New star-shaped ligands generated by evolutionary fitting the Omicron spike inner-cavity

short title: anti-Omicron evolutionary docking

Coll, J*.
Department of Biotechnology. Centro Nacional INIA-CSIC. Madrid, Spain.
* Corresponding author
Emails: julio.coll.m@csic.es; juliocollm@gmail.com
Julio Coll, orcid: 0000-0001-8496-3493

Abstract

Predictions generated by evolutionary docking of star-shaped ligands targeting the prefusion state of Omicron variants are described here. For this, one selected star-shaped molecule previously identified with the seeSAR program, was used as parent to randomly generate made-on-demand large children libraries evolutionary selected for best fitting to the Omicron spike top-to-bottom inner-cavity with the DataWarrior subprogram. The generated children docking-scores were consensed by AutoDockVina ranks normalized by molecular size and hydrophobicity. These explorations identified one new main chemotype and variants with improved specificity and exceptional nanomolar affinities, predicting aqueous soluble molecules targeting the prefusion state of Omicron spike inner-cavity of trimeric alpha-helices.

Keywords: Omicron; spikes; star-shaped molecules; evolutionary libraries; docking; DataWarrior; AutoDockVina

Introduction

Multiple variants have emerged from those initial **S**evere **A**cute **R**espiratory **S**yndrome **C**oron**a****v**irus 2 (SARS-Cov-2) isolates to the moderate dangerous Omicron variants with increased spreading power. The Omicron variants presumably resulted after selection by resistance to natural infection [1-3] or to vaccination[4]. Most recent Omicron variants are still under intensive study [5-8], but world-wide reports from several laboratories have confirmed an unusually high number of spike (S) mutations (~30) correlating with increased immune-evasion to vaccination, increase spreading rates and tighter S packing[9-11]. Most of the identified Omicron S mutations changed the SARS-Cov-2 surface sites implicated in recognition of the human ACE receptor[2, 12, 13], one of the first requirements for viral/human membrane fusion to initiate Omicron infection. Therefore, most recent studies have focused on mutations affecting the human ACE receptor binding[10, 13, 14], on Omicron surfaces by drug repurposing[15-19].

Rather than targeting Omicron receptor binding, in this work we focused on computationally designing new ligands targeting the internal top-to-bottom cavity surrounded by the trimer central α-helices using one recent evolutionary docking method. This work continued our previous efforts to propose possible blocking of the fusion-dependent α-helix internal changes that precede viral/host membrane fusion[20, 21]. Such inner-cavity has been implicated in the Omicron early prefusion-fusion switch triggered after human ACE receptor recognition [20, 21]. This inner-cavity seems to be highly conserved in Omicron variants, since only three Omicron new mutations (Q954H, N969K, L981F) [22] have been described that could directly affect the prefusion-fusion switch [23]. For this, we started from previous computationally-identified star-shaped triazine-core molecules cross docking the inner-cavity of the SARS-COV-2 wild-type all-down prefusion state isolate of 2021 [20, 21]. The predicted targeted site of star-shaped molecules were located at the upper S2 trimer which formed a central top-to-bottom cavity of 7-20 Å of diameter surrounded by 3 α-helices [24] (**Figure 1**). Cross binding of the inner-cavity could stabilize the α-helices at their closed prefusion state delaying or inhibiting the early switch to fusion, therefore reducing or blocking subsequent viral infection. Similar effects had been described before by experimental mutations at the ^{986}KV residues of the bended elbow of the S2 trimeric spikes[16, 25, 26]. Therefore our previous and present objective focused on the computational exploration of possible new star-shaped compounds which may block the Omicron early fusion with drugs rather than with mutations.

Each of the targeted inner-cavity site at the S2 subunit of the monomeric spike of SARS-CoV-2, including their Omicron variants, contained part of the HR1 heptad-repeat (910 to 988) and CH central helix (986-1033) extending from the 960 to the 1010 amino acid residues, apparently accessible only when in the prefusion all-down state, each of the S2 trimeric α-helices are bended in an spring-loaded mechanism. The α-helix bending unfolds during the early steps of fusion elongating to one trimeric coiled-coil bundle of rigid linear α-helix conformations at the postfusion states. Similar spring-loaded switch mechanisms but with other amino acid α-helix sequences have been demonstrated in many other enveloped viruses [27, 28].

Most previous experimental and/or computational search for anti-coronaviral molecules has been focused on approved drugs (drug repurposing)[29-31] to other protein targets rather to the coronavirus inner-cavity. Most of those earlier works included the inhibition of the S2 coiled-coil bundle by complementary peptides [23, 32-35], the blocking of the surface interphase of S1 with the ACE2 human receptor [18, 19, 36], and/or the active sites of viral proteases [30, 37, 38] among others. In contrast, our previous efforts to use ligands rather than mutations[20, 21], explored the inner-cavity. In the past we have used limited explorations of chemical spaces by combining computational strategies previously reported by others, including neural networks and large public libraries[39], [40-43], [44], [45-48]. The results of our previous studies, predicted subnanomolar binding ranges for Triazine-cores branched with Trihydroxyl-Triphenyls[21]. However, those star-shaped molecules tend to be highly hydrophobic and toxic. With the progression of the new Omicron variants, star-shaped molecules reducing any possible toxicity of those triazine core derivatives [49-52] while preserving the cross docking to Omicron inner-cavity would be more desirable now.

To address those requirements, this work computationally explored a wider chemical/chemotype space by generating novel libraries designed-on-demand by random evolution of one selected parent corresponding to one previously identified star-shaped molecule binding the inner-cavity of first SARS-Cov-2 isolates obtained with the seeSAR (sS) program algorithm[20, 21]. The children molecules were generated by extending the number of new star-shaped molecules while still fitting the inner-cavity of the Omicron variants with one **D**ata**W**arrior (DW) algorithm. Each of the evolutionary library run, selected thousands of new children possibilities from the parent molecule, discarding tens of thousands of other generated children that fail to fit into the inner-cavity and/or to other multiple molecular criteria. The DW selected children were filtered for absence of known toxicities and two of their top-leads re-used as new parents. Once the re-evolved non-toxic children libraries were generated, consensus docking-scores were obtained by comparing DW results with those obtained from the **A**uto**D**ock**V**ina (ADV) algorithm after being normalized for excessive molecular weight and hydrophobicity with **L**igand **E**fficiency **L**ipophilic **P**rize (LELP) metric. VENN diagrams among DW, ADV and LELP ranks were used to compare and select top-leads.

To derive the final list of top-leads, some particular methods were used for this work. For instance, seeSAR leads from our previous work [20, 21] were selected to start our previously described sequential rather than previous independent consensus strategy, similar to the one applied in our most recent work[53]. Comparisons using ranks rather than absolute docking-scores, were used as recommended by many other authors for higher prediction accuracies[54-58], [59, 60]. Finally, lead tendencies to apparent and most probably erroneous high affinities such as those only due to unspecific increase in molecular weights and hydrophobicities[61, 62], were corrected with a unique LELP parameter for ligand efficiency normalization, selected among other methods [63-66].

