteins of SARS-CoV-2. Dispensable for viral replication, those accessory proteins interact with the immune response, most of them playing a crucial role in the pathogenesis of human coronaviruses. ORF8 is unique among other beta-coronavirus, as the sequence identity between SARS-CoV and SARS-CoV-2 is low (~ 26 %)[1], probably to best adapt to human hosts [2]. During the last 2-years, studies focused on SARS-CoV-2 showed an increasing trend implicating ORF8 in many aspects of human immune evasion [3, 4] and other enigmatic functions [5]. Recently, SARS-CoV-2 ORF8 has been describe as a virulence factor[5], as the infection of mice with ORF8-deleted virus and people infected with natural variants lacking ORF8 showed a milder disease, due to the reduction of excesive levels of host pro-inflammatory cytokines. Additionally, ORF8 blood levels correlated with mortality in infected patients [6-8].

ORF8 shows a high number of variations [9-12] and its largest interactome suggested it is implicated in many different functions[13-15]. For instance, some of the effects known to be modulated by ORF8 include host endoplasmic stress [16-18], histone modification [19], SARS-CoV-2 spike S expression [20-22], inhibition of complement [23], inhibition of human interferon responses [24-27], binding to dendritic cells producer of inflammatory cytokines [28], activation of the IL17 pathway [29, 30], inhibition of antibody-dependent cellular cytotoxicity [31], activation of pro-inflammatory storm-causing cytokines trough NLPR3 [32] and dimer-dependent downregulation of MHC1 in human cytotoxic lymphocytes to favour viral immune evasion[33-35].

ORF8 is an asymmetric covalent AB homodimer of 121 amino acid monomers (15 first residues of signal peptide per monomer). The core of each monomer consists of two antiparallel β-sheets [36]. Crystallographically, two different possible dimer interfaces have been proposed, the demonstrated interface stabilized by a $^{20}C-^{20}C$ disulphide covalent bond and the hypothetical non-covalently stabilized by $^{73}YIDI$-motif hydrophobic interactions[37-39].

ORF8 monomers, contain an Ig-like fold stabilized by intramolecular disulphides ($^{25}C-^{90}C$, $^{37}C-^{102}C$)[37] and a highly-dynamic fold loop (hypervariable loop) flanked by its $^{61}C-^{83}C$ disulphide and coding for a $^{73}YIDI$ motif including a glycosylation site (^{78}N). The hypervariable loop has been implicated in the many host protein interactions mentioned above, capable of adapting its hypervariable fold to any of those different ligands [38]. However, to our knowledge, there are no crystallographic ORF8-protein complex structures reported to date.

The ORF8 homodimer disulphide highly interdigitated interface is made up of complementary A-B amino acid surfaces implicating salt bridges (^{119}D-^{115}R and ^{115}R-^{92}E), hydrogen bonds (^{120}F-^{53}K, ^{53}K-^{24}S, $18Q$-^{22}L, ^{52}R-^{121}I) and many other hydrophobic interactions [36]. The predominant natural mutants of ORF8 include the L84S mutation, but S24L and V62L mutations are also abundant [39, 40]. The L84S mutation reduced some of the dimer interacting amino acids, including those implicated in salt bridges (i.e., ^{119}D, ^{115}R), hydrogen bonds (^{120}F, ^{53}K, ^{52}R, ^{121}I), and other interface interactions (^{98}R, ^{104}I, ^{117}V)[41].

ORF8 is secreted into the blood of infected humans both in N-linked glycosylated (^{78}N) [42] and unglycosylated [43] forms, inducing the highest immunogenicity among the SARS-CoV-2 proteins, specially immunodominant is its amino-terminal α-helix[44]. The ORF8 accessory protein has been proposed to modulate the recognition of viral antigens via antigen presenting monocytes, showing a stronger interaction with CD14+ monocytes using their NLPR3 receptor.

The unglycosylated ORF8 binding to CD14+ monocytes causes a dysregulation of the inflammatory response characterized by elevated blood concentrations of interleukin-6, granulocyte-macrophage colony stimulating factor, and α-tumour necrosis factor among other cytokines (cytokine storm). The higher CD14+ binding ^{84}L variant and the milder disease outcome produced by the weaker CD14+ binding ^{84}S variant, suggested that the ratio of ^{84}L and ^{84}S may control different degrees of inflammation by modulating monocyte recognition of viral antigens depending on the host immune reactions[45].

On the other hand, it was described that ORF8 modulates the recognition of viral antigen via downregulation of MHC1. Because the down regulation of MHC1 was shown to be dimer-dependent[33-35], disruption of the dimer interface perhaps could be employed to neutralize some ORF8 effects, as recently suggested[35]. Therefore, in this work, we included the exploration of the homodimer interface of SARS-CoV-2 ORF8 as a potential target for specific antiviral ligands as a first step to develop possible therapeutic targets, rather than targeting other of the possible ORF8-dependent human factors which may be essential for some important physiological functions[32]. In contrast to previous computational work which only targeted the $^{20}C-^{20}C$ disulphide [35], an alternative strategy including blind-docking of the whole ORF8 molecule to explore any other cavity offered by the homodimer was favoured in this work. Therefore, the best docking-cavities described here were automatically selected by the programs used, minimizing any possible a priori selection.

The present work explored the co-evolution of ORF8 / ligand pair starting from an artificial pseudoligand parent, because of the affinity limitations of the previously proposed computationally defined ligands[35]. Trial and error attempts resulted in the manual design of a polyCarbon pseudoligand centered at the ORF8 molecule widely extending in 4 directions. The pesudoligand was drawn and manually docked in PyMol to be used as initial parent to start co-evolutions (more details in methods). The **D**ata**W**arrior **B**uild **E**volutionary **L**ibrary (DW-BEL) algorithms[2-5] were employed here to successfully generate pseudoligand-derived children fitting the ORF8 interface cavities. The best affinity children-scaffolds targeting the interface cavities were then selected for further co-evolutionary refinement, seeking to increase their children affinities and/or additional cavities. DW-BEL co-evolutionary dockings of those new children scaffolds as parents were capable of generating large numbers of highly specific grandchildren. The high computer memories required to perform such co-evolutionary searches was higher than in our previously reported work[2-5] indicating that targeting the ORF8 interface cavities had enormous steric difficulties, most probably due to the interdigitating amino acids between its monomer faces. The generation of children molecules during co-evolution were controlled for molecular weight, hydrophobicity [6, 7, 8-11] and toxicity risks to avoid unspecificities and/or toxic molecules[46]. The most accurate **A**uto**D**ock**V**ina (ADV) program quantitatively estimated their affinities, and explore for further docking cavity possibilities in the whole homodimer[47].

