Anticoagulant rodenticide novel candidates predicted by evolutionary docking

short title: anti-rodenticide 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

The evolutionary docking generation of high numbers of potentially novel anticoagulant rodenticides is described here. In particular, made-on-demand libraries were generated by randomly introducing small molecular variations into alphafold-modeled brodifacoum-rat VKORC1 binding-cavity. For evolution of the brodifacoum parent, criteria were mainly optimized to fit-dock the brodifacoum cavity. Libraries of specific brodifacoum-children were selected by pooling those predicting higher affinities (lower binding-scores) at various molecular weights. The selected brodifacoum-children were filtered for known toxicities, desirable high affinity to recently identified spanish resistance rat mutants and undesirable high affinities to human VKORC1. The flexible threshold-adjustable new chemotype brodifacoum-children libraries constitute an step forward towards further *in silico* fine-tuning to computationally reduce other possible unspecific off-target ecological impacts.

Keywords: VKORC1; brodifacoum; rats; evolutionary libraries; docking; rodenticides; human; mutants

Introduction

World-wide rodenticide genetic resistance among rat populations (*Rattus, sp*) are increasingly detected, mostly due to the intensive use of anticoagulants [1, 2]. For instance, resistance to warfarin (used at 58 mg/Kg LD_{50}) or coumatetralyl (16.5 mg/Kg), and both resistance and poison-death learning association when using stronger difenacoum (1.8 mg/Kg), bromadiolone (1.1 mg/Kg), difethialone (0.56 mg/Kg), flocoumafen (0.46 mg/Kg) or brodifacoum (0.26 mg/Kg)[3, 4] have ben reported. Genetic resistance constitute one of the main concerns of present rodenticide anticoagulants. The detection of anticoagulant toxic residues on several off-target predator wildlife species [1, 3, 5-7] which may include humans, add other important concerns. New anticoagulant rodenticides that could maximize on-target rat bioeffects including resistant mutants while minimizing off-target ecological impacts are necessary [8].

Most actual rodenticide anticoagulants are Vitamin K antagonists that inhibited **v**itamin **K** ep**o**xide **r**eductase **c**omplex **1** (VKORC1). Vitamin K is oxidized during carboxylation of the coagulation II, VII, IX, and X proteins. VKORC1s are found in most biological species. They are membrane enzymes that maintain the balance between active reduced Vitamin K and inactive oxidized Vitamin K. To do that, one VKORC1 cavity binds oxidized Vitamin K. Anticoagulant binding to the VKORC1 Vitamin K-binding cavity, inhibits oxidized Vitamin K binding, inactivating the carboxylation of the coagulation proteins. These causes animal death by internal bleeding.

The recent elucidation of the crystal structure of human VKORC1 complexed with Vitamin K and several anticoagulants, clarified the molecular details implicated in the oxred/redox reactions [9] (**Figure 1A,B**). After carboxylation, the inactive oxidized Vitamin K binds to the VKORC1 open top (Cys132 and Cys 135) and penetrates into the top-to-bottom central cavity formed by a 4-transmembrane helix bundle inside the cell membrane [9]. The coumarin/indandione heads of most anticoagulant molecules bind to the VKORC1 open top while their hydrophobic tails bind down into the central hydrophobic cavity through hydrogen bonds to Asn80 and Tyr139. Binding induces also the displacement of the Cys43-Cys 51 disulphide bond near the top cavity facilitating reductions of Cys132/Cys135/ and oxidized Vitamin K. After reduction, the top cavity opens up releasing reactivated Vitamin K to continue carboxylation. All actual anticoagulant rodenticides target VKORC1 by mimicking Vitamin K binding at similar binding-cavities and conformational changes[9].

Most resistances to rodenticide anticoagulants have been associated with genetic mutations (single nucleotide polymorphisms, SNPs) at the *vkorc1* gene that codes for VKORC1. Most amino acid mutations located at the top of the VKORC1 cavity [5, 10, 11, 12] caused anticoagulant resistance (**Figure 1A**). The prevalent European resistances mapped to *vkorc1* Y139F/C/S, L120Q, L128Q and F63C residues (first letter wild-type amino acid, last mutant)[5]. In Spanish regions, our group has identified mutations at S149I/T in urban (brown rat *Rattus norvegicus*) and E155K/Q in rural (black rat *Rattus rattus*) environments. These spanish mutations map down the VKORC1 cavity nearby the boundaries of the cell membrane [13] (**Figure 1A, orange spheres**). According to the alphafold modeling, all these spanish mutations cause a relative displacement of their α-helix, causing an small widening of the binding cavity (not shown).

