Exploring non-toxic co-evolutionary docking

short title: co-evolutionary docking

Coll, J*.
Biotechnology Department. 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

Drug-spaces of nine crystallographic protein / ligand models have been comparatively explored by including *Toxicity Risk* assessment during computational co-evolution. Tens of thousands children were randomly generated from parent ligands and iteratively selected for higher affinities, increased specificities and low *Toxicity Risk* using *DataWarrior* / *Build Evolutionary Library* algorithms, mimicking natural evolution. Only a few hours of co-evolution increased ~ 2-fold the numbers of non-toxic children. Top-leads predicted drug-like properties, nanoMolar affinities (confirmed by AutoDockVina), higher specificities, absence of known toxicities, and similar docking to their initial binding cavities. Tables were provided with multi-threshold-adjustable filters for alternative *in silico* explorations of this new "co-evolutionary docking" tool.

Keywords: co-evolutionary docking; novel antibiotics; anticancerigens; antidiabetes; rodenticides; heart diseases; antifibrosis

Introduction

***D**ata**W**arrior* (*DW*) ***B**uild **E**volutionary **L**ibrary* (*BEL*) co-evolution algorithms including for the first time *Toxicity Risk* assessment are described here targeting nine protein / ligand crystallographic pairs.

As recently reported [1-4], these *DW* tools mimicked natural evolution, by generating thousands of children molecules fitting their crystallographic binding-cavities. However, ~ 50 % of the generated children predicted *DW* / *Toxicity*[1-4]. More in detail, these co-evolutionary fast algorithms randomly generated tens of thousands of small molecular raw children from parent ligand molecules, rather than screening for a few hundreds in large chemical banks. Using *DW-BEL*, the abundant raw children were then rapidly ranked by best fitting both to their binding cavity and to other co-evolving criteria such as molecular weight and hydrophobicity [1-4]. Best fitted thousands of evolved children could be generated in few hours for each protein / ligand pair. In collaboration with *DW* researchers, we explored here the inclusion of *Toxicity Risk* during *DW-BEL* co-evolution to evaluate any possible improvements on the resulting children using nine protein / ligand crystallographic pairs as selected examples. To further fine-tune their accuracies, the *DW-BEL* top-lead docking-affinities were confirmed by the widely known **A**uto**D**ock**V**ina (ADV) algorithms.

Only the high speed of Java's *DW-BEL* favored the short time co-evolution of several criteria by one-by-one molecular checking tens of thousands of raw children. Only those children predicting maximal fittings were selected for the next iteration cycle. To avoid structure repetitions among the selected children, 10-50 Gb of RAM memory were required to keep track of all the numerous children generated, depending on the particular target / parent, number of runs, cycles, generations, etc. Iterations were automatically repeated from each parent during a number of cycles until their fitness to the user-defined criteria reached a plateau. Because of their stochastic nature, independent runs should be consecutively repeated from the same parent to increase prediction accuracies.

Because any computational docking screening method based solely in maximizing affinities, generates highly unspecific molecules by progressively increasing their molecular weights and hydrophobicities[5, 6], ligand efficiency parameters have been included for filtering traditional screening results. For instance, **L**igand **E**fficiency (LE), **L**igand **E**fficiency **L**ipophilic **P**rize (LELP) and many others have been previously proposed [7-10]. Our previous work[1-4] already introduced alternative co-evolution with low molecular weight / hydrophobicity criteria during *DW-BEL* to reduce such trends[1-4]. Therefore, such criteria have been also employed in the present work for similar purposes.

Because the accuracy of any individual docking programs to predict real affinities and conformations still remains challenging, consensus docking using a minimum of two different algorithms has been recommended by several authors[11-17]. Therefore, the widely employed **A**uto**D**ock**V**ina (ADV), which relies on a completely different docking algorithm than *DW*, was chosen before to be included in our previous work[1-4]. Additionally, ADV provided ~ nM Kd to quantify affinity estimations. Such estimation of affinities with consensus purposes have been also included here.

The addition of *Toxicity Risk* during the *DW-BEL* co-evolution was suggested by detection of ~ 50 % toxicities on the predicted children, during our previous work targeting Vkorc1, FtsZ, LolCE and omicron S [1-4]. Since the inclusion of *Toxicity Risk* could theoretically increase the percentage of non-toxic children and/or improve the penetration into unexplored chemical spaces, it was suggested to the *DW* researcher forum. A previously developed *DW* toxicity assessment code was then kindly included by Dr.T. Sander as a new *DW-BEL Toxicity Risk* criteria with its values and weights. To comparatively explore co-evolution with *DW-BEL* ± *Toxicity Risk*, nine protein / ligand crystallographic pairs were selected here because of their practical importance.

Many of the selected pairs (**Table 1**) required either higher specificities to reduce their physiological or ecological off-targets or improved miliMolar to nanoMolar affinities for other physiological / delivery reasons. In any of the selected pairs, a large number of candidates for alternative ligands need to be generated to select the most appropriated. Because it was not practical to experimentally or computationally screen for such large numbers of candidates[11, 12], the above commented *DW-BEL* tools may supplied significant alternatives. Combining the recent availability of newly crystallized protein / ligand models, the improvements in 3D protein modeling by alphafold algorithms[13], and the above mentioned *DW-BEL + Toxicity Risk* novel co-evolution, large numbers of parent-derived children candidates were generated for the following protein targets:

Vkorc1. The **V**itamin **K** ep**O**xide **R**eductase **C**omplex **1** [14] binds/oxidizes reduced Vitamin K to recycle animal coagulation. Vkorc1 has been targeted for coagulation control in humans and for anticoagulant rodenticides in rats[15, 16, 17, 18, 19]. However, rodent genetic resistances and off-target unspecific ecological effects, remain as main concerns. Maximal specificity for on-target rodent lethality and minimal off-target for ecological impacts[12] are desirable for new rodenticides**.** Children predicting nanoMolar affinities for wild type and resistant rat Vkorc1 and lowest affinities for human Vkorc1, could be generated from the brodifacoum parent by *DW-BEL* in our previous work[3]. However, *Toxicity Risk* was not yet available.

FtsZ. The **F**ilament **T**emperature **S**ensitive **Z**-ring proteins of bacterial cell division (i.e., *Staphylococcus aureus*) are important targets to develop novel antibiotics against increasing resistances[20, 21, 22-25]. The reference anti-staphylococcal PC190723 inhibitor binds FtsZ but at low µM ranges [26, 27, 28-30], often causing pharmacological problems and bacterial gene resistances[24, 31]. TX-derived[32] drugs/prodrugs (Taxis pharmaceuticals, Monmouth Junction, NJ, USA)[33-38] are being developed to overcome such issues, but no affinity improvements have been yet reported[21]. Previous attempts to find new FtsZ docking candidates by large computational screening, identified ~ 100 new inhibitors of bacterial cell division but only at µM affinities[39-41]. PC190723-derived *DW-BEL* children at nanoMolar affinities were described in our previous work [2], including some validations by bacterial eNTRY globularity rules for drug bacterial cell wall penetration[36-38]. *Toxicity Risk* was not available.

Sglt. The **S**odium–**Gl**ucose co-**T**ransporters 1 and 2 are sodium-coupled transport proteins that facilitate glucose food absorption (Sglt1) and glucose blood reabsorption (Sglt2)[42-44].Inhibition of Sglt1 (reference ligand LX2761) [43] blocks intestinal glucose absorption influencing cardiovascular and other diseases. Inhibition of Sglt2 (reference ligand Empaglifozin) prevents kidney reabsorption eliminating excess of blood glucose through the urine, facilitating diabetes control[44]. Sglt are similar proteins displaying 14 transmembrane helices, each having particular molecular properties and physiological targets. The reference ligands are potent binders targeting Sglt 1 and/or 2 with nM affinities. Computational targeting has been recently reported[43-47]. Exploration for alternative ligands may contribute physiological fine-tuning and/or reduction of off-target effects. No Sglt have been targeted by *DW-BEL*.

Vegfr2. The **V**ascular **E**ndothelial **G**rowth **F**actor **2** is a receptor of tyrosine kinases that have been implicated in tumor angiogenesis, a key feature of many cancers[48]. However, the number of drugs for cancer treatments targeting Vegfr2 are limited[49]. Sorafenib is the main reference ligand as an oral multi-target kinase inhibitor of tumor genesis and angiogenesis. In particular cases, Sorafenib binds Vegfr2 and prolonged survival times of chemotherapy-resistant patients. Although some docking explorations have been published[50-53], no Vegfr2 was yet targeted by *DW-BEL*.

Nkcc1. The **N**a$^+$**K**$^+$**C**l$^-$ **C**otransporter **1** is a Sodium, Potassium, Chloro co-transporter implicated in salt-hypertension, kidney reabsorption and neuronal

excitability[54-57] . Bumetanide has been used as a reference inhibitor but only at µM affinities. No Nkcc1 have been targeted by *DW-BEL*.

Hsp47. The **H**eat **S**hock **P**rotein **47** targets collagen folding and function. Hsp47 excess caused tissue fibrosis and/or cancer. The Hsp47 main binding site include some molecular species of collagen[58-61] . Although some Hsp47 inhibitors (i.e., Hs55) supressed excesive collagen synthesis in *in vitro* models[61] they were only active at µM affinities. No Hsp47 have been targeted by *DW-BEL*.

P2x3r. The **P2x3 R**eceptor is an ATP-gated cation channel that has been implicated in pulmonary fibrosis, rheumatoid arthritis, pain perception, and synaptic transmissions, among many other diseases[62, 63] . The limited success in drug development targeting these type of receptors may be due to the difficulties to compete with their ATP binding-sites [64, 65] . Additional allosteric sites may be possible alternative therapeutic targets (reference ligand AF-219). No P2x3r have been targeted by *DW-BEL*.