Because of the absence of experimental data, the predicted docking-cavities and their new ligand candidates remain hypothetical. Some of the predicted results were included as **Supplementary Material** including tables with sliders. Further co-evolutions could be applied by increasing computer memories in the future because the vast chemical space limits have not been yet reached.

Computational Methods

Starting *DataWarrior "Build Evolutionary Library"* with a home-designed parent molecular pseudoligand

The *DataWarrior (DW)* updated program was downloaded (https://openmolecules.org/datawarrior/download.html) following the details for Windows as described before [46, 48]. To explore the initial cavity an star-like carbon molecule of C89H180 formula, was drawn in the ICM Molecular editor and saved as a 2D *.sdf file. It will be called here "erizo". After trial-and-error attempts, the final erizo design contained a central carbon branched by 4 arms of 7 carbons per arm, having all carbons saturated with methyls (molecular weight 1249, logP 28.2). After drawing in MolSoft, the erizo.sdf molecule was manually docked to the ORF8 center in PyMol and saved as a *.pdb file to start *DW-BEL* co-evolutions (**Figure 1 up, red spheres** and **Figure 2 up-left**). The *DW-BEL* fits the children with flexibility around the erizo-supplied cavity, dynamically adapting its cavity to the best fit of the evolving children ligands, as visualized in their *.dwar files.

The fitness criteria preferences and their weight values for *DW Build Evolutionary Library* (*DW-BEL*) for evolutionary docking (co-evolution criteria), were: minimal *DW* docking-scores (weight 4), molecular weights <= 600 g/mol (2), cLog <=4 (1) and Toxicity risk <=1 (4). The *DW* Docking-scores used the mmff94s+ force-field algorithm[49] and specially saved as *sdf files with selected options to best preserve the 2D geometry of the docked children molecules (**Table S1**). The number of runs were limited when reaching 100 Gb of computer memory.

The raw and fitted children data were saved as *.dwar files for storage of the complete co-evolution including docking-scores, chemical properties (molecular weights and LogP hydrophobicities), fitness values and cavity-children images. Before their use, the fitted children data were further filtered using a macro to exclude any remaining toxicities and/or nasty function fragment-containing children molecules (filtering for hundreds of mutagenesis, tumorigenicity, reproductive interference, irritant, and/or nasty functions) (**Supplementary Material / NTNV.dwam** and **/ Nasty_functions.dwar**). The fitted children Special SD-files (*.sdf) were also saved as described in detail at **Table S1**. This saving method supplied *.sdf files maintaining the 3D protein cavity docked to children 3D conformers for visualization in PyMol (using its split_states command)[48] and/or maximal preservation of their 2D geometries for optimal consensus docking.

AutoDockVina docking program

The **A**uto**D**ock**V**ina (ADV) program written in Python vs3.8 included into a modified PyRx-098/PyRx-0.98/1.0 package[46] was used as described before. Some modifications were included here to employ a wide grid to fine-tune docking cavities, to compare with *DW-BEL* docking-scores, quantify affinities in approximated nM and generate protein / ligand 3D images[48, 50-52] with the recently recommended parameters for highest accuracy[53] and optimal 52.3 % success rate compared with other docking programs [47]

Briefly, *.pdbqt file conversion of ORF8 and ligands[54] were made in the mmff94s (Merck) force-field. Ligands for ADV were supplied only as DW-BEL generated children in *sdf files, carefully saved as mentioned above for maximal preservation of most of their 2D geometries after docking. ADV generates many 3D conformers using the rotatable bonds of the input ligands. To avoid their abundant interpretation errors, output ligands should be checked for conservation of the 2D structure of the conformer predicting the lowest docking-score ADV docking-scores in Kcal/mol[55],[51, 56, 57] were converted to nM affinities by the formula, $10^9 \cdot (\exp^{(Kcal/mol/0.592)})$. A grid surrounding the whole ORF8 homodimer molecule of 45x45x45 Å was automatically centered around the PyMol / centerofmass. This grid size explored any other possible docking-cavities for each children, rather than those more limited by *DW-BEL* co-evolution.

Computational manipulations

Computational manipulation software and hardware were similar to those described in detail before. References are reproduced here for convenience:

Table 1
Software and hardware used here for computational manipulations

name	version	Main use	url
DataWarrior	Updated 5.5.0 Windows/Linus	Evolutionary docking[34] Commercial ChemSpace	https://openmolecules.org/ datawarrior/download.html)
Babel & AutoDockVina	Home-adapted PyRx 098/1.0	Force-field minimization & 2D conservation	https://pyrx.sourceforge.io/
MolSoft	3.9 Win64bit	Manipulation of sdf files	https://www.molsoft.com/download.html
PyMol	2.5.7.	Visualization of molecules	https://www.pymol.org/
Discovery Studio	21.1.1.0.20298	Visualization of molecules	https://discover.3ds.com/discovery-studio-visualizer-download
OriginPro	2022	Calculations and Figures	https://www.originlab.com/
LigPlot+	2.2.8.	Amino acid bonds of docked ligands	https://www.ebi.ac.uk/thornton--rv/software/LigPlus/applicence.html
AMD Ryzen i9 computer	4 DDR4 x 32 Gb memory	47 CPU Computational hardware	https://www.pcspecialist.es/

Results

In the absence of drugs approved for ORF8, most probable its best cavities to target could be located at its AB homodimer interface, as already suggested before by targeting the [20]C disulfide bond[35]. Alternatively, here we chose a 45x45x45 Å wide grid surrounding the whole ORF8 homodimer molecule (blind-docking) to search for any other alternative cavities, leaving the docking programs the selection of the best cavities. To start such search, we employed DW-BEL co-evolution because it usually generated tens of thousands of raw-children from any parent molecule to select thousands of children fitting the supplied cavity (fitted-children). However, to start DW-BEL co-evolutions, not only one parent ligand was required but also a protein-ligand complex was required to define a flexible target cavity. After preliminary experiments demonstrated the feasibility of using pseudoligands for both of the DW-BEL requirements mentioned above, the so-called erizo pseudoligand was manually designed, optimized by several co-evolution iterations, manually drawn in MolSoft and coded into an erizo.sdf file. The optimized erizo pseudoligand was an star-like molecule with a central carbon branched by 4 arms of 7 carbons per arm, all carbons were saturated with methyls to maximize its cavity (**Figure 2 up-left**). The erizo.sdf file was merged with 7jtl.pdb at PyMol and manually moved to the center of the homodimer interface to "create" an erizo's cavity (**Figure 1, red spheres**).