In this work, the restricted parent docking-cavity, to generate a large list of children possibilities improving the DW docking-score, molecular weight and hydrophobicity, helped the generation of many alternative leads to the first one identified before with sS and selection among the vast wide chemical space, otherwise impossible to reach by any other methods. Filtering by non-toxicity, and docking to ADV, a different program with wider grid-dependence and LELP normalized, helped to develop a consecutive consensus among the DW+ADV+LELP results to generate a top-lead downsized list with presumably higher accuracies than before. The top-lead list may suggest a possible chemical synthesis for validation through experimental work.

.

Computational Methods

Wild-type Omicron 3D target model

The SARS-CoV-2 Omicron S trimer model coded by the 7tnw.pdb crystallography file (**R**esearch **C**ollaboratory for **S**tructural **B**ioinformatics, RCSB, **P**rotein **D**ata **B**ank PDB ID) of the wild-type (without the PP stabilizing mutations) all-down S prefusion state [13] was employed here. The 2021 isolate earliest wild-type SARS-COV-2 6xr8 model [20, 21] and the 7to4[13] Omicron one-up fusion-competent model, were used for comparison.

The most appropriated ligand for this work was selected among those leads identified with the seeSAR program and described before as one of the

derivatives of the 2-[4,6-bis(2-hydroxyphenyl)-1,3,5-triazin-2yl] phenol (PubChem 135616181) [13, 21]. The double bonds of the triazine core of that ligand were substituted for single bonds to reduce their toxicity while maintaining low its docking-score. The resulting parent ligand will be called TTT, since it was absent from the main public data banks. TTT was selected as the parent star-shaped molecule due to its simple symmetric structure despite predicting a low docking-score (~0.2 nM) to the α-helix inner-cavity of the wild-type (no stabilizing mutations) prefusion state of SARS-COV-2 6xr8 earliest used as a model [21].

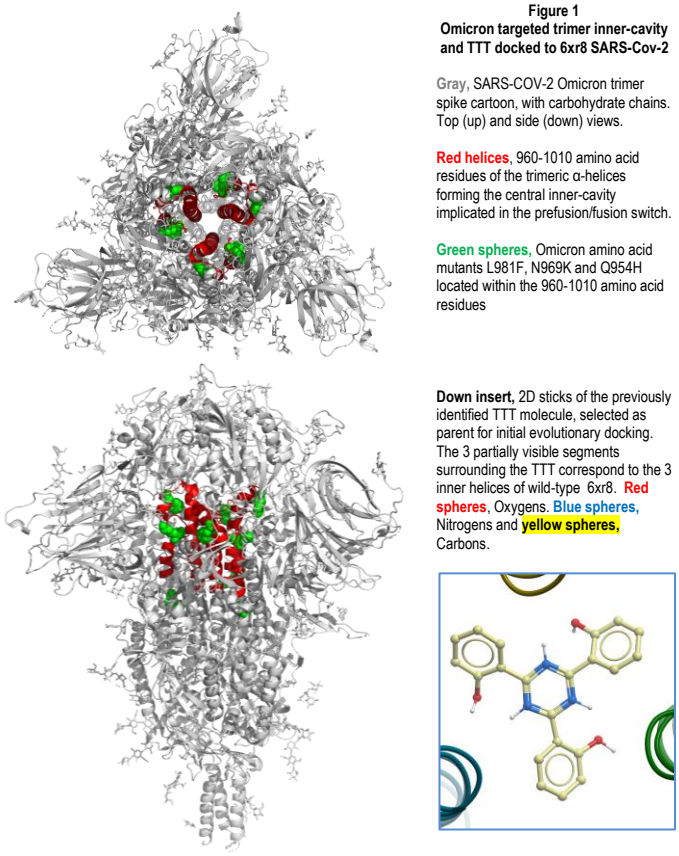

Figure 1
Omicron targeted trimer inner-cavity and TTT docked to 6xr8 SARS-Cov-2

Gray, SARS-COV-2 Omicron trimer spike cartoon, with carbohydrate chains. Top (up) and side (down) views.

Red helices, 960-1010 amino acid residues of the trimeric α-helices forming the central inner-cavity implicated in the prefusion/fusion switch.

Green spheres, Omicron amino acid mutants L981F, N969K and Q954H located within the 960-1010 amino acid residues

Down insert, 2D sticks of the previously identified TTT molecule, selected as parent for initial evolutionary docking. The 3 partially visible segments surrounding the TTT correspond to the 3 inner helices of wild-type 6xr8. Red spheres, Oxygens. Blue spheres, Nitrogens and yellow spheres, Carbons.

The "Build Evolutionary Library"

The "Build evolutionary library" subprogram included into the DataWarrior (DW) written in Java (dw550win.zip for Windows) was locally downloaded (https://openmolecules.org/datawarrior/download.html) following the details described by DW and our previous work [53].

Restricting the DW docking to the user-supplied ligand-protein cavity complex defined by their amino acid side-chains, DW ranks those ligands by predicting their lowest unitless docking-scores. To generate new children molecules from selected parents, DW randomly added small molecular modifications generating 128 children molecules per generation, rank them by their natural product properties and fitting them to user-defined multiple criteria. The best 8 fitting molecules are then selected for additional modifications into the next generation. Generations automatically continued until a fitness plateau was reached [53]. The user-defined fitness criteria and weight values were described in the Figure legends. The docking subprogram of DW ("Dock structures into protein cavity") used its proprietary mmff94s+ force-field version[67] for energy minimization to best preserve the geometry of the output molecules.

Usually, 50 to 150 generations (500-2000 unique children molecules) were automatically generated per parent molecule reaching ~ 0.7 to 0.9 fitness to the multiple criteria. The raw results were saved as *.dwar files for storage of the complete evolutionary data including fitness evolution and cavity-complex images. The raw children data were then filtered by excluding predicted high and low toxic DW chemical properties (mutagenesis, tumorigenicity, reproductive interference, irritant, and nasty functions). The toxic-filtered children were saved as both *.dwar and special *.sdf (vs3) files maintaining all evolutionary information including the 3D protein cavity docked to children ligand conformers for more detailed visualization in PyMol (using its split_states command)[53].

The AutoDockVina docking program

The AutoDockVina (ADV) program written in Python vs3.8 and included in the PyRx-098/PyRx-1.0 package was used as described before https://pyrx.sourceforge.io/). Briefly, ADV docking was performed after *.pdbqt file conversion of the trimer protein and ligands[68] using the mmff94s (Merck) force-field for energy minimization [67, 69, 70]. Comparison of ligand *InChiKeys* before and after docking monitored for possible alterations of molecular geometries[67, 69-72]. ADV generates many possible conformers using the rotatable bonds of the input ligands and selects those with the lower docking-scores for output[73];[74]. Only the conformer with the lowest Kcal/mol docking-score per ligand was selected for the present studies.