LolEC. The **L**ipoprotein **O**uter membrane **L**ocalization **EC** protein complexes control traffic of bacterial lipoproteins. Deep-learning screening based on experimental inhibitory training-model developed from data on the Drug Repurposing Hub library[66] , recently discovered Abaucin[67] . Abaucin, targeted LolEC, showed anti-*A.baumannii* activities at low µM ranges [68-74] and suppressed mice infection. *A. baumannii* is a Gram-negative multi-drug resistant bacteria causing health care-associated infections world-wide[75-77] with high mortality rates[78,69] *A. baumannii* needs new drug discoveries[79-82] possibly targeting the E protein as in *E.coli*[83] . Higher affinities may help further improved Abaucin-derivatives to reduce off-target physiological problems. Alphafold *E.coli* modeling of *A. baumannii* LolCE was employed to generate *DW-BEL* Abaucin-children in our previous work[1] but *Toxicity Risk* was not available.

Computational Methods

Protein / ligand pair models

Whenever possible the protein/ligand pdb files were downloaded from crystallographic 3D models from the RCSB-PDB bank (**R**esearch **C**ollaboratory for **S**tructural **B**ioinformatics-**P**rotein **D**ata **B**ank) corresponding solved structures in complex with their reference ligand (**Table 1**). The complex pdb corresponding to rat Vkorc1[3] and *A.baumannii* LolEC[1] were alphafold modeled from their human and *E.coli* targets, respectively**.** Alphafold modeling were performed by Sokrypton Colab Alphafold ipynb (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb)[13] .

Table 1
Target / parent pairs selected for performance tests of *DW-BEL ± Toxicity Risk*

Protein	Original specie	RCSB, pdb	Ref.	Alphafold modelled	Ligand	Affinity, ~ nM	Ref*
Vkorc1	human	6WVH	[14]	Rat	Brodifacoum	100	[3]
FtsZ	*S.aureus*	4DXD	[84, 85]		PC190723	10000	[2]
Sglt1	human	7WMV	[43]		LX1761	2	
Vegfr	human	2OH4	[53, 86]		Sorafenib	2400	
Nkcc1	human	7S1X	[57]		Bumetanide	500	
Hsp47	human	3ZHA	[59]		HS55	55000	
Sglt2	human	7VSI	[44]		Empaglifozin	2	
P2x3r	human	5YVE	[64, 65]		AF-219	330	
LolEC	*E.coli*	7ARH	[67]	*A.baumannii*	Abaucin	50000	[1]

Vkorc1, **V**itamin **K** epo**x**ide **R**eductase **C**omplex **1**; targeting anticoagulant rodenticides
FtsZ, **F**ilament **T**emperature **S**ensitive **Z**-ring; targeting cell division filaments of resistant *S. aureus*.
Sglt1, **S**odium-**GL**ucose **T**ransporter **1**; targeting reabsorption of intestinal sodium and glucose to control diabetes
Vegfr, **V**ascular **E**ndothelial **G**rowth **F**actor **R**eceptor; targeting tyrosine kinases to inhibit tumour angiogenesis
Nkcc1, **N**a$^+$**K**$^+$**C**l$^-$ **C**otransporter 1; targeting salt-related hypertension and neuronal excitability
Hsp47, **H**eat **S**hock **P**rotein **47**; targeting collagen folding and function implicated in tissue fibrosis and cancer
Sglt2, **S**odium-**GL**ucose **T**ransporter **2**; targeting reabsorption of kidney sodium and glucose to control diabetes
P2x3r, **P2X3 R**eceptor ; targeting ATP-gated cation channels implicated in rheumatoid arthritis and pain
LolEC, **L**ipoprotein **O**uter membrane **L**ocalization **EC**; targeting lipoprotein trafficking in *A. baumannii*
Ref., original references describing the *.pdb files coding for the selected crystallographic models.
Ref*, our previous references using *DW-BEL* without *Toxicity Risk*.

Generation of co-evolutionary children libraries

The last updated **D**ata**W**arrior (*DW*) versions were downloaded (https://openmolecules.org/datawarrior/download.html) for the generation of user-defined libraries. The updates included the dw550win.zip (windows) and/or dw550x.zip (Linus/MaxOS) which were substituted at the *DW* local directory. *DW* was launched from the *DataWarrior.exe*. Alternatively, *DW* was launched from a *startJarWin.bat* file kindly provided by Dr.T.Sander to best set ~ 60 Gb RAM memory requirements (-Xmx60g).

The *DW* docking (*DW/Chemistry/Dock Structures Into Protein Cavity*) uses the improved *mmff94s+* force-field[87] for energy minimization to best preserve children molecular geometries (most of them Nitrogen rings and double bonds) (https://cheminfo.github.io/openchemlib-s/classes/ForceFieldMMFF94.html and https://github.com/cheminfo/ openchemlib-js/lob/e88e8a0/types.d.ts#L3334).

DW-BEL algorithms generated large numbers of random raw children and selected those fitting the protein-ligand-cavity of the corresponding target/parent pairs by co-evolution with additional user-defined preferences. For that, each individual 2D parent structure was opened from an *.sdf file with *DW/File/Open*, selected and copied to *DW/Chemistry/BEL/Root generation compounds/from the clipboard*. In these studies, each parent molecule was employed without any selection of substructures to protect them from being changed (*lazo tool*). Then, each of the target / parent complexes coded into *.pdb files without CONECT lines, was opened into the *DW/BEL/Add Criterion/ Docking score/Load Protein Cavity From PDB-File,* to select for the target docking cavity. Other user-selected *Add Criterion* -> were included.

Taken into account the random nature of the generation of thousands of children, 3 consecutive runs for each pair were chosen. Since each consecutive run initiated a fresh new parent co-evolution, to avoid duplicates previous children were kept in RAM memory by the algorithms. Most of the runs required 10-50 Gb RAM memory, variations depending on each pair. Monitoring heap memory by Java Jconsole garbage collector (https://download.oracle.com/java/19/latest/jdk-19_windows-x64_bin.msi), were used to control for possible memory excess causing program crashes[2, 4] The number of raw children increased proportionally from 1 to 3 runs. The number of automatically stopped runs was limited to reduce demands for excessive memory resources.

DW randomly added small molecular modifications to the parent molecule, generating 128 children molecules per generation (*Compounds per cycle*). Children molecules were generated by randomly adding small molecular modifications to their parent. Java's Mutator applied random changes choosing from a list applying single atom replacements, atom insertions, single/double bond changes, atom migrations, ring aromatization/reduction, etc (https://github.com/Actelion/openchemlib/blob/master/src/main/java/com/actelion/research/chem/Mutator.java). The modifications were applied to the parent molecule, the resulting children ranked by user-selected fitting criteria and the best fitting children molecules listed for further modifications in the next generation. After each generation, a calculated weighted sum of all the user-selected fitting criteria, ranks each children by their fitness. In this study, 16 best children fitting user-set criteria were selected for each new generation (*Compounds survive a cycle*). Parent-children generations proceeded automatically until a fitness plateau after few hours. The *DW* docking-scores were expressed in unit-less relative values expanding from -20 to -140 ranks. The more negative, the higher affinity. The children were compared with and without *Toxicity Risk*.

The fitness criteria values and their weights applied in this study were: minimal *Docking Score* (weight = 4), *Molecular weight* <=400 g/mol (2), c*LogP* <=3 (1) and ± *Toxicity Risks* <= 1 (4).

DW *Toxicity Risk* assessment and *Nasty functions*

Toxicity risk assessments were developed by DW years ago to locate substructures within chemical structures divided in four classes (*Mutagenesis, Tumorigenicity, Reproductive Interference, Irritant*). These risks are alerts, not reliable toxicity predictions, nor 100 % free of any toxic effects. To assess toxicity prediction's reliability, *DW* ran a set of known toxic compounds and compare them to known non-toxic compounds to evaluate whether they have high, medium or low risk of being toxic. Structural fragment analysis generated toxicity alerts based on the **R**egistry of **T**oxic **E**ffects of **C**hemical **S**ubstances (RTECS data base) (https://www.cdc.gov/niosh/docs/97-119/default.html) and on > 3000 traded drugs non-toxic compounds (https://github.com/thsa/datawarrior/blob/master/src/html/properties/properties.html).

Previously defined *DW Nasty functions* are a list of small fragments having known physiological interference problems (https://openmolecules.org/forum/index.php?t=msg&th=662&start=0&). The corresponding *.*dwar* file contains the latest *DW Nasty functions* kindly supplied by Dr.T.Sander (**Supplementary Material: Nasty_functions.*dwar***)

Saving *DW-BEL* children libraries

A user-designed *DW* Macro called NTN (**Supplementary Material: NoToxiNasty.*dwam***) was developed to filter the generated raw children files by any remaining *DW* toxicity risk (*Mutagenesis, Tumorigenicity, Reproductive interference*, and *Irritant*), including the *Nasty functions*. The files were saved as *.*dwar* files for storage of the complete evolutionary data and their slider filters[2, 4] . The children coded into the *.*dwar* filtered by NTN macro were named as NTN children.The NTN children were also saved as special *.sdf (vs3) files, maintaining evolutionary information using *File/Save Special/SD-File*, selecting *Docked Protonation: Structure column, Docking pose: Atom coordinates* and including *Cavity and Natural Ligand*. After filtering for their toxicities added by the NTN macro, NTN children *.sdf files allowed their opening in PyMol, the use of its split-states command and visualization of the complexes (more details at the *DW* forum beginning on February 3th, 2023) (https://openmolecules.org/forum/index.php?t=msg&th=632&start=0&).