Wide enough to include the previously proposed disulfide target[35], the erizo and the erizo's targeted cavity were uploaded to DW-BEL. Co-evolutions rapidly began trimming, changing, adding and selecting the erizo's carbon atoms to generate thousands of children best fitting to any docking cavity located nearby those defined by the erizo (**Figure 2, blue stars**). Using this strategy, and thanks to the adaptability of the cavity during co-evolutions, two docking cavities were identified at the ORF8 asymmetric homodimer. The example described here in more detail, randomly generated 46822 raw-children (erizo's children) of which only 1409 were fitted-children (**Figure 2, blue stars**). One of the docking cavities was located at the AB interface (53 % of the fitted-children) while the other was located at its hypervariable A loop (**Figure 1, down**). Because the hypervariable A loop has been implicated in other protein-protein interactions, the AB ORF8 interface cavity was chosen for this first study.
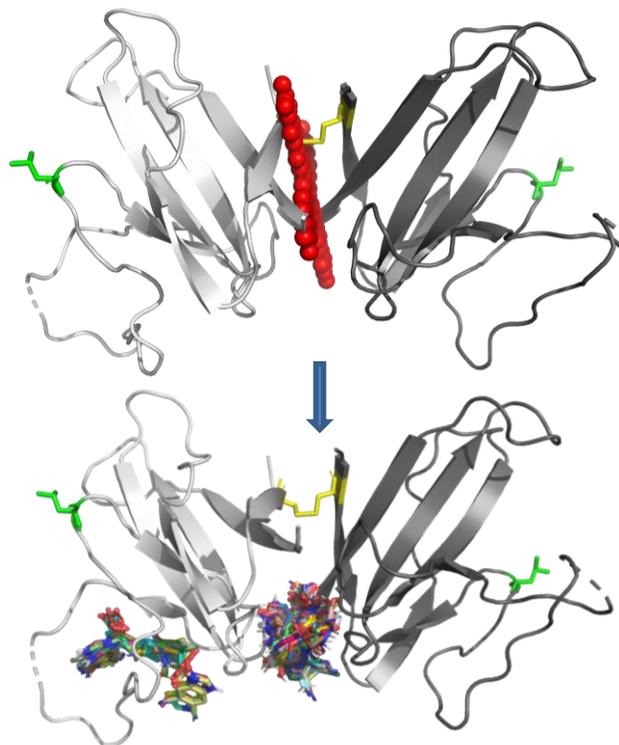
.

Figure 1
Erizo (up) and erizo's top-children (down)
The DW-BEL co-evolution was made from erizo's parent and erizo's cavity at ORF8 (7jtl.pdb) to generate erizo's children. The ADV docking of erizo's children was performed by blind-docking around the ORF8 center and surrounding the whole ORF8 molecule. The ORF8 molecule and top-children were drawn at PyMol.
Light gray cartoon, chain A of the ORF8 homodimer. **Dark gray cartoon,** chain B of the ORF8 homodimer.
Yellow sticks, [20]C-[20]C disulfide Cysteins between the A and B chains
Up red spheres, user-designed and manually docked "erizo" pseudoligand to start *DW-BEL* co-evolutions
Green sticks, amino acid positions of ORF8 variants L[84]S
Multicolor thin sticks, mapping of erizo's derived 100 top-children docking to:
cavity A (left hypervariable loop at monomer A) and cavity AB (interface of the AB homodimer).
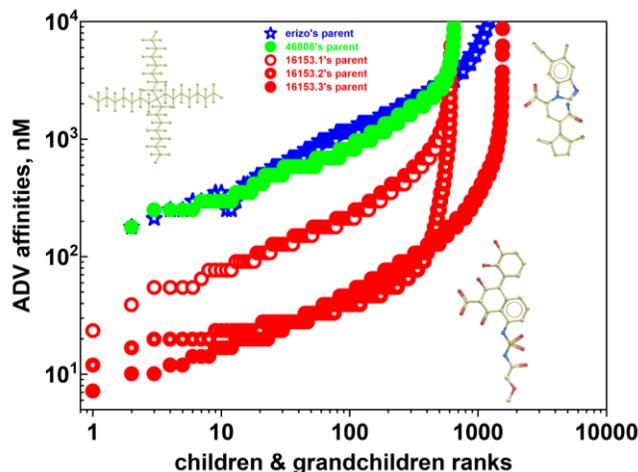
Figure 2
ADV affinity ranks of DW-BEL co-evolutions targeting the ORF8 AB cavity
2D molecular geometries drawn in MolSoft of the erizo (up-left) and the erizo's children 16153 and 46808 (right, down and up respectively), were represented.
Blue stars, erizo's children (3 runs generated 1409 fitted-children)
Green closed circles, grandchildren from 46808 (1 run generated 652 fitted-grandchildren)
Red open circles, 1-run from 16153 generated 602 fitted-grandchildren (16153.1).
Red half-closed circles, 2-runs from 16153 generated 616 fitted-grandchildren (16153.2)
Red closed circles, 3-runs from 16153 generated 1552 fitted-grandchildren (16153.3)

The erizo's children targeting the AB interface cavity predicted maximal affinities ~ 100 nM (**Figure 2, blue stars**) and two main scaffolds. To explore for higher affinities, two top-children were selected (16153 and 46808), each representing one different scaffold (**Figure 2, right** 2D molecular geometries and **Table S2, children**).