A grid of 50x50x50 Å around the automatically centered trimer at 199x199x186 (z corrected) by PyRx/ADV was selected to explore any best docking-cavity alternatives, in contrast to the restricted cavity for DW docking.

To correct for the unspecificities caused by increased molecular weights and hydrophobicities[61, 62], the **L**igand **E**fficiency **L**ipophilic **P**rize (LELP)[77] unique parameter was chosen among other alternatives [63-66]. LELP was applied after converting the Kcal/mol of the ADV output values to Ki in nM by the formula $10^9 * (\exp^{(Kcal/mol/0.592)})$. LELP was then calculated from the nM values by DW chemical property options.

.

Other computational programs

To identify and predict seeSAR leads used for the initial parent definition, the seeSAR vs.10 package (https://www.biosolveit.de/SeeSAR/) was used as described before [20, 21]. The DataWarrior (Osiris, vs5.5.0.Idorsia Ph.Ltd, https://openmolecules.org/datawarrior/download.html)[75] and MolSoft (ICM Molbrowser vs3.9Win64bit, https://www.molsoft.com/download.html) were used for docking and manipulating the *.sdf files, as described before [53]. The Origin program (OriginPro 2022, 64 bit, Northampton, MA, USA) (https://www.originlab.com/) was used for calculations and drawing Figures. The predicted trimer-ligand complexes were visualized in PyRx 098/PyRx1.0 (Mayavi), Discover Studio Visualizer v21.1.0.20298 (Dassault Systemes Biovia Corp, 2020, https://discover.3ds.com/discovery-studio-visualizer-download) and PyMOL 2.5.3 (https://www.pymol.org/). Hydrophobic and Hydrogen-bonded amino acid interactions predicted by the docked ligand complexes were identified by LigPlus vs2.2.8 (https://www.ebi.ac.uk/thornton-srv/software/LigPlus/download.htm), and their LigPlot results were visualized in PyMol. Venn diagrams were build up and optimized using a web tool (https://bioinformatics.psb.ugent.be/webtools/Venn/). A multithreading multi-core i9 (47 CPU) PCSpecialist computer (AMD Ryzen Threadripper 3960X) with 64 Gb of RAM (Corsair Vengeance DDR4 at 3200 MHz, 4 x 16 GB) was used for all the computational work (https://www.pcspecialist.es/).

Results

The wild-type S trimer spike of SARS-Cov-2 Omicron without PP stabilizing mutations (model 7tnw), was selected for this work. Comparison of the 7tnw S Omicron with the earliest SARS-COV-2 wild-type all-down 6xr8 resulted in an RMSD of 2.8 Å. Despite such small differences, ADV docking-scores of the initial TTT ligand were 0.20 and 0.67 nM for 6xr8 and 7tnw (Omicron), respectively. These results suggested that the L981F, N969K and Q954H mutations of the Omicron variants that mapped inside the 960-1010 residues inner-cavity targeted in this work have some influence in the ADV docking-scores and therefore would require additional studies.

DataWarrior evolutionary docking was selected to generate large amounts of ligands fitting the 960-1010 targeted Omicron inner-cavity. Preliminary evolutionary results with the selected TTT-parent explored possible simplifications varying their central core ring of 6 atoms (three Carbons and three Nitrogens) (top 2D structures at **Figure 2**) with core alternatives such as trimeric cores of three Nitrogens, three Carbons or one central Nitrogen. However none of the children generated from any of the core alternatives mentioned above, reduced the initial TTT-parent binding-scores (results not shown). After many additional tests, optimal fitting criteria were adjusted to maintain the TTT-parent 6-atom ring structure allowing random evolution of any atom (including the central 6-atom core). Additional fitting criteria were included maintaining molecular weights, reducing hydrophobicity, maintaining the maximal number of rings and allowing for a higher number of Nitrogen and Oxygens (**Figure 2**, legend). According to the results, most of the randomly generated and fitting selected TTT-children predicted lower docking-scores when compared to their corresponding TTT-parent (compare dashed line and circle profile at **Figure 2,** gray).

The TTT-children predicted two main chemotype top-leads displaying one common central core of 6-atoms and three variable atom branches containing one 6-carbon ring each. The results confirmed that any other possible cores displaying lower numbers of atoms were again not selected by the evolutionary docking when extending random variations to any of the atoms of the initial parent. One of the top-lead 6-atom chemotypes contained cores with two Nitrogens (**Figure 2up,** green 9990-child) and the other chemotype contained cores with one Nitrogen and one Oxygen, the rest being 4 Carbons (**Figure 2up,** blue 4976-child). Compared to the TTT-parent, both the 4976 and 9990 TTT-children

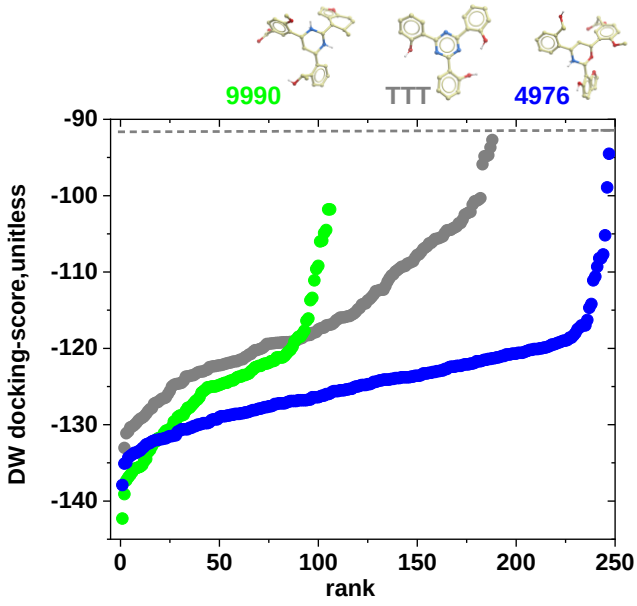

Figure 2
Rank profiles of TTT-, 9990- and 4976-children
The represented children have no known toxicities. **Inserts above**, 3D structures of the parents used for each of the evolutionary dockings including the 9990 and 4976 TTT-children (**Table S1**). **Red spheres**, Oxygens. **Blue spheres**, Nitrogens and **yellow spheres**, Carbons. Their unique identification numbers were automatically assigned by their generation order from the parent during evolution.
Horizontal dashed gray line, DW docking-score of the TTT- parent.
Gray, TTT-parent profile fitted to <500 Dalton, logP<2, ring-count <4, Nitrogens<6, Oxygens<6
Blue, TTT 4976-children profile fitted to <550 Dalton, logP<2, ring-count<4, Nitrogens<3, Oxygens <9
Green, TTT 9990-children profile fitted to <550 Dalton, logP<2, ring-count<4, Nitrogens<3, Oxygens <9

predicted lower DW and ADV docking-scores, lower hydrophobicities, and higher specificities, despite their slightly higher molecular weights (**Table S1**). The 9990-child predicted a higher number of hydrophobic contacts and one Hydrogen bond compared to the TTT-parent (no Hydrogen bonds) and to the 4796-child (two Hydrogen bonds) (**Table S2**).