Consensus with AutoDockVina docking

Auto**D**ock**V**ina (ADV) dockings were performed by user-modified PyRx1.0 package (https://pyrx.sourceforge.io/)[88] , as described before[1]. To prepare for ADV, both crystallographic *.pdb protein targets without any ligand and their

corresponding NTN children *.sdf files were force-field minimized and charge-converted to *.pdbqt files. PyRx1.0 OpenBabel were selected choosing the mmff94s (Merck) force-field for energy minimization. The conservation of children molecular geometries were checked by comparing their *DW* InChiKeys before and after minimization. For ADV docking 45x45x45 Å grids automatically centred into the target proteins was chosen to explore any ADV docking-cavity alternatives. ADV generates rotatable conformers from input ligands and selects those with the lower docking-scores during iterations [89]. The conformer predicting the lowest binding-score is selected for approximated outputs expressed in Kcal/mol. Experimental accuracies of ± 2.8 Kcal/mol are predicted for ADV [88], while repetition of ADV docking to the same protein target were < 10 % of docking-scores (n=3-10)[90]

Computational programs

The ***B**uild **E**volutionary **L**ibrary* (*BEL*) algorithm using *Dock Structures into Docking Cavity* of **D**ata**W**arrior (*DW*) [46] written in Java were used. *DW* dw550win.zip for Windows and dw550x.zip for Linus, were downloaded from the 10 August 2023 updates (https://openmolecules.org/datawarrior/download.html), following *DW* guides and our previous work[2, 4]. The Python 3.8 written **A**uto**D**ock**V**ina (ADV) included in a user-modified PyRx1.0 package (https://pyrx.sourceforge.io/) was run for docking-score consensus with *DW*. ADV docking-score outputs in Kcal/mol were converted to nM Ki by the formula 10^9 *($\exp^{(Kcal/mol/0.592)}$). MolSoft (ICM Molbrowser vs3.9Win64bit, https://www.molsoft.com/ download.html) was used for manipulating the *.sdf files[2] and drawing 2D ligand structures. Origin (OriginPro 2022, 64 bit, Northampton, MA, USA) (https://www.originlab.com/) was used for calculations and drawings. The predicted structures were visualized in PyRx098/PyRx1.0 (Mayavi), Discover Studio Visualizer v21.1.0.20298 (Dassault Systemes Biovia Corp, 2020, https://discover.3ds.com/discovery-studio-visualizer-download) and PyMOL2.5.3 (https://www.pymol.org/). A multithreading multi-core i9 (47 CPU) AMD Ryzen Threadripper 3960X (PCSpecialist) computer, provided with 64 to 128 Gb of RAM (Corsair Vengeance DDR4, 4 x 16 or 4 x 32 GB) (https://www.pcspecialist.es/) was used to run the programs.

Results

Optimization of *Toxicity Risk* values/weights during *DW-BEL*

For optimization of the *Toxicity Risk,* the Vkorc1-brodifacoum protein / ligand (target / parent) complex was selected because of their high affinity and recently published crystallographic human/rat models[21,3].

To generate their co-evolutionary children, the *DW-BEL Docking-Score* criteria was selected with its highest weight (weight = 4). To increase specificity, *Molecular weight* <=400 g/mol (4), hydrophobicity *clogP* <=3 (1) and several *Toxicity Risk* values (4), were also selected. Other variables were Create *compounds like: Approved drugs,* and *Total round count = 1*. The preliminary conclusions of the prediction results with several *Toxicity Risk* values, were the following (**Figure S1**):

i) **Number of raw children**. There were 13033-15827 raw children randomly generated for *Toxicity Risk* values between 0 to 2. Lower 5695 raw children were predicted with 4 as *Toxicity Risk* value.

ii) Number of fitted children. There were 1393-2691 children (~ 10 % of the raw children) that fitted the user-selected criteria with values >0.89 (1.0 being the maximum). In contrast, only 977 children were predicted with 4 as *Toxicity Risk value.*

iii) Number of NTN children. There were 1251-2330 **N**o**T**oxi**N**a**s**ty (NTN) children for *Toxicity Risk* values between 0.01 to 2, displaying maximal *DW* affinities (minimal docking-scores) to its docking cavity between -116 to -127. In contrast, there were none NTN children predicted with 4 as *Toxicity Risk* value.

iv) **Percentages of NTN children**. The percentages of NTN children with *Toxicity Risk* values between 0.01 to 1, were between 82.6 to 98.8 % (100 % could only be expected if the criteria were set as thresholds rather than preferences).

The results commented above should be taken as preliminary because only one target / parent pair, one run for estimating stochastic variations and preference criteria. Within those limits, the *Toxicity Risk* value of 1 was chosen for the rest of the work with other target /parent pairs, since it generated ~ 2.6 fold higher number of NTN children than without *Toxicity Risk* (**Figure S1**, **red**). For the following *DW-BEL* work, variables predicting a high number of children were preferred.

***DW-BEL* of target / parent pair models**

Docking Score (weight of 4), *Molecular weight* <=400 g/mol (2), *clogP* <= 3 (1) ± *Toxicity Risk* <= 1 (4) were chosen for *DW-BEL* co-evolution of 9 target / parent pairs, selected by their practical importance (**Table 1**). In particular, one *.pdb file per each complex pair coding their crystallographic 3D binding cavities (protein and ligand), and one 2D *.sdf file coding for the ligand were uploaded to *DW-BEL*. Three consecutive runs of automatically decided number of cycles were performed for each target / parent pair ± *Toxicity Risk,* to best deal with the stochastic generation of raw children. According to what was briefly mentioned above, in this study the number of *Compounds surviving a cycle* were chosen as 16, higher than the 8 default number, to generate maximal numbers of children. Before applying the final NTN macro filtering, one Vkorc1 representative target / raw children sample analysis showed that all children containing Toxic groups and Nastic functions were reduced with *Toxicity Risk* (**Table S1**). Therefore the remaining Toxic groups and/or Nasty function children for all 9 target / parent pairs were totally eliminated by NTN filtering. Results predicted by *DW-BEL* ± *Toxicity Risk* could be summarized as follows (**Table 2 and Figure 1**):

i) Number of raw children. There were 23183-51950 (mean 35582 ± 7643) raw children generated after 3 consecutive runs. As expected, there were no significative differences between with *Toxicity Risk* (36075 ± 8534) or without (35088 ± 7125). The consecutive runs took ~ 200-400 cycles during 6 to 24 h to automatically finish, depending on each target/parent pair. The number of raw children generated per cycle remained nearly constant at ~ 100 raw children per cycle (except for Vegfr).

ii) Number of fitted children. There were 3606-8238 (mean 5702±1211) children fitting the user-selected criteria, which corresponded to ~ 16 % of the raw children. There were no significative differences with (5687 ± 1338) or without (5718 ± 1151) *Toxicity Risk*.

iii) Number of common NTN children ± *Toxicity Risk*. The number of common NTN children ± *Toxicity Risk* were < 5.2 % (< 221 NTN children). These low numbers suggest that *Toxicity Risk* during *DW-BEL* alters children co-evolution pathways predicting children from alternative chemical spaces.

iv) Percentage of NTN children. There were significative ~ 2-fold differences between the percentage of NTN children with (85.2 ± 19.1 %) compared to without (46.5 ± 20.3 %) *Toxicity Risk.* These results suggest that *Toxicity Risk* reduced the number of children with remaining *Toxic* and/or *Nasty Functions* and therefore increased the numbers of NTN children.

Table 2
Comparison of the *DW-BEL* co-evolutions of the protein target / parent pairs of Table 1.

Protein Target	Number Of Raw Children	Number / cycle	Number Of Fitted Children	NTN, Number	NTN children, %	NTN, Common
Vkorc1	**+** 40476	101.2	6425	**5307**	**82.6**	101
	- 42322	100.5	6963	2911	41.8	
FtsZ	**+** 35440	99.0	5737	**5307**	**92.5**	203
	- 39389	98.2	6568	4328	65.9	
Sglt1	**+** 28341	101.2	4484	**4157**	**92.7**	91
	- 30043	104.7	4598	1554	33.8	
Vegfr	**+** 51950	265.0	8238	**2916**	**35.4**	152
	- 39830	100.5	6340	241	3.8	
Nkcc1	**+** 38668	101.4	6185	**5641**	**91.2**	130
	- 27533	106.3	5089	3125	61.4	
Hsp47	**+**42697	107.8	6377	**5937**	**93.1**	119
	- 40188	101.4	6344	2481	39.1	
Sglt2	**+** 23183	106.3	3606	**3390**	**94.0**	170
	- 25445	101.4	4018	2696	67.1	
P2x3r	**+** 31442	101.0	4987	**4468**	**89.6**	106
	- 42972	99.0	6998	3149	45.0	
LolEC	**+** 32483	102.1	5147	**4967**	**96.5**	221
	- 28078	99.6	4548	2792	61.4	

DW-BEL co-evolution with (**+**) or without (**-**) *Toxicity Risk* <=1 (weight = 4), *Docking Score* (4), *Molecular weight* <=400 g/mol (2), and *clogP* <=3 (1). The fitted children were filtered by the NTN macro.
Number Of Raw Children, randomly generated children from parent by Java's Mutator**.**
Number / cycle, calculated by the formula, number of raw children / number of cycles
Number Of Fitted Children, number of children fitting the user-set criteria
NTN number, number of children ± *Toxicity Risk* filtered by the NTN macro
NTN children %, calculated by the formula, 100 x NTN Number / Number Of Fitted children.
Black bold and **green backgrounds,** NTN Number and % with *Toxicity Risk* (**+**)
NTN Common, Number of NTN children structures in common ± *Toxicity Risk* as calculated by *DW/Chemistry/Find Similar Compounds/Other File/ Structure[Exact]*

v) **Study of 100 NTN children top-leads**. Most of the 100 NTN children top-leads docking-scores were of similar mean values as shown by their low standard deviations (**Figure 1**). All NTN children top-leads ± *Toxicity Risk* predicted mean DW affinities higher (lower docking-scores) than those of their parents (**Figure 1, yellow hatched bars** and **white hatched bars** compared to **open bars,** respectively). There were top-lead NTN children that predicted higher mean DW affinities with than without *Toxicity Risk*, such as Vkorc1, Sglt1, Vegfr, Nkcc1 and LolEC. In contrast, there were also top-lead NTN children that predicted higher mean DW affinities only without *Toxicity Risk* such as FtsZ, Hsp47, Sglt2, and P2x3r (**Figure 1, yellow hatched bars** compared to **white hatched bars**). These results show that affinity improvements are not necessarily associated with *Toxicity Risk* but it depends on the target / parent pair under these criteria.