The corresponding DW-BEL-ADV grandchildren from 16153 and 46808 predicted improved affinities only when generated from 16153 (Compare **Figure 2, red circles** with **Figure 2, green circles,** and **Table S2**, 16153 and 46808-grandchildren). To test whether grandchildren affinities could be further incremented (lower docking-scores), additional co-evolutions were performed from the 16153 parent using increased number of consecutive runs. Because during consecutive runs, memories are kept for each of the raw- and fitted-grandchildren, their affinities (lower docking-scores) are expected to increase for each additional run. The maximal number of runs was limited to 3 by the computer memory actually available (120 Gb). The docking-score profiles monitored by plotting them *versus* the fitted-grandchildren ID number predicted that the co-evolution of 16153 for 3 consecutive runs maximized their final affinities (**Figure S1, red circles**). These improved DW-BEL affinity results were confirmed by ADV docking (**Figure 2, red circles**). However, because two different scaffolds, one each for 2- and 3-run profiles, were generated, one representative top-grandchild from each of the profiles were chosen for additional studies.

Most of the 2-run grandchildren predicted a unique 4 ring scaffold docking to the same AB cavity (**Figure 3, green sticks**), despite many other cavity possibilities offered by the ADV blind-docking wide grid. The 2-run top-grandchild was 4670 predicting ~ 16.7 nM affinity (**Table S2**). One of its extreme rings predicted 2 Cl atoms and a characteristic short -CO-N-SO2 - link in between 2 other extreme rings, an heritage of the 16153 parent (**Figure 4, 4670 right**). Seven oxygens contributed to lowering its LogP to 2.0 (**Table S2**). The main interactions between 4670 and the AB interface cavity, targeted [115]R at both A and B chains by two Hydrogen bonds (**Figure 4, 4670 left**), and salt bridges between the A and B chains ([119]D-[115]R and [115]R-[92]E, respectively). There were also other amino acids implicated in the AB interface that were targeted by 4670 such as those forming Hydrogen bonds to [94]K and [93]P as well as [120]F (Hydrogen bond [120]F-[53]K). The predicted 4670 competition formed by two Hydrogen bond predictions with [115]R, together with the rest of interactions may explain its nanoMolar affinities.

In contrast, most 3-run grandchildren predicted a more compact scaffold with 5 rings of top-grandchild 21422 with affinities of ~ 7.1 nM, docking to similar but not exact AB cavities than 4670 (**Figure 3, red sticks**).Four of the rings were together and displaying six oxygens which may contribute to lowering its LogP to 2.7 (**Table S2**). The important interactions with [115]R with both A and B chains predicting two Hydrogen bonds were also found for 21422 (**Figure 4, 21422 right**).

Pending of experimental confirmation, all the grandchildren, their ORF8 docked cavities, their low nanoMolar affinity predictions mentioned above and/or those of other of the top-grandchildren, could modify the conformation of the ORF8 homodimer interface.
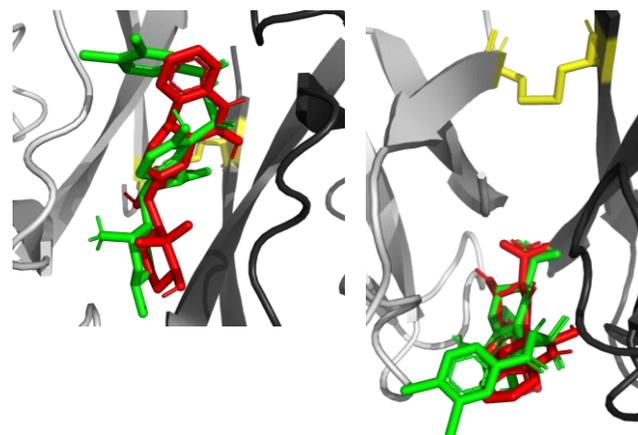

Figure 3
Bottom- (left) and side- (right) views of docked 4670 and 21422 top-grandchildren
Yellow sticks, [20]C-[20]C disulphide bond
Grey cartoons, carbon backbone of ORF8 (7jtl.pdb) asymmetric homodimer. **Ligh**, chain A. **Dark**, chain B
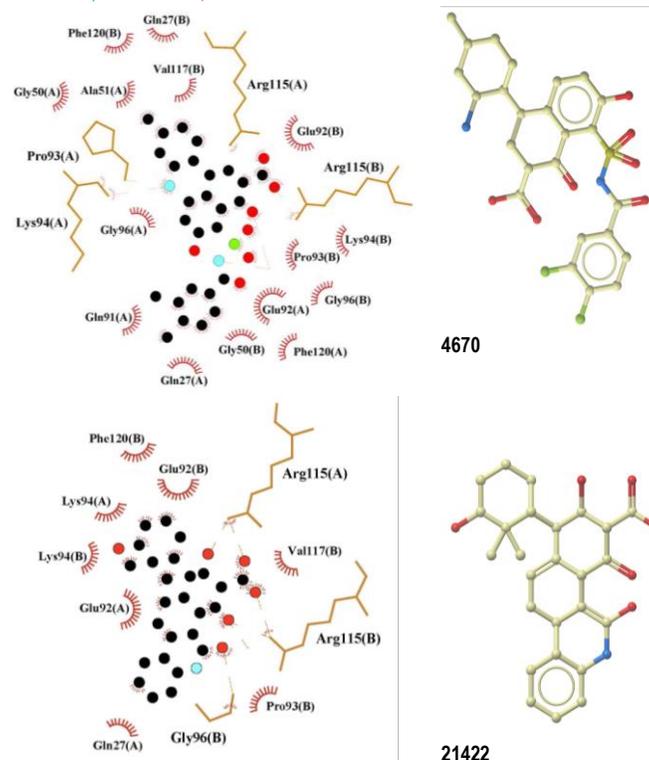**Green sticks,** 4670. **Red sticks,** 21422.

4670

21422

Figure 4
ORF8 amino acids around 4 Å of 4670 and 21422 top-grandchildren (left) and 2D schemes (right)
The 4670- and 21422-targeted amino acids docked into the ORF8 homodimer (7jtl.pdb, Chains A and B) were described by their 3 letter code followed by their ORF8 chain in parenthesis.
Left) Amino acid predictions drawn by LigPlus
Red circles, Oxygens. **Green circles,** Nitrogens. **Cyan circle,** Sulphur. **Black circles,** Carbons and Cl.
Brown sticks, ORF8 side-chains forming Hydrogen bonds (**brown hatched lines,** Hydrogen bonds).
Right) 2D chemical structures drawn by MolSoft.
Red circles, Oxygens. **Green circle,** Sulphur. **Blue circles,** Nitrogens.
Light green circles and sticks, Carbons and bonds

Discussion

Generation of children by DW-BEL co-evolution and their affinity ranking by ADV blind-docking have been combined here to unrevealed new ORF8 ligands. This automatic generation by co-evolution rather than by selection among databank compounds, was preferred to predict new molecules with some possibilities to inhibit the interferences of ORF8 SARS-COV-2 protein to the human immune response. Results successfully identified some of the ORF8 docking cavities and predict some novel putative scaffolds with many ligand variations.