When the two TTT-children top-leads were independently re-used as new parents for additional evolutionary dockings, the generated 9990-children predicted more and lower DW docking-score profiles (**Figure 2, green profiles**) compared to the TTT-parent (**Figure 2, gray profiles**) or to the 4976-children (**Figure 2, blue profiles**). Therefore, further studies were focused to the 9990-children.

Because no algorithm is still capable of predicting docking-scores with good enough accuracies for each different chemotype, and DW explores only the initially supplied docking cavity, a minimal consensus of docking-scores was searched including also exploration of other alternative protein cavities. For that, the well known AutoDockVina (ADV) program was selected, mainly because **i)** ADV may detect any other docking cavity alternatives inside different size grids (such as the one selected here for 50x50x50 Å) and **ii)** ADV reports their docking-score output results in Kcal/mol values (~ΔG docking energies), which allows for normalization of specificity corrections. Normalization corrects for apparent low docking-scores which may be due only to unspecific high molecular size or hydrophobicity, as reported in many other docking systems (and also observed during this work). The LELP ligand efficiency unique parameter was chosen here for specificity normalization. LELP was calculated within the DW program (LogP/LE ratios), taken into account both the logP lipophilicity (as reliable prediction of hydrophobicity) and the number of non-hydrogen atoms (proportional to their molecular weight).

The comparisons between ADV *versus* DW docking-scores by ranks, predicted abundant top-leads below the values of the 9990-parent including exceptionally low ADV binding-scores (**Figure 3A** and **Table S1**). To select for the best top-lead alternatives, 9990-children common the top-lead ranks for DW, ADV and LELP were graphically selected using VENN diagrams (**Figure 3B** and **Table S1**). The thresholds of DW, ADV and LELP values to be included into the VENN diagrams were optimized to result in a low practical number of top-leads, which in this work was arbitrarily set to 6. The following thresholds for the VENN comparison were optimized with the following values, **i**) DW docking-scores from -142 to -124, resulting in 52 top-leads, **ii**) DW docking-scores 0.1 to 1.3 nM, resulting in 26 top-leads and **iii)** LELP values from -3.6 to 2.9 (nearer to 0), resulting in 26 top-leads. Detailed study of the resulting 6 top-leads showed that only one main chemotype was predicted maintaining the common central 6-atom core with 2 Nitrogens and varying chemical structures in its three arms. Some of their molecular characteristics (**Table S2**) and the corresponding 2D structures (**Figure 3B**) were represented. PyMol visualization of their docked complexes with the Omicron inner-cavity confirmed the crossbinding of the three α-helices (representative **Figure 4, 4040** and not shown). LigPlus localized their amino acid ligand interactions of the corresponding docked predictions (**Table S2**). Given their exceptional docking-scores of most of the 9990-children, the remaining unexplored chemical spaces and the difficulties to exactly reproduce randomly generated data, all the 9990-children and their predicted properties were included for any interested readers to explore further possibilities of any other top-lead selection alternatives (**Table S3**).

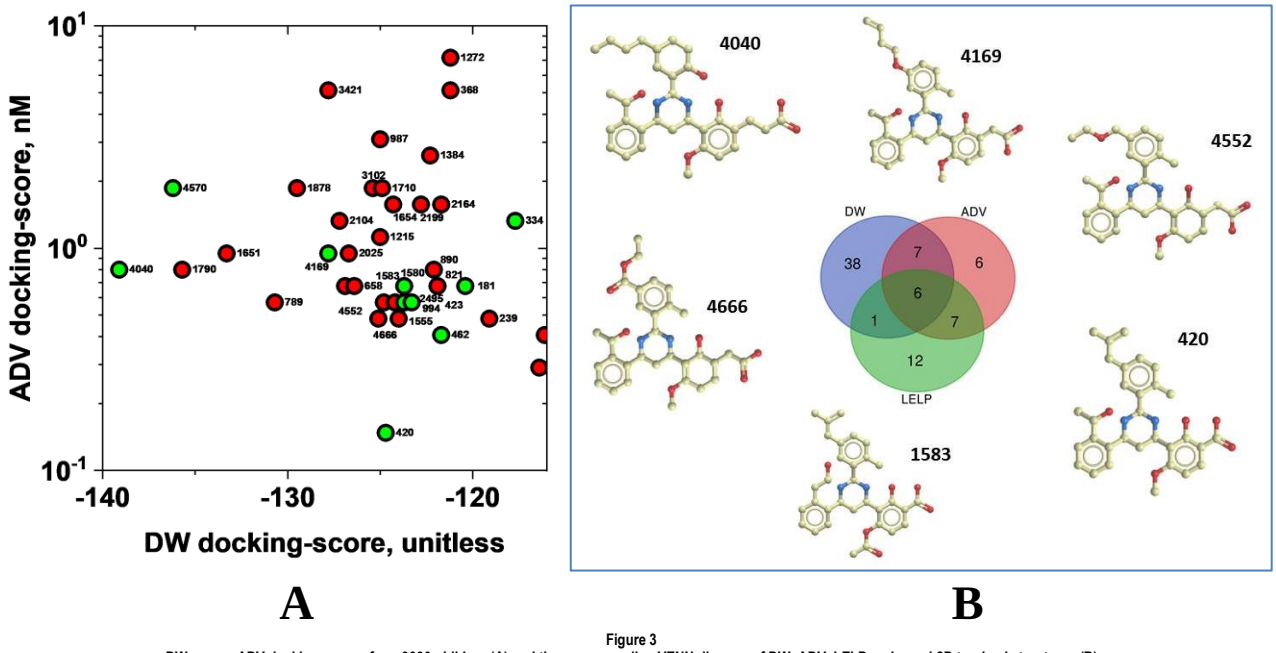

Figure 3
DW *versus* ADV docking-scores from 9990-children (A) and the corresponding VENN diagram of DW+ADV+LELP ranks and 2D top-lead structures (B)
The 9990-child evolved from TTT were DW re-evolved using 9990 as parent (**Figure 2, green**).
The newly generated 9990-children were ADV docked (using a 50x50x50 Å grid surrounding the 690-1010 target of the Omicron inner-cavity) and LELP normalized.
A) DW *versus* ADV leads comparison. Red + green circles, selected top-leads between -140 to -115 DW and 10^{-1} to 10 nM ADV docking-scores. **Green** circles, top-leads predicting the lowest LELP values between -2.2 to 2.2. **Numbers** beside circles corresponded to each of the children according to the DW evolutionary generation order.
B) VENN diagram and 2D structures of 9990-children, **Blue VENN**, DW. **Red VENN**, ADV. **Green VENN**, LELP. The 2D structures corresponded to the 6 top-lead 9990-children common to DW, ADV and LELP: **Red spheres**, Oxygens. **Blue spheres**, Nitrogens. **Yellow spheres**, Carbons.
Further details of all the 107 children evolved from the 9990-parent, including DW, ADV and LELP results, are described in supplementary **Table S3**