To improve prediction accuracies, the 100 top-leads were re-docked by **A**uto**D**ock**V**ina (ADV). The accuracy of ADV docking increased after the application of DW mmff94s+ force-field to children during *DW-BEL* (by correcting torsion angles) [87] reducing the aberrant geometries generated by alternative force-fields (not shown).

The comparison of DW and ADV docking scores of the 100 top-leads could be summarized as follows:

i) Despite their different DW / ADV algorithms, correlation trends were observed when comparing their corresponding top-lead docking-score means (**Figure S2**).

ii) *DW* affinities of each of the pairs grouped around similar values (~ ±10 unitless). In contrast, the corresponding ADV docking-scores estimations spreaded throughout values from 10 to 10^3 nM (**Figure 2**).

iii) The highest DW and ADV consensed affinities were predicted for FtsZ and Sglt2 (corresponding to lowest <10 nM ADV docking-scores). Sglt1 and Vegfr have only a few children predicting high affinities (< 10 nM ADV docking-scores). In comparison, the top-leads corresponding to Nkcc1, Hsp47 and P2x3r, predicted lower ADV affinities (corresponding to high ~10^2-10^4 nM docking-scores).

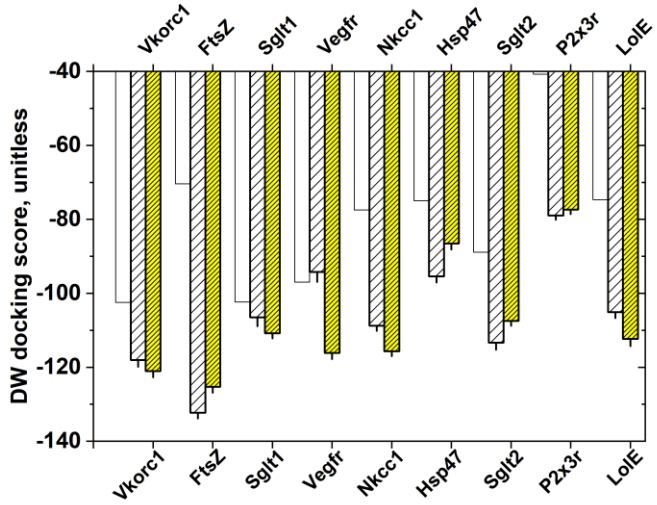

Figure 1
Comparison of NTN children binding-scores of 100 top-leads ± *Toxicity Risk*

Docking-score means ± standard deviations calculated for 100 top-leads from each pair.
NTN children were obtained after fitted children were filtered by the NTN macro.
Open bars, *DW* of parent molecules (their standard deviations were too low to be drawn)
White hatched bars, *DW-BEL* NTN children without *Toxicity Risk*.
Yellow intense-hatched bars, DW-BEL NTN children with *Toxicity Risk*.

PyMol drawings were generated for each of the target / parents (**Figure S2,** gray cartoons and **red sticks**, respectively) and compared to their ADV 10 top-lead children (**Figure S2, green sticks**). Comparisons of 2D structures between parents and their best top-lead children molecules were also compared in MolSoft (**Figure 3**).

Top-lead NTN children were selected by plotting *DW versus* ADV docking-scores (not shown) and by PyMol selecting those predicting docking to their binding cavities from their 10 ADV top-leads (included into the **Supplementary Material** *.pse files). Nevertheless, most of the 10 top-leads mapped to their crystallographic binding cavities (**Figure S2, green)** similarly to their parents (**Figure S2, red**). These high level of preservations were possible due to the *DW* particular *mmff94s+* force-field. Thus, *mmff94s+* minimization preserved most of the children molecular geometries. In contrast, other force-field methods like gaff, mmff94, uff, and/or ghemical, only preserved ~ 50 % of the initial children molecular geometries, as evidenced by *InChiKeys* comparisons (not shown).

Although most NTN children increased the previously qualified as "high" ~ µM affinities of their parents (**Figure 1**), alternative co-evolution criteria may still be explored for further improvements in selected target / parent pairs. For instance, the Nkcc1, Hsp47 and P2x3r, may be some examples requiring improvements. In particular, the lower affinities of Hsp47 and P2x3r may be explained by their incomplete binding cavity. It is possible that targeting their protein surfaces rather than internal cavities, resulted in too weak interactions for co-evolution (**Figure S2**). In this regard, collagen-Hsp47 interactions[73-76], may be required to define a complete binding cavity, rather than only a Hsp47 surface. Additional work may help to clarify these possibilities.

Comparison of the 2D structures from NTN top-leads with those from their parents showed different chemical scaffolds (**Figure 3**). *Molecular weight* and *clogP* were either maintained or lowered in the NTN top-leads compared to those of their parents (included into the **Supplementary Material** at their *.*dwar* files).

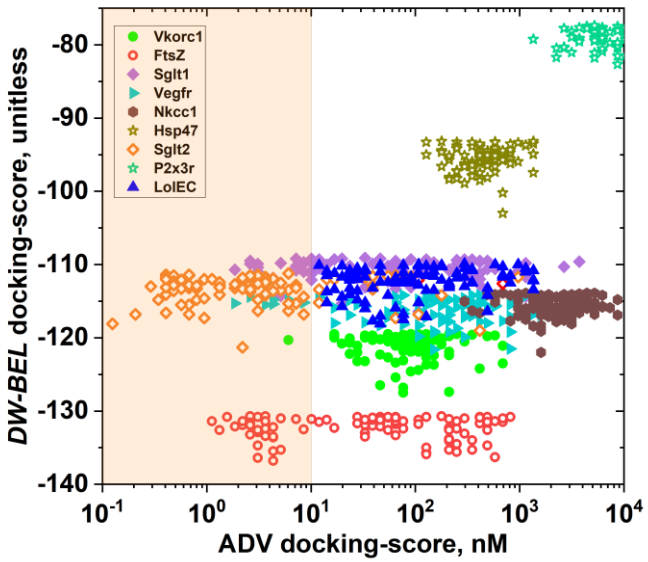

Figure 2
Comparison of DW and ADV docking-scores predicted from 100 NTN children top-leads

The 100 NTN children top-leads generated by *DW-BEL* ± *Toxicity Risk* (**Table 2**), were re-docked by ADV and their corresponding docking-scores represented in nM. Only the top-leads between -140 to -85 *DW-BEL* and 10^{-1} to 10^4 ADV predicting the lowest *DW-BEL* docking-scores ± *Toxicity Risk* were represented
Solid symbols, with *Toxicity Risk*. **Open symbols**, without *Toxicity Risk*.
Orange background, TNT children predicting < 10 nM ADV docking-scores
Green circles, rat Vkorc1 **Green triangles,** Vegfr **Orange-open diamonds,** Sglt2
Red-open circles, *A.baumannii* FtsZ **Brown hexagons,** Nkcc1 **Green-open stars,** P2xr3
Purple diamonds, Sglt1 **Green-Brown stars,** Hsp47 **Blue triangles,** *S.aureus* LolEC

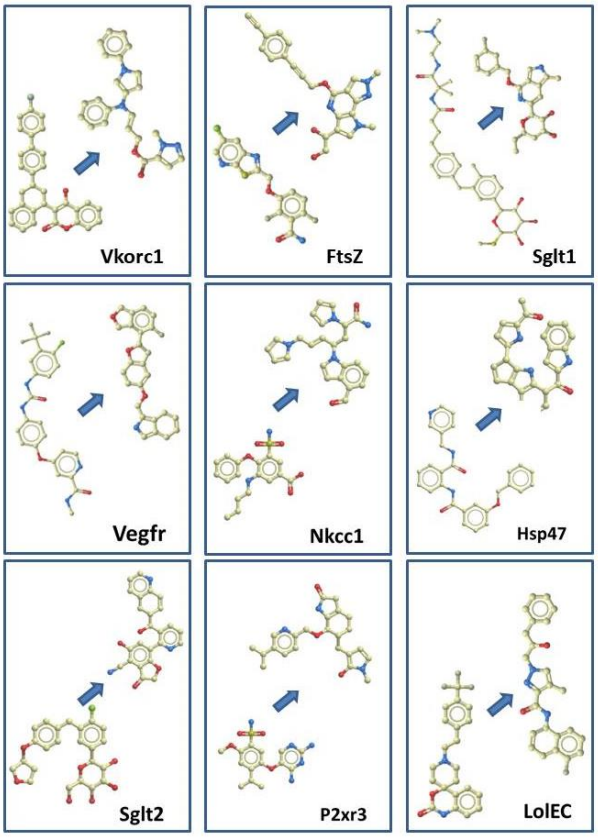

Figure 3
2D scheme of parents and each of their derived best top-lead NTN children

The best top-lead NTN children were drawn in MolSoft and identified at the legend bottom by their *DW-BEL* target name (**black**), parent (**red**) + generation number (**green**) and ADV docking-scores (**~ nM**). More detail of the complexes can be visualized amplifying the Figure view and/or opening their corresponding *.*dwar* files at the **Supplementary Material**. **Yellow circles and sticks**, Carbons. **Red circles,** Oxygens. **Blue circles,** Nitrogens and/or Sulphurs. **Green circles,** Halogens. **Blue Arrows,** go from the parent to the best top-lead child.
Vkorc1, Brodifacoum + 29611 (16.7 nM)
FtsZ, PC190723 + 12135 (1.1 nM)
Sglt1, LX1761 + 25743 (1.9 nM)
Vegfr, Sorafenib + 33763 (1.9 nM)
Nkcc1, Bumetanide + 23967 (688.2 nM)
Hsp47, Hs55 + 14736 (127.1 nM)
Sglt2, Empaglifozin + 23912 (0.1 nM)
P2xr3, Gefapixan + 30828 (1352.6 nM)
LolEC, Abaucin + 14761 (11.9 nM)

Discussion

This work explored the *Toxicity Risk* effects on *DW-BEL* co-evolutions to generate optimal children affinities (*DW Docking Score*) and specificities (*DW-BEL* / *Molecular weight* / *cLogP*) for nine protein/ligand (target / parent) pairs.