Results showed only two possible docking cavities on ORF8, one into the hypervariable loop in the A chain and the other in the dimer AB chain interface. Co-evolution into the AB interface cavity showed that: **i)** the affinities predicted by a first round of co-evolutions were relatively low, at the 8-9 Kcal/mol (1400-250 nM) ranges (similar to those predicted by previously reported hypothetical disulphide

disruptors[35]) and **ii)** the numbers of grandchildren that could be generated from top-children (16153 or 46808), were more in the hundreds rather than in the thousand ranges, as usually expected . Thus, DW-BEL co-evolutions using other protein-ligand pairs in our hands usually generated thousands of fitted-children distributed in different scaffolds[48, 58, 58, 59, 46]. In contrast, targeting the ORF8 interface generated hundreds of fitted-grandchildren within unique or few scaffolds while demanding much higher computer memories than in other examples. The strong steric constrains of the highly interdigitated ORF8 interface may explain the evolutionary docking difficulties experienced. Despite those difficulties, hundreds of unique fitted-grandchildren with few scaffold variations predicted docking to their initial docking interface cavity. It was also notice that in contrast to previous work[35], no ligand molecules were predicted around the disulphide link, even after the widest ADV blind-docking, most probably due to higher steric constrains.

The DW-BEL co-evolutions applied in this work have benefit from improvements that have been incorporated step-by-step during the last months to their algorithms targeting different protein / ligand models [48, 58, 58, 59, 46]. However, it seems likely that further explorations into the vast chemical space[38, 39] to search for additional ORF8 ligand alternatives, would demand higher computer memories. In particular, to overcome the steric difficulties of higher affinity insertions into the narrow space between the A and B ORF8 interface, computers with more speed and memories would be required.

Apart from the hypothetical pseudoligand trick used here to start co-evolutions, other limitations may also include the explorations using fixed docking-cavity geometries. Considering the possible influences of amino acid side-chain mobilities on the ORF8 docking affinity estimations, those mobilities may add small variations to best accommodate any of the predicted ligands or new ones. However, their higher memory demands would made the explorations of such possibilities difficult to apply to the high numbers of grandchildren molecules identified here. Alternatively, other more adaptable or flexible docking cavities may still exist on ORF8 such as those in: **i)** the interface of the hypervariable loop A with another ORF8 monomer, such as it was earlier suggested by their surface complementarity identified in crystallographic models with another ORF8 monomer [36] or with **ii)** the interfaces with some of the many protein molecular surfaces of human proteins with already demonstrated interactions with ORF8 and/or those that have been computationally predicted by interactome studies. Further computational work could further extent these hypothetical alternatives.

Supporting information

Table S1
DW special SD-File save options to conserve 2D geometries of DW-BEL children

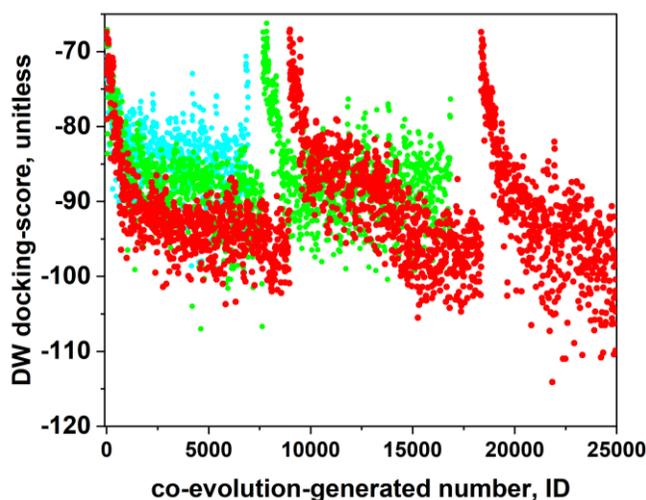| Save special options	1	2
Structure column	Docked protonation state	Structure
SD-file version	Version 3	Version 3
Atom coordinates	Docking pose	3D (1st of multiple)
☐ Include.....	Cavity & Natural Ligand	RefCompound
Compound name column	ID	ID
2D geometry conservation	MAXIMAL	MINIMAL

Figure S1
Docking-score evolution *vs* ID of fitted-children at consecutive runs
The co-evolution randomly generated unique children from the 16153 parent. As raw-children were generated they were assigned a consecutive ID number. The fitted-children are then selected. Each run re-starts co-evolution from the initial parent by avoiding duplicates.
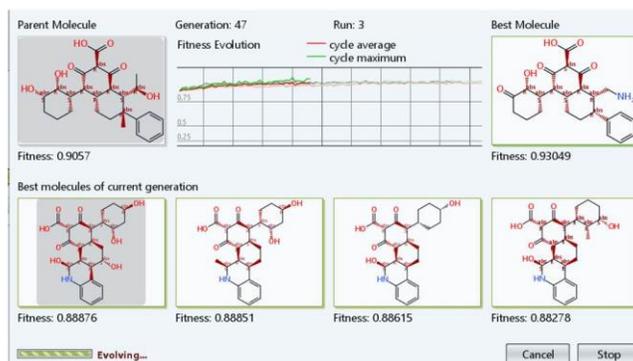**Blue,** runs = 1. **Green,** runs = 2. **Red,** runs = 3.