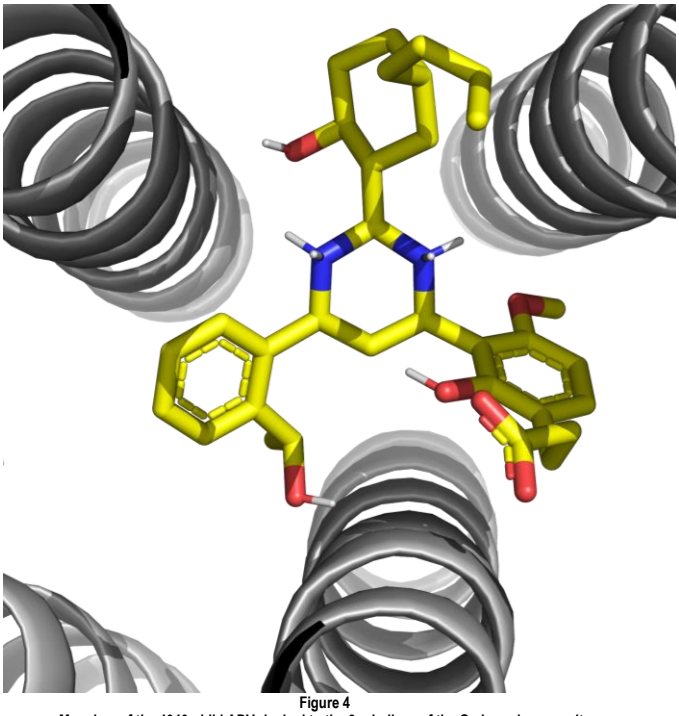

Figure 4
Mapping of the 4040-child ADV-docked to the 3 α-helices of the Omicron inner-cavity

Discussion

Minimal docking-scores consecutively obtained from sS, DW and ADV programs have been used to screen the many parent-like children molecules generated from Ss previously-identified star-shaped molecules by DW evolutionary libraries fitting the inner-cavity of prefusion states of Omicron variants. Rank consensus between the predictions of the programs have identify common top-leads among DW (two consecutive parent runs), and ADV-LELP-normalized ranks using VENN diagrams.

A high number of parent-like star-shaped children molecules, estimated between 10000-30000 molecules per parent, were randomly generated by DW evolutionary fitting the Omicron inner-cavity target defined at earliest SARS-Cov-2 isolates by previous sS. Such enriched numbers of randomly generated molecules were automatically selected for best fitting to the Omicron inner-cavity, and additional criteria including molecular weight and hydrophobicity. The numbers of potentially useful children molecules was downsized by excluding those with DW-predicted high and low toxic properties and nasty functions (~20-74% of the children molecules). Two top-leaders from these children-molecules were selected for a second round of evolutionary fitting. For that, the top-leads chosen to represent two different chemotypes containing one central core of 6-atom rings with either two Nitrogens (9990-child) or one Nitrogen and one Oxygen (4976-child), were re-used as parents to generate hundreds of secondary children. Although, despite using this evolutionary powerful approach, most of the corresponding vast chemical/chemotype space[78, 79] fitting the Omicron inner-cavity still remains to be explored, a similar large number of star-shaped parent-like molecules would have been impossible to retrieve throughout screening of any of the existent public chemical libraries (Mcule, chemSpace, Zinc, PubChem, Chembl, etc).

Since new predictions from a unique algorithm such as DW could be limited by their restricted docking-cavity, wider grid-dependent ADV docking together with LELP normalization specificity were employed to improve prediction accuracies. Ranks rather than absolute docking-score comparisons were selected to derive consensus among DW, ADV and LELP leads [53, 69-72]. To our knowledge, the VENN diagrams used here for the first time, required trial-and-error attempts to optimize the numbers of leads from each program to identify common ranked leads. In this particular case, selecting 24-48 % of the top-ranks were required to propose 6 top-leads. The 6 DW, ADV and LELP top-leads predicted a similar chemotype, despite the high number of possibilities explored as discussed above, while the best top-lead was the 4040-child. Thus, the 4040-child reduced the LELP value (increased specificity) from 4.2 to 1.6 compared to its TTT-parent, despite increasing its molecular weight (363 to 554 Dalton). Additionally, the 4040-child, lowered the cLogP increasing water solubility (from 3.1 to 0.5), increased the DW affinity (-92 to -139), and maintained the ADV affinity (0.7 to 0.8 nM) (**Table S1**). Maximal numbers of Hydrogen and hydrophobic ADV interactions with the amino acids of the Omicron inner-cavity were also among the improvements of the 4040-child (**Table S2**).

Known limitations of the predictions mentioned above, include for example, fixed docking-cavities, eliminated water molecules, chemotype-dependent unreliabilities, and docking discrepancies due to force-field energy minimization algorithms (despite selecting mmff94s and mmff94s+ force-fields to best maintain molecular geometries after docking). Additionally, any exploration of the vast chemical/chemotype space, including the ones reported here, will always be incomplete. Some of the limitations together with unexpected *in vivo* physiological variables, may still limit the accuracy of these predictions.

The results described here, identified one new chemotype predicting exceptionally low docking-score ranges with high specificity while conserving its targeting to the inner-cavity of the trimer α-helices of Omicron variants. The results remain to be confirmed after chemical synthesis (perhaps similar to those recently reported [80] would allow experimental validation.

Supporting information

Table S1
Resume of molecular properties of parents and top-lead children chemotypes

ID	number children	non toxic	MW	cLogP	ADV Kcal/mol	ADV nM	DW score	LELP
TTT- parent	897	187	363	3.1	-12.5	0.7	-92	4.2
4976-parent	337	250	503	1.4	-13.3	0.2	-121	3.8
9990-parent	353	107	490	0.2	-12.9	0.3	-116	0.7
4040-child			554	0.5	-12.4	0.8	-139	1.6
4169-child			546	0.1	-12.3	0.9	-127	0.4
1583-child			544	0.9	-12.6	0.5	-125	2.8
4552-child			538	-1.1	-12.5	0.6	-127	-3.6
420-child			516	0.7	-13.4	0.1	-124	2.0
4666-child			552	-1.1	-12.7	0.4	-125	-3.4

The TTT-children were generated from the TTT parent (**Figure 2**). The 4076- and 9990 TTT-children top-lead chemotypes were chosen as parents for additional evolutions. Six 9990-children were selected as top-leads (lower part of the Table). Each of the child numbers were automatically assigned by the DW generation order during the 9990 evolutionary docking.
Added note. Targeting the fusion-competent one-up Omicron 7to4 model inner-cavity [13], increased to 150-490 nM (500-1000-fold) the corresponding ADV binding-scores predicted and the docked top-lead complexes were displaced from the center to one of the sides of the inner-cavity.