As may be expected, the inclusion of optimized *Toxicity Risk* increased ~ 2-fold the numbers of *DW-BEL* NTN children. Furthermore, under the same conditions, most of their corresponding top-lead *DW Docking Scores* were higher that those of their parents.

However, there were also some unexpected results. **First**, the numbers of fitted children were unaffected with *Toxicity Risk,* suggesting one possible internal control by the *DW-BEL* algorithms. **Second**, there were few common children between the chemical structures of NTN children predicted ± *Toxicity Risk*, suggesting the induction of independent co-evolution pathways extending differently throughout the chemical space. **Third**, although there were children predicting highest affinities with *Toxicity Risk* (Vkorc1, Sglt1, Vegfr, Nkcc1, LolEC), there were also other children in which the highest affinities were predicted without *Toxicity Risk* (FtsZ, Hsp47, Sglt2, P2x3r). The stochastic nature of raw children generations, the differently limited protein cavity shapes / volumes, the hydrophobicity variations among children and/or other unknown target steric limitations, could explain those differences. Overall, these results highlight the existence of a vast chemical space to be further explored [21, 22].

Most docking-cavities predicted by *DW-BEL* ± *Toxicity Risk* were similar to the crystallographic binding-cavities, suggesting a probable conservation of their ligand biological activities. However, confirmation of these hypothesis will require chemical synthesis and subsequent experimental tests.

The proposed no-toxic co-evolution has enriched the number of candidates predicting higher affinities and specificities in only a few hours of computation (mostly due to the fast Java's code). All these would had been impossible for any other more "traditional" computational screening (i.e., AutoDockVina, Yasara, seeSAR, etc) of largest chemical libraries (i.e., Mcule, ChemSpace, Zinc, PubChem, Chembl, etc) which would have required much more computational time. On the other hand, most actual machine-learning approaches to docking[91, 92]; including those employing the new transformer methods[78, 91,93-95], are actually limited in their accuracies because of the reduced numbers of examples of small-drugs protein interactions required for model training[96]. Some hands-on experience of the state-of-the-art of machine-learning trained models, was acquired by employing Hots proposed methods[91] Hots predicted both docking-cavities and ligand docking-scores solely from protein amino acid sequences[91]. Despite Hots successful identification of the Vkorc1 binding cavity, Hots predicted only one possible ligand candidate. When forced to predict 150 ligand docking-scores[3], several days were required to computational complete the docking. Furthermore, the generated docking-scores correlated poorly when compared to DW or ADV docking programs (results not shown). It may be concluded that programs by machine learning as docking alternatives are still on development. In contrast, the alternative random generation / selection algorithms proposed here, could generate higher numbers of candidates much faster and including *Toxicity Risk* (and/or other criteria) during their co-evolution (libraries-on-demand or generative biology by "co-evolutionary docking"). The question may remain as to whether or not, any evolutionary approach similar to the one described here, could be hybridized with any machine-learning for further improvements. Similar "hybrid" methods may be most important to predict protein-protein predictions[96, 97].

To allow to any potential readers for additional exploration of "non-toxic co-evolutionary docking", thousands of NTN children molecules derived from nine target / parent pairs were downsized to their 100 top-leads and their data included into *DW* / **.dwar* files. These **.dwar* files also included AutoDockVina docking-score data with their corresponding nanoMolar affinity predictions.

Larger *DW-BEL* libraries could be generated by additional runs and/or adding alternative co-evolutionary fitting criteria, and/or to adapt to other possible parent interactions maintaining any of their substructures (*lazo* variability). Some of the *DW* algorithms ready to use in *DW-BEL*, presently include not only *Docking score*, *Molecular weight, Hydrophobicity,* and *Toxicity Risk* but also *Basic nitrogen counts, Acidic oxygen counts, Ring count, Structural (dis)similarity, Conformers similarity, Molecular flexibility, Molecular complexity, Molecular shape*, and other. Inclusion of any new *DW* criteria may be requested into the *DW* forum (https://openmolecules.org/ datawarrior/download.html).

The described *DW-BEL* "non-toxic co-evolutionary docking" results predicted high numbers of TNT children with nanoMolar affinities, new chemotypes, high specificity and conservation of previously defined binding cavities. Because exploration of the vast chemical/chemotype space appear almost endless, further *DW-BEL* "co-evolutionary docking" criteria may possibly identify more candidates. For instance, it should be also possible to further explore novel artificially-build docking cavities (i.e., preliminary results with poly-benzene *de novo* ligands), co-evolving with alternative chemical synthesis pathway-preferences, and/or discarding off-target affinities during co-evolution.

Supporting information

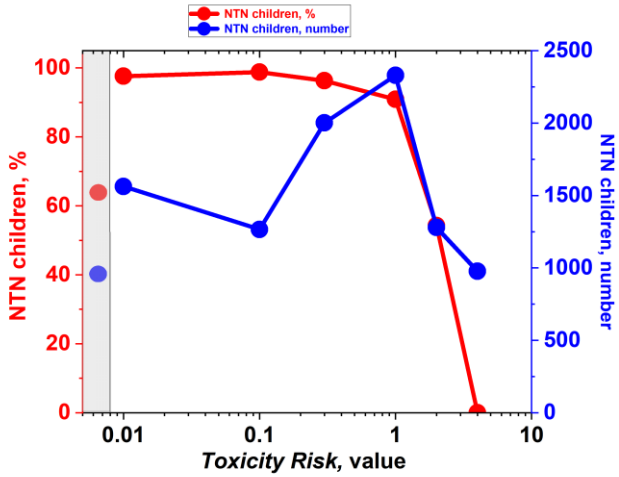

Figure S1
NTN children predicted using different *Toxicity Risk* values

The rat Vkorc1-brodifacoum complex was selected as target. The parent was brodifacoum. The co-evolution criteria included: *Cycle*: *automatic*, *Total run count* = *1*, *Compounds per cycle* = *128*, *Compounds survive a cycle* = *16* , *Create compounds like* = *Approved drugs*. The *Add Criterion* included: *Docking Score* = rat Vkorc1-brodifacoum, *Docking Score* of weight = 4 (4), *Molecular weight* <= 400g/mol (2), *cLogP*<=3 (1), and ± *Toxicity Risk* <= Variable (4). The % NTN children were calculated by the formula, 100 * NTN children / Fitted children.
Left gray background, NTN children without adding any *Toxicity Risk*.
Red circles and line, NTN children with *Toxicity Risk*, %.
Blue circles and line, NTN children with *Toxicity Risk*, number.

Table S1
Example of distribution of *Toxicities* and *Nasty functions* of raw children ± *Toxicity Risk*

Vkorc1 Raw children numbers	Mutagenic	Tumorigenic	Reproductive Effective	Irritant	Nasty Functions
+ 6425	110	14	112	228	696
- 6963	1859	1954	1571	613	1994

The numbers of 4 *Toxicity* groups and *Nastic functions* present in raw children after being *DW-BEL* generated ± *Toxicity Risk* co-evolution were compared by individually manual filtering by *DW/Chemistry/From Chemical Structure/Calculate Properties/Tox*.
Some of the raw children were classified in several of the groups (not shown).

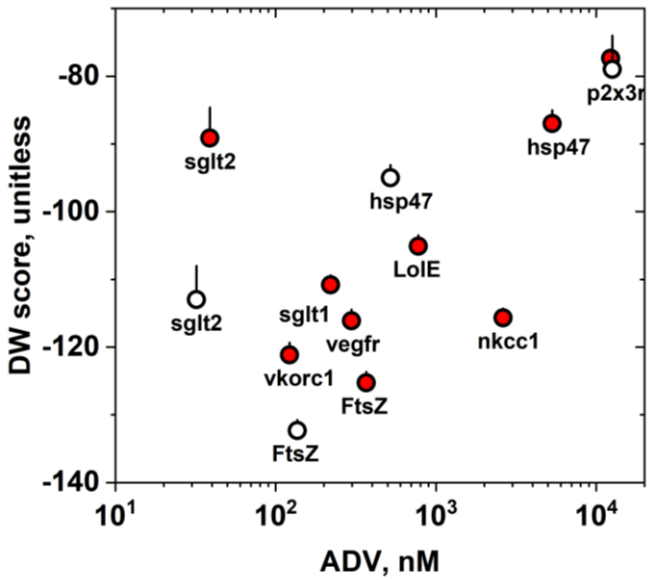

Figure S2
Predicted linear correlation between *DW* and ADV means of NTN children top-leads (n=100).

The *DW-BEL* criteria ± *Toxicity Risk* were applied during co-evolution as described in **Figure 2.** The Fitted children generated were finally filtered by the NTN macro**.** Means ± standard deviations (n=100) were calculated from data (**Figure 2)** and other data not shown. **Red circles,** *DW-BEL* means with *Toxicity Risk*. **Open circles**, *DW-BEL* means without *Toxicity Risk*.