Figure S2
PrintScreen example of monitoring one of the DW-BEL co-evolutions

Table S2
Molecular properties of selected children and representative top-grandchildren

ID	MW	LogP	DW score	ADV Kcal/mol	nM	scaffold	ring number
16153.3	486	-0.9	-79.8	-8.9	295.1		
21422	453	2.7	-85.8	-11.1	7.2	XXX	5
18063	551	1.4	-92.6	-10.9	10.1	---	5
21918	455	1.5	-94.2	-10.9	10.1	XXX	5
21529	455	1.5	-93.2	-10.8	11.9	XXX	5
19654	441	1.2	-100.8	-10.8	11.9	XXX	5
16153.2							
4670	568	2.0	-98.5	-10.6	16.7	XXX	4
4975	546	1.7	-97.7	-10.4	23.4	XXX	4
4625	566	1.9	-107.0	-10.3	27.7	XXX	4
5943	516	0.8	-101.6	-10.3	27.7	XXX	4
6413	536	1.1	-98.5	-10.3	27.7	XXX	4
16153.1							
1083	515	2.2	-79.9	-10.4	23.5	XXX	3
3767	499	0.1	-87.4	-10.1	39.0	---	5
4574	499	0.4	-91.8	-9.9	54.6	---	4
916	497	1.7	-89.0	-9.9	54.6	---	5
958	512	0.9	-83.7	-9.9	54.6	---	4
46808	401	1.9	-84.7	-9.0	249.5		
3956	540	2.1	-113.0	-9.2	178.2	XXX	3
382	431	3.7	-88.9	-9.0	249.8	---	4
847	425	1.9	-83.3	-9.0	249.8	---	4
1217	431	3.6	-86.3	-9.0	249.8	---	4
1223	419	3.3	-78.6	-9.0	249.8	---	4

The 46808 and 16153 erizo's children were co-evolved to generated fitted grandchildren. Grandchildren 16153 were generated by increasing the number of consecutive runs.
XXX, same scaffold than the corresponding top-grandchild (**green letters**).
----, other scaffold(s)
Supporting Materials included DW tables with 100 ADV top-grandchildren for 16153.3. Tables were provided with threshold slider-filters to their *DW* and ADV docking-scores, Molecular weights and clogP properties to select particular threshold combinations.

Supporting Materials

- NTNV.dwam. A DW macro developed to automatically save, label and eliminate children molecules generated during any DW-BEL co-evolution which contained known Toxicity risks (Mutagenesis, Tumorigenicity, Reproductive Interference, Irritant) and/or any of the numerous Nasty Functions (see **Nasty_functions.dwar**). The NTNV macro uses *.sdf or *.dwar files as inputs, asks for user-renaming the input *.dwar file and renamed and saved the corresponding *.sdf file. This saving method supplied *.sdf files maintaining the 3D protein cavity docked to children 3D conformers for visualization in PyMol (using its split_states command) and/or maximal preservation of their 2D geometries for optimal consensus docking (**Table S1**). More than ~ 3000 traded drugs were taken by DW as low toxicity reference (https://github.com/thsa/datawarrior/ blob/master/src/html/ properties/properties.html). Additional information on the DW Toxicity risks evaluated can be found at the **R**egistry of **T**oxic **E**ffects of **C**hemical **S**ubstances (RTECS) (https://www.cdc.gov/niosh/docs/97-119/default.html).

- ***Nasty_functions.dwar.*** Previously defined DW Nasty functions list of small chemical fragments having known physiological interference problems, kindly supplied by Dr.T.Sander of DW (https://openmolecules.org/forum/ index.php?t=msg&th=662&start=0&).

- **100top2-grandchildren.*dwar***. These *.dwar DW tables contain 100 top-grandchildren selected by their ADV affinities from the 2-runs of 16153 . Tables are provided with threshold slider-filters to their DW and ADV docking-scores, Molecular weights and logP properties, to select for particular threshold combinations. The *.dwar files can be opened in DW available at https://openmolecules.org/datawarrior/download.htm.

- 100top2-grandchildren.pse. 100 top-grandchildren ADV complexes from the 2-runs of 16153 with the ORF8 asymmetrical homodimer crystalographic 7jtl model to be visualized in PyMol vs2.5.3.

- **100top3-grandchildren.*dwar***. These *.dwar DW tables contain 100 top-grandchildren selected by their ADV affinities from the 3-runs of 16153. Tables are provided with threshold slider-filters to their DW and ADV docking-scores, Molecular weights and logP properties, to select for particular threshold combinations. The *.dwar files can be opened in DW available at https://openmolecules.org/datawarrior/download.htm.

- 100top3-grandchildren.pse. 100 top-grandchildren ADV complexes from the 3-runs of 16153 with the ORF8 asymmetrical homodimer crystalographic 7jtl model to be visualized in PyMol vs2.5.3.

Funding

Competing interests

The author declares no competing interests

Authors' contributions

JC designed, performed and analyzed the dockings and drafted the manuscript.

Acknowledgements

References

[1] Takatsuka, H., *et al.* In silico Analysis of SARS-CoV-2 ORF8-Binding Proteins Reveals the Involvement of ORF8 in Acquired-Immune and Innate-Immune Systems. *Front Med (Lausanne).* 2022, 9: 824622 1http://dx.doi.org/0.3389/fmed.2022.824622.

[2] Arduini, A., *et al.* SARS-CoV-2 ORF8: A Rapidly Evolving Immune and Viral Modulator in COVID-19. *Viruses.* 2023, 15: http://dx.doi.org/10.3390/v15040871.

[3] Hassan, S.S., *et al.* A unique view of SARS-CoV-2 through the lens of ORF8 protein. *Comput Biol Med.* 2021, 133: 104380 http://dx.doi.org/10.1016/j.compbiomed.2021.104380.

[4] Kohyama, M., *et al.* SARS-CoV-2 ORF8 is a viral cytokine regulating immune responses. *Int Immunol.* 2023, 35: 43-52 http://dx.doi.org/10.1093/intimm/dxac044.

[5] Zinzula, L. Lost in deletion: The enigmatic ORF8 protein of SARS-CoV-2. *Biochem Biophys Res Commun.* 2021, 538: 116-124 http://dx.doi.org/10.1016/j.bbrc.2020.10.045.

[6] Akaishi, T., *et al.* Insertion/deletion hotspots in the Nsp2, Nsp3, S1, and ORF8 genes of SARS-related coronaviruses. *BMC Ecol Evol.* 2022, 22: 123 http://dx.doi.org/10.1186/s12862-022-02078-7.

[7] Twohig, K.A., *et al.* Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *Lancet Infect Dis.* 2022, 22: 35-42 http://dx.doi.org/10.1016/S1473-3099(21)00475-8.

[8] Young, B.E., *et al.* Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. *Lancet.* 2020, 396: 603-611 http://dx.doi.org/10.1016/S0140-6736(20)31757-8.

[9] Hassan, S.S., *et al.* An issue of concern: unique truncated ORF8 protein variants of SARS-CoV-2. *PeerJ.* 2022, 10: e13136 http://dx.doi.org/10.7717/peerj.13136.