Table S2
Amino acid residues of Omicron predicting interactions with TTT-parent and children leads

	position	Aa	Aa	TTT	4976	9990	4040	4169	1583	4552	420	4666
A	756	Y	Tyr					H				H
	759	F	Phe									
	970	F	Phe									
	995	R	Arg									
	998	T	Thr									
	999	G	Gly									
	1001	L	Leu									
	1002	N	Asn		H				H			
B	756	Y	Tyr									
	759	F	Phe									
	970	F	Phe									
	995	R	Arg				H					
	998	T	Thr				H			H		H
	999	G	Gly									
	1001	L	Leu									
	1002	N	Asn				H					
	1005	N	Gln									
C	756	Y	Tyr									
	759	F	Phe									
	970	F	Phe									
	995	R	Arg									
	998	T	Thr		H	H		H				
	999	G	Gly									
	1001	L	Leu									
	1002	N	Asn									

Aa, Amino acid residues of the inner-cavity made of 3 α-helices (A,B,C) of the S spike of Omicron implicated in ADV dockings to common star-shaped top-leads selected from the VENN analysis.
Colored rectangles, amino acid residues predicted as ligands to the top-leads by LigPlus
H, predicted Hydrogen bonds by LigPlus.
Yellow rectangles, TTT- parent.
Blue rectangles, 4976- child from TTT used as second parent .
Green rectangles, 9990-child from TTT used as second parent .
Light green rectangles, 9990-children top-leads

Table S3
Table S3- 107 children from 9990.sdf
Table containing the properties of all non-toxic children derived from the 9990, including 2D graphs, Molecular Name(Numbers), generation, DW docking-score, Molecular weight, ADV Vina Binding Affinity in Kcal/mol and nM, and LE, LLE and LELP ligand efficiency values derived from ADV nM.

Funding
The work was carried out without any external financial contribution

Competing interests
The author declares no competing interests

Authors' contributions
JC designed, performed and analyzed the docking data and drafted the manuscript.

Acknowledgements
Thanks are specially due to Dr. T.Sander and J.Wahl of Idorsia Pharmaceuticals Ltd. at Allschwil (Switzerland) for their help during the first steps on the windows version into the Data Warrior Build Evolutionary Library. Thanks are also due to Dr. A. Villena from the University of Leon (Spain) for advise.

References

[1]Gobeil, S.M., *et al.* Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity.*Science.* 2021, 373: http://dx.doi.org/10.1126/science.abi6226.

[2]Hong, Q., *et al.* Molecular basis of receptor binding and antibody neutralization of Omicron.*Nature.* 2022, 604: 546-552 http://dx.doi.org/10.1038/s41586-022-04581-9.

[3]Duerr, R., *et al.* Delta-Omicron recombinant escapes therapeutic antibody neutralization.*bioRxiv.* 2022: http://dx.doi.org/10.1101/2022.04.06.487325.

[4]Braet, S.M., *et al.* Timeline of changes in spike conformational dynamics in emergent SARS-CoV-2 variants reveal progressive stabilization of trimer stalk with altered NTD dynamics.*Elife.* 2023, 12: http://dx.doi.org/10.7554/eLife.82584.

[5]Callaway, E. Coronavirus variant XBB.1.5 rises in the United States - is it a global threat?*Nature.* 2023, 613: 222-223 http://dx.doi.org/10.1038/d41586-023-00014-3.

[6]Zhao, L.P., *et al.* Rapidly identifying new coronavirus mutations of potential concern in the Omicron variant using an unsupervised learning strategy.*Sci Rep.* 2022, 12: 19089 http://dx.doi.org/10.1038/s41598-022-23342-2.

[7]Kimura, I., *et al.* Virological characteristics of the SARS-CoV-2 Omicron BA.2 subvariants, including BA.4 and BA.5.*Cell.* 2022, 185: 3992-4007 e16 http://dx.doi.org/10.1016/j.cell.2022.09.018.

[8]Saito, A., *et al.* Virological characteristics of the SARS-CoV-2 Omicron BA.2.75 variant.*Cell Host Microbe.* 2022, 30: 1540-1555 e15 http://dx.doi.org/10.1016/j.chom.2022.10.003.

[9]Cao, Y., *et al.* BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection.*Nature.* 2022, 608: 593-602 http://dx.doi.org/10.1038/s41586-022-04980-y.

[10]Stalls, V., *et al.* Cryo-EM structures of SARS-CoV-2 Omicron BA.2 spike.*Cell Rep.* 2022, 39: 111009 http://dx.doi.org/10.1016/j.celrep.2022.111009.

[11]Focosi, D., *et al.* The Omicron variant of concern: Diversification and convergent evolution in spike protein, and escape from anti-Spike monoclonal antibodies.*Drug Resist Updat.* 2022, 65: 100882 http://dx.doi.org/10.1016/j.drup.2022.100882.

[12]Cao, W., *et al.* Biomechanical characterization of SARS-CoV-2 spike RBD and human ACE2 protein-protein interaction.*Biophys J.* 2021, 120: 1011-1019 http://dx.doi.org/10.1016/j.bpj.2021.02.007.

[13]Zhang, J., *et al.* Structural and functional impact by SARS-CoV-2 Omicron spike mutations.*Cell Rep.* 2022, 39: 110729 http://dx.doi.org/10.1016/j.celrep.2022.110729.

[14]Gobeil, S.M., *et al.* Structural diversity of the SARS-CoV-2 Omicron spike.*Mol Cell.* 2022, 82: 2050-2068 e6 http://dx.doi.org/10.1016/j.molcel.2022.03.028.

[15]Ebrahimi, M., *et al.* Computational repurposing approach for targeting the critical spike mutations in B.1.617.2 (delta), AY.1 (delta plus) and C.37 (lambda) SARS-CoV-2 variants using exhaustive structure-based virtual screening, molecular dynamic simulations and MM-PBSA methods.*Comput Biol Med.* 2022, 147: 105709 http://dx.doi.org/10.1016/j.compbiomed.2022.105709.

[16]Henderson, R., *et al.* Controlling the SARS-CoV-2 Spike Glycoprotein Conformation.*bioRxiv.* 2020: http://dx.doi.org/10.1101/2020.05.18.102087.

[17]Makela, A.R., *et al.* Intranasal trimeric sherpabody inhibits SARS-CoV-2 including recent immunoevasive Omicron subvariants.*Nat Commun.* 2023, 14: 1637 http://dx.doi.org/10.1038/s41467-023-37290-6.

[18]Pirolli, D., *et al.* Virtual screening and molecular dynamics simulations provide insight into repurposing drugs against SARS-CoV-2 variants Spike protein/ACE2 interface.*Sci Rep.* 2023, 13: 1494 http://dx.doi.org/10.1038/s41598-023-28716-8.

[19]Pirolli, D., *et al.* Targeting SARS-CoV-2 Spike Protein/ACE2 Protein-Protein Interactions: a Computational Study.*Mol Inform.* 2021, 40: e2060080 http://dx.doi.org/10.1002/minf.202060080.