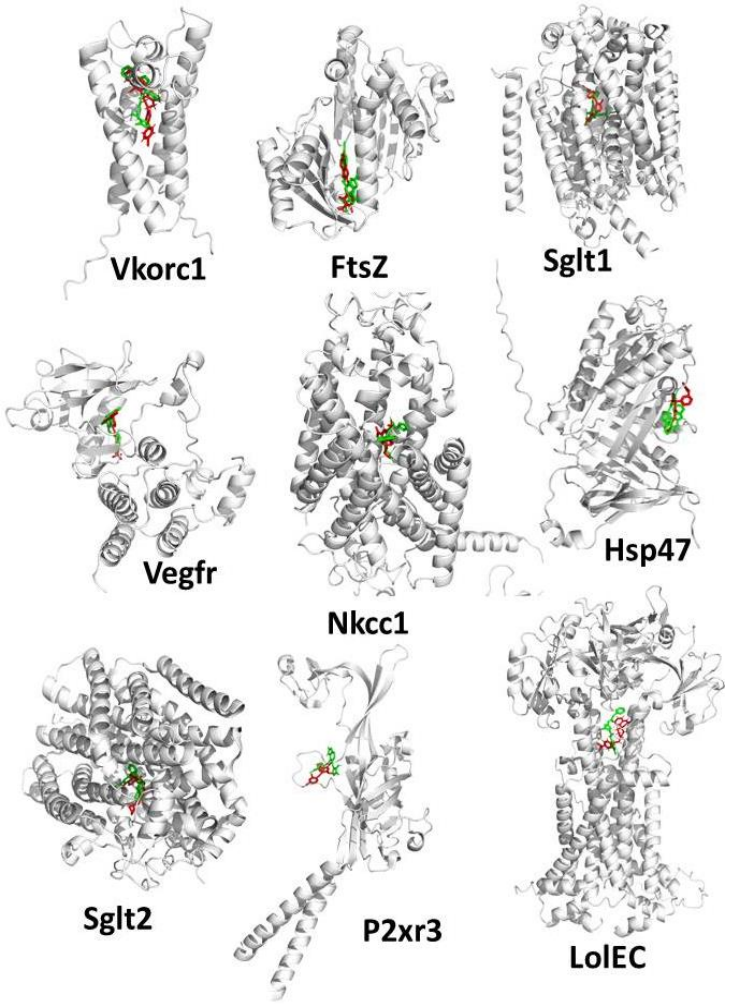

Figure S3
Representative children docking- (green sticks) and crystallographic ligand binding- (red sticks) cavities

The complexes were drawn in PyMol vs2.5.3. The best top-lead NTN children were identified at this legend bottom by their *DW-BEL* target name (**black**), parent (**red**) + generation number (**green**). More detail of the complexes can be visualized amplifying the Figure view and/or visualizing 10 complexes within PyMol vs2.5.3. opening their corresponding *.pse files (**Supplementary Material**). Gray cartoons, 3D protein targets. **Red sticks**, *DW-BEL* parents. **Green sticks**, best of top-lead children selected among 100 top-leads.

Vkorc1, Brodifacoum + 29611
FtsZ, PC190723 + 12135
Sglt1, LX1761 + 25743
Vegfr, Sorafenib + 33763
Nkcc1, Bumetanide + 23967
Hsp47, Hs55 + 14736
Sglt2, Empaglifozin + 23912
P2xr3, Gefapixan + 30828
LolEC, Abaucin + 14761

Funding: This research was not externally funded .

Competing interests
The author declare no competing interests

Authors' contributions
JC performed and analyzed the dockings, and drafted the manuscript. .

Acknowledgements
Thanks are due to Dr. Thomas Sander of *DW* (Idorsia Pharmaceuticals Ltd. at Allschwil, Switzerland), for including/adapting the *Toxicity Risk* criterion code to *DW-BEL* and supplying additional information about the *DW Toxicity* and *Nasty Functions*. Special thanks are also due to Norwid Behrnd of *DW* for extensive valuable discussions and contributions to interpret *DW-BEL Toxicity Risk* data and to fine tune / develop the NTN filtering macros. Thanks are also due to Dr. A. Villena from the University of Leon (Spain) for his valuable help with most recent bibliographic updates on actual machine learning methods.

Supplementary Material

- NoToxiNasty.*dwam*. A *DW* macro *.dwam file developed to save, label and eliminate any children molecules co-evolutionary generated by the *DW-BEL* containing *T*oxicity (*Mutagenesis, Tumorigenicity, Reproductive Interference, Irritant*) and/or *Nasty Functions* (see ***Nasty_functions.dwar***). This macro labels and retains the children having no detectable risks. The macro uses *.sdf or *.*dwar* files as inputs, user-renamed the input *.*dwar* file and renamed and saved the corresponding NTN *.*dwar* and toxic-labelled *.sdf files (they require filtering outside *DW*). More than ~ 3000 traded drugs were employed as low toxicity example data (https://github.com/thsa/datawarrior/blob/master/src/html/properties/properties.html. Additional information on *DW Toxicity* can be found at the **R**egistry of **T**oxic **E**ffects of **C**hemical **S**ubstances (RTECS data base) (https://www.cdc.gov/niosh/docs/97-119/default.html).

- ***Nasty_functions.dwar*.** List of previously defined *DW Nasty functions* 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&).

- **Vkorc1.*dwar*, FtsZ.*dwar*, Sglt1.*dwar*, Vegfr.*dwar*, Nkcc1.*dwar*, Hsp47.*dwar*, Sglt2.*dwar*, P2x3r.*dwar*, LolEC.*dwar***. These *.dwar *DW* tables contain 100 DW top-leads selected as NTN children corresponding to nine target / parent pairs. They are provided with threshold *slider-filters* to their *DW* and ADV docking-scores, *Molecular weights* and *clogP* properties. By moving the *slider-filters* located at the right of the *DW* Table, the best fitting children to particular threshold combinations could be selected and further studied. Each *.*dwar* file can be opened by downloading *DW* free access at https://openmolecules.org/datawarrior/download.htm. The *.*dwar* files can be also saved as special *.sdf (vs3) files, maintaining their 3D protein cavity docked to children conformers so that they can be opened in PyMol using its split-states command (more details at the *DW* forum from February 3th, 2023, https://openmolecules.org/forum/index.php?t=msg&th=632&start=0&).

- **Vkorc1.pse, FtsZ.pse, Sglt1.pse, Vegfr.pse, Nkcc1.pse, Hsp47.pse, Sglt2.pse, P2x3r.pse, LolEC.pse**. The *DW-BEL* best and 9 additional top-lead children complexes with their corresponding protein targets can be visualized in PyMol vs2.5.3. by opening their corresponding *.pse files (best top-lead children compared to initial ligands represented in their target proteins in **Figure S3**).

References

[1]Coll, J. Could Acinetobacter baumannii Lol-abaucin docking be improved?*ChemRxiv*. 2023: https://chemrxiv.org/engage/chemrxiv/article-details/649aa71aba3e99daef1d1756.

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

[3]Coll, J.M. Anticoagulant rodenticide novel candidates predicted by evolutionary docking.*ChemRxiv*. 2023: https://chemrxiv.org/engage/chemrxiv/article-details/6479b8cfbe16ad5c57577cce.

[4]Coll, J.M. New star-shaped ligands generated by evolutionary fitting the Omicron spike inner-cavity.*ChemRxiv*. 2023: https://chemrxiv.org/engage/chemrxiv/article-details/6479b8cfbe16ad5c57577cce.

[5]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.

[6]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.

[7]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.

[8]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.

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

[10]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.

[11]Coll, J.M. Explorando compuestos rodenticidas con inteligencia artificial.*Jornada TecnicaANECPLA. Avances en elcontrol de roedores y en los estudios de resistencia*. . 2021: On line presentation

[12]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.

[13]Mirdita, M., *et al.* ColabFold: making protein folding accessible to all.*Nat Methods*. 2022, 19: 679-682 http://dx.doi.org/10.1038/s41592-022-01488-1.

[14]Liu, S., *et al*. Structural basis of antagonizing the vitamin K catalytic cycle for anticoagulation.*Science*. 2021, 371: http://dx.doi.org/10.1126/science.abc5667.

[15]Buckle, A., *et al.* Resistance testing and the effectiveness of difenacoum against Norway rats (Rattus norvegicus) in a tyrosine139cysteine focus of anticoagulant resistance, Westphalia, Germany.*Pest Manag Sci*. 2013, 69: 233-9 http://dx.doi.org/10.1002/ps.3373.

[16]Diaz, J.C., *et al*. Analysis of vkorc1 polymorphisms in Norway rats using the roof rat as outgroup.*BMC Genet*. 2010, 11: 43 1471-2156-11-43 [pii], http://dx.doi.org/10.1186/1471-2156-11-43.

[17]Pelz, H.J., *et al.* The genetic basis of resistance to anticoagulants in rodents.*Genetics*. 2005, 170: 1839-47 genetics.104.040360 [pii], http://dx.doi.org/10.1534/genetics.104.040360.

[18]Czogalla, K.J., *et al*. Structural Modeling Insights into Human VKORC1 Phenotypes.*Nutrients*. 2015, 7: 6837-51 nu7085313 [pii], http://dx.doi.org/10.3390/nu7085313.

[19]Bermejo-Nogales, A., *et al*. VKORC1 single nucleotide polymorphisms in rodents in Spain.*Chemosphere*. 2022, 308: 136021 http://dx.doi.org/10.1016/j.chemosphere.2022.136021.

[20]Huecas, S., *et al*. Energetics and geometry of FtsZ polymers: nucleated self-assembly of single protofilaments.*Biophys J*. 2008, 94: 1796-806 http://dx.doi.org/10.1529/biophysj.107.115493.