[10] Bykova, A., *et al.* The 29-nucleotide deletion in SARS-CoV: truncated versions of ORF8 are under purifying selection. *BMC Genomics.* 2023, 24: 387 http://dx.doi.org/10.1186/s12864-023-09482-3.

[11] Brandt, D., *et al.* Multiple Occurrences of a 168-Nucleotide Deletion in SARS-CoV-2 ORF8, Unnoticed by Standard Amplicon Sequencing and Variant Calling Pipelines. *Viruses.* 2021, 13: http://dx.doi.org/10.3390/v13091870.

[12] DeRonde, S., *et al.* Identification of a Novel SARS-CoV-2 Strain with Truncated Protein in ORF8 Gene by Next Generation Sequencing. *Res Sq.* 2021: http://dx.doi.org/10.21203/rs.3.rs-413141/v1.

[13] Vinjamuri, S., *et al.* SARS-CoV-2 ORF8: One protein, seemingly one structure, and many functions. *Front Immunol.* 2022, 13: 1035559 http://dx.doi.org/10.3389/fimmu.2022.1035559.

[14] Sadegh, S., *et al.* Exploring the SARS-CoV-2 virus-host-drug interactome for drug repurposing. *Nat Commun.* 2020, 11: 3518 http://dx.doi.org/10.1038/s41467-020-17189-2.

[15] Messina, F., *et al.* COVID-19: viral-host interactome analyzed by network based-approach model to study pathogenesis of SARS-CoV-2 infection. *J Transl Med.* 2020, 18: 233 http://dx.doi.org/10.1186/s12967-020-02405-w.

[16] Liu, P., *et al.* SARS-CoV-2 ORF8 reshapes the ER through forming mixed disulfides with ER oxidoreductases. *Redox Biol.* 2022, 54: 102388 http://dx.doi.org/10.1016/j.redox.2022.102388.

[17] Rashid, F., *et al.* The ORF8 protein of SARS-CoV-2 induced endoplasmic reticulum stress and mediated immune evasion by antagonizing production of interferon beta. *Virus Res.* 2021, 296: 198350 http://dx.doi.org/10.1016/j.virusres.2021.198350.

[18] Wang, X., *et al.* SARS-CoV-2 ORF8 Protein Induces Endoplasmic Reticulum Stress-like Responses and Facilitates Virus Replication by Triggering Calnexin: an Unbiased Study. *J Virol.* 2023, 97: e0001123 http://dx.doi.org/10.1128/jvi.00011-23.

[19] Lehrer, S. and Rheinstein, P.H. Alignment of human KAT2A (GCN5) histone acetyltransferase and SARS-CoV-2 Orf8 viral proteins. *Chronic Dis Transl Med.* 2022, 9: 263-5 http://dx.doi.org/10.1002/cdt3.56.

[20] Chou, J.M., *et al.* The ORF8 Protein of SARS-CoV-2 Modulates the Spike Protein and Its Implications in Viral Transmission. *Front Microbiol.* 2022, 13: 883597 http://dx.doi.org/10.3389/fmicb.2022.883597.

[21] Hussain, M., *et al.* Immunoinformatic analysis of structural and epitope variations in the spike and Orf8 proteins of SARS-CoV-2/B.1.1.7. *J Med Virol.* 2021, 93: 4461-4468 10.1002/jmv.26931.

[22] Kim, I.J., *et al.* SARS-CoV-2 protein ORF8 limits expression levels of Spike antigen and facilitates immune evasion of infected host cells. *J Biol Chem.* 2023, 299: 104955 http://dx.doi.org/10.1016/j.jbc.2023.104955.

[23] Kumar, J., *et al.* SARS-CoV-2-encoded ORF8 protein possesses complement inhibitory properties. *J Biol Chem.* 2023, 299: 102930 http://dx.doi.org/10.1016/j.jbc.2023.102930.

[24] Fernandes, M.F., *et al.* Effect of cannabidiol on apoptosis and cellular interferon and interferon-stimulated gene responses to the SARS-CoV-2 genes ORF8, ORF10 and M protein. *Life Sci.* 2022, 301: 120624 http://dx.doi.org/10.1016/j.lfs.2022.120624.

[25] Chen, J., *et al.* Severe Acute Respiratory Syndrome Coronavirus 2 ORF8 Protein Inhibits Type I Interferon Production by Targeting HSP90B1 Signaling. *Front Cell Infect Microbiol.* 2022, 12: 899546 http://dx.doi.org/10.3389/fcimb.2022.899546.

[26] Geng, H., *et al.* SARS-CoV-2 ORF8 Forms Intracellular Aggregates and Inhibits IFNgamma-Induced Antiviral Gene Expression in Human Lung Epithelial Cells. *Front Immunol.* 2021, 12: 679482 http://dx.doi.org/10.3389/fimmu.2021.679482.

[27] Li, J.Y., *et al.* The ORF6, ORF8 and nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. *Virus Res.* 2020, 286: 198074 http://dx.doi.org/10.1016/j.virusres.2020.198074.

[28] Hamdorf, M., *et al.* The unique ORF8 protein from SARS-CoV-2 binds to human dendritic cells and induces a hyper-inflammatory cytokine storm. *J Mol Cell Biol.* 2023: http://dx.doi.org/10.1093/jmcb/mjad062.

[29] Lin, X., *et al.* Unconventional secretion of unglycosylated ORF8 is critical for the cytokine storm during SARS-CoV-2 infection. *PLoS Pathog.* 2023, 19: e1011128 http://dx.doi.org/10.1371/journal.ppat.1011128.

[30] Wu, X., *et al.* Viral Mimicry of Interleukin-17A by SARS-CoV-2 ORF8. *mBio.* 2022, 13: e0040222 http://dx.doi.org/10.1128/mbio.00402-22.

[31] Beaudoin-Bussieres, G., *et al.* SARS-CoV-2 Accessory Protein ORF8 Decreases Antibody-Dependent Cellular Cytotoxicity. *Viruses.* 2022, 14: http://dx.doi.org/10.3390/v14061237.

[32] Wu, X., *et al.* Secreted ORF8 induces monocytic pro-inflammatory cytokines through NLRP3 pathways in patients with severe COVID-19. *iScience.* 2023, 26: 106929 http://dx.doi.org/10.1016/j.isci.2023.106929.