[20]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.

[21]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.

[22]Yang, K., *et al.* Structural conservation among variants of the SARS-CoV-2 spike postfusion bundle.*Proc Natl Acad Sci U S A.* 2022, 119: e2119467119 http://dx.doi.org/10.1073/pnas.2119467119.

[23]Hu, Y., *et al.* Design and characterization of novel SARS-CoV-2 fusion inhibitors with N-terminally extended HR2 peptides.*Antiviral Res.* 2023, 212: 105571 http://dx.doi.org/10.1016/j.antiviral.2023.105571.

[24]Kalathiya, U., *et al.* Highly Conserved Homotrimer Cavity Formed by the SARS-CoV-2 Spike Glycoprotein: A Novel Binding Site.*J Clin Med.* 2020, 9: http://dx.doi.org/10.3390/jcm9051473.

[25]Hsieh, C.L., *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes.*Science.* 2020: science.abd0826 [pii], http://dx.doi.org/10.1126/science.abd0826.

[26]Xiong, X., *et al.* A thermostable, closed SARS-CoV-2 spike protein trimer.*Nat Struct Mol Biol.* 2020: http://dx.doi.org/10.1038/s41594-020-0478-5, 10.1038/s41594-020-0478-5 [pii].

[27]Carr, C.M. and Kim, P.S. A spring-loaded mechanism for the conformational change of influenza hemagglutinin.*Cell.* 1993, 73: 823-832 http://dx.doi.org/10.1016/0092-8674(93)90260-W.

[28]Pallesen, J., *et al.* Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen.*Proc Natl Acad Sci U S A.* 2017, 114: E7348-E7357 1707304114 [pii], http://dx.doi.org/10.1073/pnas.1707304114.

[29]Kandeel, M., *et al.* Repurposing of FDA-approved antivirals, antibiotics, anthelmintics, antioxidants, and cell protectives against SARS-CoV-2 papain-like protease.*J Biomol Struct Dyn.* 2020: 1-8 http://dx.doi.org/10.1080/07391102.2020.1784291.

[30]Kandeel, M. and Al-Nazawi, M. Virtual screening and repurposing of FDA approved drugs against COVID-19 main protease.*Life Sci.* 2020, 251: 117627 http://dx.doi.org/10.1016/j.lfs.2020.117627.

[31]Bakowski, M.A., *et al.* Drug repurposing screens identify chemical entities for the development of COVID-19 interventions.*Nature Communications.* 2021, 12: 3309-3323 http://dx.doi.org/10.1038/s41467-021-23328-0.

[32]Xia, S., *et al.* A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike.*Sci Adv.* 2019, 5: eaav4580 http://dx.doi.org/10.1126/sciadv.aav4580, aav4580 [pii].

[33]Xia, S., *et al.* Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein.*Cell Mol Immunol.* 2020, 17: 765-767 http://dx.doi.org/10.1038/s41423-020-0374-2, 10.1038/s41423-020-0374-2 [pii].

[34]Cannalire, R., *et al.* SARS-CoV-2 Entry Inhibitors: Small Molecules and Peptides Targeting Virus or Host Cells.*Int J Mol Sci.* 2020, 21: ijms21165707 [pii], http://dx.doi.org/10.3390/ijms21165707.

[35]Tang, T., *et al.* Coronavirus membrane fusion mechanism offers a potential target for antiviral development.*Antiviral Res.* 2020, 178: 104792 S0166-3542(20)30206-0 [pii], http://dx.doi.org/10.1016/j.antiviral.2020.104792.

[36]Wu, C., *et al.* Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods.*Acta Pharm Sin B.* 2020: http://dx.doi.org/10.1016/j.apsb.2020.02.008, S2211-3835(20)30299-9 [pii].

[37]Ruan, Z., *et al.* SARS-CoV-2 and SARS-CoV: Virtual Screening of Potential inhibitors targeting RNA-dependent RNA polymerase activity (NSP12).*J Med Virol.* 2020: http://dx.doi.org/10.1002/jmv.26222.

[38]Tsuji, M. Potential anti-SARS-CoV-2 drug candidates identified through virtual screening of the ChEMBL database for compounds that target the main coronavirus protease.*FEBS Open Bio.* 2020, 10: 995-1004 http://dx.doi.org/10.1002/2211-5463.12875.

[39]de Souza Neto, L.R., *et al.* In silico Strategies to Support Fragment-to-Lead Optimization in Drug Discovery.*Front Chem.* 2020, 8: 93 http://dx.doi.org/10.3389/fchem.2020.00093.

[40]Liu, T., *et al.* Break Down in Order To Build Up: Decomposing Small Molecules for Fragment-Based Drug Design with eMolFrag.*J Chem Inf Model.* 2017, 57: 627-631 http://dx.doi.org/10.1021/acs.jcim.6b00596.

[41]Patel, H., *et al.* SAVI, in silico generation of billions of easily synthesizable compounds through expert-system type rules.*Sci Data.* 2020, 7: 384 http://dx.doi.org/10.1038/s41597-020-00727-4.

[42]Polishchuk, P. CReM: chemically reasonable mutations framework for structure generation.*J Cheminform.* 2020, 12: 28 http://dx.doi.org/10.1186/s13321-020-00431-w.

[43]Jahnke, W., *et al.* Fragment-to-Lead Medicinal Chemistry Publications in 2019.*J Med Chem.* 2020, 63: 15494-15507 http://dx.doi.org/10.1021/acs.jmedchem.0c01608.

[44]Rifaioglu, A.S., *et al.* DEEPScreen: high performance drug-target interaction prediction with convolutional neural networks using 2-D structural compound representations.*Chem Sci.* 2020, 11: 2531-2557 http://dx.doi.org/10.1039/c9sc03414e.

[45]Yang, X., *et al.* ChemTS: an efficient python library for de novo molecular generation.*Sci Technol Adv Mater.* 2017, 18: 972-976 http://dx.doi.org/10.1080/14686996.2017.1401424.

[46]Spiegel, J.O. and Durrant, J.D. AutoGrow4: an open-source genetic algorithm for de novo drug design and lead optimization.*J Cheminform.* 2020, 12: 25 http://dx.doi.org/10.1186/s13321-020-00429-4.

[47]Chevillard, F., *et al.* Interrogating dense ligand chemical space with a forward-synthetic library.*Proc Natl Acad Sci U S A.* 2019, 116: 11496-11501 http://dx.doi.org/10.1073/pnas.1818718116.

[48]Volochnyuk, D.M., *et al.* Evolution of commercially available compounds for HTS.*Drug Discov Today.* 2019, 24: 390-402 S1359-6446(18)30242-3 [pii], http://dx.doi.org/10.1016/j.drudis.2018.10.016.