[21]Artola, M., *et al.* The structural assembly switch of cell division protein FtsZ probed with fluorescent allosteric inhibitors.*Chem Sci*. 2017, 8: 1525-1534 http://dx.doi.org/10.1039/c6sc03792e.

[22]Sharma, A.K., *et al.* A salt bridge mediated resistance mechanism to FtsZ inhibitor PC190723 revealed by a single step cell-based screen..*bioRxiv*. 2022: http://doi.org/10.1101/2022.04.06.487355.

[23]Ruiz, F.M., *et al.* FtsZ filament structures in different nucleotide states reveal the mechanism of assembly dynamics.*PLoS Biol*. 2022, 20: e3001497 http://dx.doi.org/10.1371/journal.pbio.3001497.

[24]Fujita, J., *et al.* Structural Flexibility of an Inhibitor Overcomes Drug Resistance Mutations in Staphylococcus aureus FtsZ.*ACS Chem Biol*. 2017, 12: 1947-1955 http://dx.doi.org/10.1021/acschembio.7b00323.

[25]Wagstaff, J.M., *et al.* A Polymerization-Associated Structural Switch in FtsZ That Enables Treadmilling of Model Filaments.*mBio*. 2017, 8: http://dx.doi.org/10.1128/mBio.00254-17.

[26]Artola, M., *et al.* Effective GTP-replacing FtsZ inhibitors and antibacterial mechanism of action.*ACS Chem Biol*. 2015, 10: 834-43 http://dx.doi.org/10.1021/cb500974d.

[27]Huecas, S., *et al.* Targeting the FtsZ Allosteric Binding Site with a Novel Fluorescence Polarization Screen, Cytological and Structural Approaches for Antibacterial Discovery.*J Med Chem*. 2021, 64: 5730-5745 http://dx.doi.org/10.1021/acs.jmedchem.0c02207.

[28]Tan, C.M., *et al.* Restoring methicillin-resistant Staphylococcus aureus susceptibility to beta-lactam antibiotics.*Sci Transl Med*. 2012, 4: 126ra35 http://dx.doi.org/10.1126/scitranslmed.3003592.

[29]Matsui, T., *et al.* Structural reorganization of the bacterial cell-division protein FtsZ from Staphylococcus aureus.*Acta Crystallogr D Biol Crystallogr*. 2012, 68: 1175-88 http://dx.doi.org/10.1107/S0907444912022640.

[30]Elsen, N.L., *et al.* Mechanism of action of the cell-division inhibitor PC190723: modulation of FtsZ assembly cooperativity.*J Am Chem Soc*. 2012, 134: 12342-5 http://dx.doi.org/10.1021/ja303564a.

[31]Stokes, N.R., *et al.* An improved small-molecule inhibitor of FtsZ with superior in vitro potency, drug-like properties, and in vivo efficacy.*Antimicrob Agents Chemother*. 2013, 57: 317-25 http://dx.doi.org/10.1128/AAC.01580-12.

[32]Bryan, E., *et al.* Structural and Antibacterial Characterization of a New Benzamide FtsZ Inhibitor with Superior Bactericidal Activity and In Vivo Efficacy Against Multidrug-Resistant Staphylococcus aureus.*ACS Chem Biol*. 2023: http://dx.doi.org/10.1021/acschembio.2c00934.

[33]Butler, M.S. and Paterson, D.L. Antibiotics in the clinical pipeline in October 2019.*J Antibiot (Tokyo)*. 2020, 73: 329-364 http://dx.doi.org/10.1038/s41429-020-0291-8.

[34]Jornada, D.H., *et al.* The Prodrug Approach: A Successful Tool for Improving Drug Solubility.*Molecules*. 2015, 21: 42 http://dx.doi.org/10.3390/molecules21010042.

[35]Czaplewski, L.G., *et al.* Antibacterial alkoxybenzamide inhibitors of the essential bacterial cell division protein FtsZ.*Bioorg Med Chem Lett*. 2009, 19: 524-7 http://dx.doi.org/10.1016/j.bmcl.2008.11.021.

[36]Haydon, D.J., *et al.* Creating an antibacterial with in vivo efficacy: synthesis and characterization of potent inhibitors of the bacterial cell division protein FtsZ with improved pharmaceutical properties.*J Med Chem*. 2010, 53: 3927-36 http://dx.doi.org/10.1021/jm9016366.

[37]Haydon, D.J., *et al.* An inhibitor of FtsZ with potent and selective anti-staphylococcal activity.*Science*. 2008, 321: 1673-5 http://dx.doi.org/10.1126/science.1159961.

[38]Kaul, M., *et al.* An FtsZ-targeting prodrug with oral antistaphylococcal efficacy in vivo.*Antimicrob Agents Chemother*. 2013, 57: 5860-9 http://dx.doi.org/10.1128/AAC.01016-13.

[39]Vergoños-Tomas, A. Inhibidores de la Proteína de División Celular Bacteriana FtsZ Dirigidos al Sitio de Unión del Nucleótido.*Ph.D. Thesis, Autonomous University of Madrid, Madrid, Spain*. 2017: https://repositorio.uam.es/handle/10486/680161?show=full&locale-attribute=en.

[40]Ramirez-Aportela, E. Dinámica de los Filamentos de FtsZ y Búsqueda Racional de Inhibidores Sintéticos con Actividad Antibacteriana.*Ph.D. Thesis, Autonomous University of Madrid, Madrid, Spain*. 2017: https://repositorio.uam.es/handle/10486/677469.

[41]Sun, N., *et al.* Rational design of berberine-based FtsZ inhibitors with broad-spectrum antibacterial activity.*PLoS One*. 2014, 9: e97514 http://dx.doi.org/10.1371/journal.pone.0097514.

[42]Cui, W., *et al.* Structures of human SGLT in the occluded state reveal conformational changes during sugar transport.*Nat Commun*. 2023, 14: 2920 http://dx.doi.org/10.1038/s41467-023-38720-1.

[43]Niu, Y., *et al.* Structural mechanism of SGLT1 inhibitors.*Nat Commun*. 2022, 13: 6440 http://dx.doi.org/10.1038/s41467-022-33421-7.

[44]Niu, Y., *et al.* Structural basis of inhibition of the human SGLT2-MAP17 glucose transporter.*Nature*. 2022, 601: 280-284 http://dx.doi.org/10.1038/s41586-021-04212-9.

[45]Macalalad, M.A.B. and Gonzales, A.A., 3rd. In Silico Screening and Identification of Antidiabetic Inhibitors Sourced from Phytochemicals of Philippine Plants against Four Protein Targets of Diabetes (PTP1B, DPP-4, SGLT-2, and FBPase).*Molecules*. 2023, 28: http://dx.doi.org/10.3390/molecules28145301.

[46]Bhattacharya, S., *et al.* An exhaustive perspective on structural insights of SGLT2 inhibitors: A novel class of antidiabetic agent.*Eur J Med Chem*. 2020, 204: 112523 http://dx.doi.org/10.1016/j.ejmech.2020.112523.

[47]Feng, C., *et al.* Effect of SGLT2 inhibitor on renal function in patients with type 2 diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials.*Int Urol Nephrol*. 2019, 51: 655-669 http://dx.doi.org/10.1007/s11255-019-02112-6.

[48]Taghour, M.S., *et al.* Discovery of new quinoline and isatine derivatives as potential VEGFR-2 inhibitors: design, synthesis, antiproliferative, docking and MD simulation studies.*J Biomol Struct Dyn*. 2023: 1-16 http://dx.doi.org/10.1080/07391102.2022.2164356.

[49]Deng, Y.Y., *et al.* Comparison of the efficacy and safety of fruquintinib and regorafenib in the treatment of metastatic colorectal cancer: A real-world study.*Front Oncol*. 2023, 13: 1097911 http://dx.doi.org/10.3389/fonc.2023.1097911.

[50]Elkaeed, E.B., *et al.* Discovery of New VEGFR-2 Inhibitors: Design, Synthesis, Anti-Proliferative Evaluation, Docking, and MD Simulation Studies.*Molecules*. 2022, 27: http://dx.doi.org/10.3390/molecules27196203.

[51]Yousef, R.G., *et al.* Discovery of new nicotinamides as apoptotic VEGFR-2 inhibitors: virtual screening, synthesis, anti-proliferative, immunomodulatory, ADMET, toxicity, and molecular dynamic simulation studies.*J Enzyme Inhib Med Chem*. 2022, 37: 1389-1403 http://dx.doi.org/10.1080/14756366.2022.2070744.

[52]Eissa, I.H., *et al.* Design, molecular docking, in vitro, and in vivo studies of new quinazolin-4(3H)-ones as VEGFR-2 inhibitors with potential activity against hepatocellular carcinoma.*Bioorg Chem*. 2021, 107: 104532 http://dx.doi.org/10.1016/j.bioorg.2020.104532.

[53]Dahab, M.A., *et al.* Semi-synthesized anticancer theobromine derivatives targeting VEGFR-2: in silico and in vitro evaluations.*J Biomol Struct Dyn*. 2023: 1-20 10.1080/07391102.2023.2219333.

[54]Savardi, A., *et al.* Preclinical Development of the Na-K-2Cl Co-transporter-1 (NKCC1) Inhibitor ARN23746 for the Treatment of Neurodevelopmental Disorders.*ACS Pharmacol Transl Sci*. 2023, 6: 1-11 http://dx.doi.org/10.1021/acsptsci.2c00197.

[55]Yang, X., *et al.* Structure of the human cation-chloride cotransporter NKCC1 determined by single-particle electron cryo-microscopy.*Nat Commun*. 2020, 11: 1016 http://dx.doi.org/10.1038/s41467-020-14790-3.

[56]Liu, R., *et al.* Role of NKCC1 and KCC2 in Epilepsy: From Expression to Function.*Front Neurol*. 2019, 10: 1407 http://dx.doi.org/10.3389/fneur.2019.01407.