[33] Zhang, Y., *et al.* The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-Iota. *Proc Natl Acad Sci U S A.* 2021, 118: http://dx.doi.org/10.1073/pnas.2024202118.

[34] Chaudhari, A.M., *et al.* Defective ORF8 dimerization in SARS-CoV-2 delta variant leads to a better adaptive immune response due to abrogation of ORF8-MHC1 interaction. *Mol Divers.* 2023, 27: 45-57 http://dx.doi.org/10.1007/s11030-022-10405-9.

[35] Selvaraj, C., *et al.* SARS-CoV-2 ORF8 dimerization and binding mode analysis with class I MHC: computational approaches to identify COVID-19 inhibitors. *Brief Funct Genomics.* 2023, 22: 227-242 http://dx.doi.org/10.1093/bfgp/elac046.

[36] Flower, T.G., *et al.* Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. *Proc Natl Acad Sci U S A.* 2021, 118: http://dx.doi.org/10.1073/pnas.2021785118.

[37] Cheng, Y. and Peng, X. In silico study on the effects of disulfide bonds in ORF8 of SARS-CoV-2. *Phys Chem Chem Phys.* 2022, 24: 16876-16883 http://dx.doi.org/10.1039/d2cp01724e.

[38] Wu, F., *et al.* Glycosylated, Lipid-Binding, CDR-Like Domains of SARS-CoV-2 ORF8 Indicate Unique Sites of Immune Regulation. *Microbiol Spectr.* 2023, 11: e0123423 http://dx.doi.org/10.1128/spectrum.01234-23.

[39] Pereira, F. Evolutionary dynamics of the SARS-CoV-2 ORF8 accessory gene. *Infect Genet Evol.* 2020, 85: 104525 http://dx.doi.org/10.1016/j.meegid.2020.104525.

[40] Laha, S., *et al.* Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. *Infect Genet Evol.* 2020, 85: 104445 http://dx.doi.org/10.1016/j.meegid.2020.104445.

[41] Islam, S., *et al.* Structural and functional effects of the L84S mutant in the SARS-COV-2 ORF8 dimer based on microsecond molecular dynamics study. *J Biomol Struct Dyn.* 2023: 1-18 http://dx.doi.org/10.1080/07391102.2023.2228919.

[42] Matsuoka, K., *et al.* SARS-CoV-2 accessory protein ORF8 is secreted extracellularly as a glycoprotein homodimer. *J Biol Chem.* 2022, 298: 101724 http://dx.doi.org/10.1016/j.jbc.2022.101724.

[43] Lin, X., *et al.* ORF8 contributes to cytokine storm during SARS-CoV-2 infection by activating IL-17 pathway. *iScience.* 2021, 24: 102293 http://dx.doi.org/10.1016/j.isci.2021.102293.

[44] Hachim, A., *et al.* ORF8 and ORF3b antibodies are accurate serological markers of early and late SARS-CoV-2 infection. *Nat Immunol.* 2020, 21: 1293-1301 http://dx.doi.org/10.1038/s41590-020-0773-7.

[45] Chen, X., *et al.* Crystal Structures of Bat and Human Coronavirus ORF8 Protein Ig-Like Domain Provide Insights Into the Diversity of Immune Responses. *Front Immunol.* 2021, 12: 807134 http://dx.doi.org/10.3389/fimmu.2021.807134.

[46] Coll, J.M. Exploring non-toxic co-evolutionary docking. *ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/6512b162ade1178b2424c325: http://dx.doi.org/10.26434/chemrxiv-2023-rr5b0.

[47] Bryant, P., *et al.* Structure prediction of protein-ligand complexes from sequence information with Umol. *bioRxiv.* 2023: http://dx.doi.org/10.1101/2023.11.03.565471.

[48] Coll, J.M. Evolutionary-docking targeting bacterial FtsZ. *ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/6405c36fcc600523a3bcb679: http://dx.doi.org/10.26434/chemrxiv-2023-ld9d3

[49] Wahl, J., *et al.* Accuracy evaluation and addition of improved dihedral parameters for the MMFF94s. *J Cheminform.* 2019, 11: 53 http://dx.doi.org/10.1186/s13321-019-0371-6.

[50] Coll, J. Star-shaped Triazine-derivatives: would they crossbind SARS-CoV-2 spike helices? *ChemRxiv.* 2021, https://chemrxiv.org/engage/chemrxiv/article-details/6133c1096563696d9d222bbd: http://dx.doi.org/10.33774/chemrxiv-2021-xb6sx-v2.

[51] Lorenzo, M.M., *et al.* Would it be possible to stabilize prefusion SARS-CoV-2 spikes with ligands? *ChemRxiv.* 2021: http://dx.doi.org/10.26434/chemrxiv.13453919.v2.

[52] Bermejo-Nogales, A., *et al.* Computational ligands to VKORC1s and CYPs. Could they predict new anticoagulant rodenticides? *BioRxiv.* 2021: http://dx.doi.org/10.1101/2021.01.22.426921.

[53] Agarwal, R. and Smith, J.C. Speed vs Accuracy: Effect on Ligand Pose Accuracy of Varying Box Size and Exhaustiveness in AutoDock Vina. *Mol Inform.* 2023, 42: e2200188 http://dx.doi.org/10.1002/minf.202200188.

[54] Dallakyan, S. and Olson, A.J. Small-molecule library screening by docking with PyRx. *Methods Mol Biol.* 2015, 1263: 243-50 http://dx.doi.org/10.1007/978-1-4939-2269-7_19.

[55] Morris, G.M., *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 2009, 30: 2785-91 http://dx.doi.org/10.1002/jcc.21256.

[56] Huey, R., *et al.* A semiempirical free energy force field with charge-based desolvation. *J Comput Chem.* 2007, 28: 1145-52 http://dx.doi.org/10.1002/jcc.20634.

[57] Trott, O. and Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010, 31: 455-61 http://dx.doi.org/10.1002/jcc.21334.

[58] Coll, J.M. Anticoagulant rodenticide novel candidates predicted by evolutionary docking. *ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/6479b8cfbe16ad5c57577cce: http://dx.doi.org/10.26434/chemrxiv-2023-gh4xl-v2.

[59] Coll, J. Could Acinetobacter baumannii Lol-abaucin docking be improved? *ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/649aa71aba3e99daef1d1756: http://dx.doi.org/10.26434/chemrxiv-2023-962ht.