[49]Malebari, A.M., *et al.* The Antiproliferative and Apoptotic Effect of a Novel Synthesized S-Triazine Dipeptide Series, and Toxicity Screening in Zebrafish Embryos.*Molecules.* 2021, 26: http://dx.doi.org/10.3390/molecules26041170.

[50]El-Faham, A., *et al.* Di- and tri-substituted s-triazine derivatives: Synthesis, characterization, anticancer activity in human breast-cancer cell lines, and developmental toxicity in zebrafish embryos.*Bioorg Chem.* 2020, 94: 103397 http://dx.doi.org/10.1016/j.bioorg.2019.103397.

[51]Semren, T.Z., *et al.* Oxidative stress in triazine pesticide toxicity: a review of the main biomarker findings.*Arh Hig Rada Toksikol.* 2018, 69: 109-125 http://dx.doi.org/10.2478/aiht-2018-69-3118.

[52]Saka, M., *et al.* Chronic toxicity of 1,3,5-triazine herbicides in the postembryonic development of the western clawed frog Silurana tropicalis.*Ecotoxicol Environ Saf.* 2018, 147: 373-381 http://dx.doi.org/10.1016/j.ecoenv.2017.08.063.

[53]Coll, J.M. Evolutionary-docking targeting bacterial FtsZ.*ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/6405c36fcc600523a3bcb679:

[54]Palacio-Rodriguez, K., *et al.* Exponential consensus ranking improves the outcome in docking and receptor ensemble docking.*Sci Rep.* 2019, 9: 5142 http://dx.doi.org/10.1038/s41598-019-41594-3.

[55]Scardino, V., *et al.* Combination of pose and rank consensus in docking-based virtual screening: the best of both worlds.*RSC Adv.* 2021, 11: 35383-35391 http://dx.doi.org/10.1039/d1ra05785e.

[56]Plewczynski, D., *et al.* Can we trust docking results? Evaluation of seven commonly used programs on PDBbind database.*J Comput Chem.* 2011, 32: 742-55 http://dx.doi.org/10.1002/jcc.21643.

[57]Plewczynski, D., *et al.* VoteDock: consensus docking method for prediction of protein-ligand interactions.*J Comput Chem.* 2011, 32: 568-81 http://dx.doi.org/10.1002/jcc.21642.

[58]Blanes-Mira, C., *et al.* Comprehensive Survey of Consensus Docking for High-Throughput Virtual Screening.*Molecules.* 2022, 28: http://dx.doi.org/10.3390/molecules28010175.

[59]Ochoa, R., *et al.* dockECR: Open consensus docking and ranking protocol for virtual screening of small molecules.*J Mol Graph Model.* 2021, 109: 108023 http://dx.doi.org/10.1016/j.jmgm.2021.108023.

[60]Rosignoli, S. and Paiardini, A. DockingPie: a consensus docking plugin for PyMOL.*Bioinformatics.* 2022, 38: 4233-4234 http://dx.doi.org/10.1093/bioinformatics/btac452.

[61]Hopkins, A.L., *et al.* Ligand efficiency: a useful metric for lead selection.*Drug Discov Today.* 2004, 9: 430-1 http://dx.doi.org/10.1016/S1359-6446(04)03069-7.

[62]Cavalluzzi, M.M., *et al.* Ligand efficiency metrics in drug discovery: the pros and cons from a practical perspective.*Expert Opin Drug Discov.* 2017, 12: 1087-1104 http://dx.doi.org/10.1080/17460441.2017.1365056.

[63]Hopkins, A.L., *et al.* The role of ligand efficiency metrics in drug discovery.*Nat Rev Drug Discov.* 2014, 13: 105-21 http://dx.doi.org/10.1038/nrd4163.

[64]Murray, C.W., *et al.* Validity of ligand efficiency metrics.*ACS Med Chem Lett.* 2014, 5: 616-8 http://dx.doi.org/10.1021/ml500146d.

[65]Kenny, P.W. The nature of ligand efficiency.*J Cheminform.* 2019, 11: 8 http://dx.doi.org/10.1186/s13321-019-0330-2.

[66]Garcia-Sosa, A.T., *et al.* Design of multi-binding-site inhibitors, ligand efficiency, and consensus screening of avian influenza H5N1 wild-type neuraminidase and of the oseltamivir-resistant H274Y variant.*J Chem Inf Model.* 2008, 48: 2074-80 http://dx.doi.org/10.1021/ci800242z.

[67]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.

[68]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.

[69]Lewis-Atwell, T., *et al.* Comparing the Performances of Force Fields in Conformational Searching of Hydrogen-Bond-Donating Catalysts.*J Org Chem.* 2022, 87: 5703-5712 http://dx.doi.org/10.1021/acs.joc.2c00066.

[70]Loeffler, H.H., *et al.* Reproducibility of Free Energy Calculations across Different Molecular Simulation Software Packages.*J Chem Theory Comput.* 2018, 14: 5567-5582 http://dx.doi.org/10.1021/acs.jctc.8b00544.

[71]Wang, Y., *et al.* End-to-end differentiable construction of molecular mechanics force fields.*Chem Sci.* 2022, 13: 12016-12033 http://dx.doi.org/10.1039/d2sc02739a.

[72]Mobley, D.L., *et al.* Escaping Atom Types in Force Fields Using Direct Chemical Perception.*J Chem Theory Comput.* 2018, 14: 6076-6092 http://dx.doi.org/10.1021/acs.jctc.8b00640.

[73]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.

[74]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.

[75]Sander, T., *et al.* DataWarrior: an open-source program for chemistry aware data visualization and analysis.*J Chem Inf Model.* 2015, 55: 460-73 http://dx.doi.org/10.1021/ci500588j.

[76]Zhao, Z., *et al.* Omicron SARS-CoV-2 mutations stabilize spike up-RBD conformation and lead to a non-RBM-binding monoclonal antibody escape.*Nat Commun.* 2022, 13: 4958 http://dx.doi.org/10.1038/s41467-022-32665-7.

[77]Keseru, G.M. and Makara, G.M. The influence of lead discovery strategies on the properties of drug candidates.*Nat Rev Drug Discov.* 2009, 8: 203-12 http://dx.doi.org/10.1038/nrd2796.

[78]Polishchuk, P.G., *et al.* Estimation of the size of drug-like chemical space based on GDB-17 data.*J Comput Aided Mol Des.* 2013, 27: 675-9 http://dx.doi.org/10.1007/s10822-013-9672-4.

[79]Mroz, A.M., *et al.* Into the Unknown: How Computation Can Help Explore Uncharted Material Space.*J Am Chem Soc.* 2022, 144: 18730-18743 http://dx.doi.org/10.1021/jacs.2c06833.

[80]Azzouzi, M., *et al.* Design, synthesis, and computational studies of novel imidazo[1,2-a]pyrimidine derivatives as potential dual inhibitors of hACE2 and spike protein for blocking SARS-CoV-2 cell entry.*Journal of Molecular Structure.* 2023, 1285: 135525 http://dx.doi.org/10.1016/j.molstruc.2023.135525.