[57]Zhao, Y., *et al.* Structural basis for inhibition of the Cation-chloride cotransporter NKCC1 by the diuretic drug bumetanide.*Nat Commun*. 2022, 13: 2747 http://dx.doi.org/10.1038/s41467-022-30407-3.

[58]Jalan, A.A., *et al.* Chain alignment of collagen I deciphered using computationally designed heterotrimers.*Nat Chem Biol*. 2020, 16: 423-429 http://dx.doi.org/10.1038/s41589-019-0435-y.

[59]Cai, H., *et al.* Identification of HSP47 Binding Site on Native Collagen and Its Implications for the Development of HSP47 Inhibitors.*Biomolecules*. 2021, 11: http://dx.doi.org/10.3390/biom11070983.

[60]Yoshida, M., *et al.* Structure-Activity Relationship Study on Col-003, a Protein-Protein Interaction Inhibitor between Collagen and Hsp47.*Chem Pharm Bull (Tokyo)*. 2020, 68: 220-226 http://dx.doi.org/10.1248/cpb.c19-00634.

[61]Miyamura, T., *et al.* Small molecule inhibitor of HSP47 prevents pro-fibrotic mechanisms of fibroblasts in vitro.*Biochem Biophys Res Commun*. 2020, 530: 561-565 http://dx.doi.org/10.1016/j.bbrc.2020.07.085.

[62]Gu, Y., *et al.* Epac-protein kinase C alpha signaling in purinergic P2X3R-mediated hyperalgesia after inflammation.*Pain*. 2016, 157: 1541-1550 http://dx.doi.org/10.1097/j.pain.0000000000000547.

[63]Chen, Y., *et al.* p38 MAPK mediates glial P2X7R-neuronal P2Y1R inhibitory control of P2X3R expression in dorsal root ganglion neurons.*Mol Pain*. 2015, 11: 68 http://dx.doi.org/10.1186/s12990-015-0073-7.

[64]Kang, K.M., *et al.* AI-based prediction of new binding site and virtual screening for the discovery of novel P2X3 receptor antagonists.*Eur J Med Chem*. 2022, 240: 114556 http://dx.doi.org/10.1016/j.ejmech.2022.114556.

[65]Wang, J., *et al.* Druggable negative allosteric site of P2X3 receptors.*Proc Natl Acad Sci U S A*. 2018, 115: 4939-4944 http://dx.doi.org/10.1073/pnas.1800907115.

[66]Corsello, S.M., *et al.* The Drug Repurposing Hub: a next-generation drug library and information resource.*Nat Med*. 2017, 23: 405-408 http://dx.doi.org/10.1038/nm.4306.

[67]Liu, G., *et al.* Deep learning-guided discovery of an antibiotic targeting Acinetobacter baumannii.*Nat Chem Biol*. 2023: http://dx.doi.org/10.1038/s41589-023-01349-8.

[68]Usjak, D., *et al.* Targeting outer membrane protein A (OmpA) - inhibitory effect of 2'-hydroxychalcone derivatives on Acinetobacter baumannii and Candida albicans dual-species biofilm formation.*Biofouling*. 2023: 1-11 http://dx.doi.org/10.1080/08927014.2023.2215693.

[69]Kumar, A., *et al.* Targeting multi-drug-resistant Acinetobacter baumannii: a structure-based approach to identify the promising lead candidates against glutamate racemase.*J Mol Model*. 2023, 29: 188 http://dx.doi.org/10.1007/s00894-023-05587-4.

[70]Leao, P.V.S., *et al.* Riparin-B as a Potential Inhibitor of AdeABC Efflux System from Acinetobacter baumannii.*Evid Based Complement Alternat Med*. 2023, 2023: 1780838 http://dx.doi.org/10.1155/2023/1780838.

[71]Kumari, P., *et al.* Heterocyclic Diaryliodonium-Based Inhibitors of Carbapenem-Resistant Acinetobacter baumannii.*Microbiol Spectr*. 2023, 11: e0477322 http://dx.doi.org/10.1128/spectrum.04773-22.

[72]Gopikrishnan, M. and George Priya Doss, C. Molecular docking and dynamic approach to screen the drug candidate against the Imipenem-resistant CarO porin in Acinetobacter baumannii.*Microb Pathog*. 2023, 177: 106049 http://dx.doi.org/10.1016/j.micpath.2023.106049.

[73]Eduvirgem, J., *et al.* Antimicrobial and antibiofilm activities of desloratadine against multidrug-resistant Acinetobacter baumannii.*Future Microbiol*. 2023, 18: 15-25 http://dx.doi.org/10.2217/fmb-2022-0085.

[74]Liu, H., *et al.* In vitro analysis of synergistic combination of polymyxin B with 12 other antibiotics against MDR Acinetobacter baumannii isolated from a Chinese tertiary hospital.*J Antibiot (Tokyo)*. 2023, 76: 20-26 http://dx.doi.org/10.1038/s41429-022-00573-z.

[75]Wu, H.J., *et al.* Drug-resistant Acinetobacter baumannii: From molecular mechanisms to potential therapeutics (Review).*Exp Ther Med*. 2023, 25: 209 http://dx.doi.org/10.3892/etm.2023.11908.

[76]Novovic, K. and Jovcic, B. Colistin Resistance in Acinetobacter baumannii: Molecular Mechanisms and Epidemiology.*Antibiotics (Basel)*. 2023, 12: http://dx.doi.org/10.3390/antibiotics12030516.

[77]Rahimzadeh, G., *et al.* Genotypic Patterns of Multidrug-Resistant Acinetobacter baumannii: A Systematic Review.*Adv Biomed Res*. 2023, 12: 56 http://dx.doi.org/10.4103/abr.abr_434_22.

[78]Sun, C., *et al.* Resistance mechanisms of tigecycline in Acinetobacter baumannii.*Front Cell Infect Microbiol*. 2023, 13: 1141490 http://dx.doi.org/10.3389/fcimb.2023.1141490.

[79]Mittal, K.R., *et al.* Multidrug-Resistant Acinetobacter baumannii: An Emerging Aspect of New Drug Discovery.*Recent Adv Antiinfect Drug Discov*. 2023, 18: 29-41 http://dx.doi.org/10.2174/2772434417666220912120726.

[80]Giannouli, M., *et al.* Molecular epidemiology of carbapenem-resistant Acinetobacter baumannii strains in intensive care units of multiple Mediterranean hospitals.*J Antimicrob Chemother*. 2009, 63: 828-30 http://dx.doi.org/10.1093/jac/dkp032.

[81]Gallagher, P. and Baker, S. Developing new therapeutic approaches for treating infections caused by multi-drug resistant Acinetobacter baumannii: Acinetobacter baumannii therapeutics.*J Infect*. 2020, 81: 857-861 http://dx.doi.org/10.1016/j.jinf.2020.10.016.

[82]Perez, F., *et al.* Global challenge of multidrug-resistant Acinetobacter baumannii.*Antimicrob Agents Chemother*. 2007, 51: 3471-84 http://dx.doi.org/10.1128/AAC.01464-06.

[83]Tang, X., *et al.* Structural basis for bacterial lipoprotein relocation by the transporter LolCDE.*Nat Struct Mol Biol*. 2021, 28: 347-355 http://dx.doi.org/10.1038/s41594-021-00573-x.

[84]Andreu, J.M., *et al.* The Search for Antibacterial Inhibitors Targeting Cell Division Protein FtsZ at Its Nucleotide and Allosteric Binding Sites.*Biomedicines*. 2022, 10: http://dx.doi.org/10.3390/biomedicines10081825.

[85]Andreu, J.M., *et al.* The antibacterial cell division inhibitor PC190723 is an FtsZ polymer-stabilizing agent that induces filament assembly and condensation.*J Biol Chem*. 2010, 285: 14239-46 http://dx.doi.org/10.1074/jbc.M109.094722.

[86]Hasegawa, M., *et al.* Discovery of novel benzimidazoles as potent inhibitors of TIE-2 and VEGFR-2 tyrosine kinase receptors.*J Med Chem*. 2007, 50: 4453-70 http://dx.doi.org/10.1021/jm0611051.

[87]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.

[88]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.

[89]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.

[90]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.

[91]Grechishnikova, D. Transformer neural network for protein-specific de novo drug generation as a machine translation problem.*Sci Rep*. 2021, 11: 321 http://dx.doi.org/10.1038/s41598-020-79682-4.

[92]Ivanenkov, Y.A., *et al.* Chemistry42: An AI-Driven Platform for Molecular Design and Optimization.*J Chem Inf Model*. 2023, 63: 695-701 http://dx.doi.org/10.1021/acs.jcim.2c01191.

[93]Ferruz, N., *et al.* ProtGPT2 is a deep unsupervised language model for protein design.*Nat Commun*. 2022, 13: 4348 http://dx.doi.org/10.1038/s41467-022-32007-7.

[94]Ferruz, N. and Hocker, B. Dreaming ideal protein structures.*Nat Biotechnol*. 2022, 40: 171-172 http://dx.doi.org/10.1038/s41587-021-01196-9.

[95]Ferruz, N., *et al.* From sequence to function through structure: Deep learning for protein design.*Comput Struct Biotechnol J*. 2023, 21: 238-250 http://dx.doi.org/10.1016/j.csbj.2022.11.014.

[96]Mock, M., *et al.* AI can help to speed up drug discovery - but only if we give it the right data.*Nature*. 2023, 621: 467-470 http://dx.doi.org/10.1038/d41586-023-02896-9.

[97]Tsaban, T., *et al.* Harnessing protein folding neural networks for peptide-protein docking.*Nat Commun*. 2022, 13: 176 http://dx.doi.org/10.1038/s41467-021-27838-9.