A large number of new candidates need to be generated to search for new anticoagulant rodenticides. Those high numbers would be required to filter those candidates predicting high affinities to on-target rat VKORC1 and resistant mutants, and to exclude those predicting unspecific off-target effects (i.e, humans). Because it is not practical to experimentally screen such large numbers of candidates, alternative computational screenings were proposed [8, 14]. The recent availability of crystallized human VKORC1-brodifacoum models[9], the accuracy improvements in modeling protein structures by alphafold[15], the identification of brodifacoum-resistance mutants in spanish rats[13], the recent evolutionary docking possibilities and the preliminary anticoagulant computationally studies reported before [8, 13, 16], have been combined here to computationally generate and select new candidates to rat VKORC1 and their spanish mutants filtering out those that may affect humans.

The evolutionary docking belongs to the self-supervised generation methodologies that bypass the previous screening of large chemical data banks. Thus evolutionary docking can be used to augment the repertoire of specific ligands by targeting a defined docking-cavity using low hydrophobic criteria to increase their specificities to reduce off-target effects. To optimize the evolutionary docking, the brodifacoum chemotype was selected here as the initial anticoagulant parent from which randomly generate multiple children derivatives. The brodifacoum anticoagulant was selected because is among the most potent rodenticides due to its hydrophobic tail (logP = 8, penetrating the VKORC1 binding cavity) and because a crystallized binding model has been recently elucidated (**Figure 1**). Furthermore, the brodifacoum chemotype shares hydroxycoumarin-derived chemotypes common to other anticoagulant rodenticides including warfarine, coumatetralyl, bromadiolone, difenacoum, difethialone, and flocoumafen. Nevertheless, brodifacoum high hydrophobicity may contribute to increase unespecificity (by binding to any hydrophobic protein surfaces), which may translate in an increase of off-target effects to other biological species.

To generate multiple brodifacoum-children (libraries-on-demand), the recently released DataWarrior (DW) Build Evolutionary Library subprogram was chosen. DW first randomly generated multiple possible brodifacoum children and then selects those best fitting to a provided protein binding-cavity. Although, other alternatives are being developed using novel transformer technologies[17], in this work we have focused on generating a large number of brodifacoum-children molecules by DW. Those children candidates predicting high affinity (low docking-scores) for wild-type and mutant VKORC1 and low affinities for human VKORC1 were selected by DW docking. Filtration thresholds help to decide the limits most appropriated for selecting the candidates. These results will facilitate further computational ecological studies to reduce any predictable off-target bioeffects. Computationally predicting chemical synthesis pathway alternatives will be the next step before experimental confirmation studies. The identification of possible chemical synthesis pathways are being facilitated by the advent of modern retrosynthesis predictions[18], including those provided by the IBM RXN server (https://rxn.res.ibm.com/).

Computational Methods

Rat VKORC1, and resistance-mutant alphafold models

The amino acid sequences of rat VKORC1 were downloaded from the NCBI GenBank (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucleotide). Since brown (*Rattus norvegicus,* NP_976080) and black (*Rattus* rattus, XP_032748408) wild-type rats differed only at amino acid residue I90 or L90, respectively, and both amino acids had similar hydrophobic side-chains, the brown rat VKORC1 sequence was selected for this study. To be 3D modeled, the spanish mutations S149I/T and E155K/Q [13], were introduced into the VKORC1 wild-type amino acid sequences. For the 3D modeling of rat VKORC1, wild-type and spanish resistant mutant sequences were submitted to Sokrypton Colab Alphafold and models 1 selected (https://colab.research.google.com/ github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb)[15]. PyMol 3D similarities, were estimated by **R**oot **S**quare **M**ean **D**ifferences (RMSD)(**Figure 1C,D**). The crystallographic brodifacoum conformer was transferred by alignment of the brodifacoum-human

VKORC1 model[9] to each of the rat VKORC1 wild-type and mutant models by PyMol. The pdb files containing the VKORC1 ATOMS-TER and HETATM-END brodifacoum lines, required elimination of the CONECT lines for the VKORC1-brodifacoum docking cavity to be recognized by the DW docking.

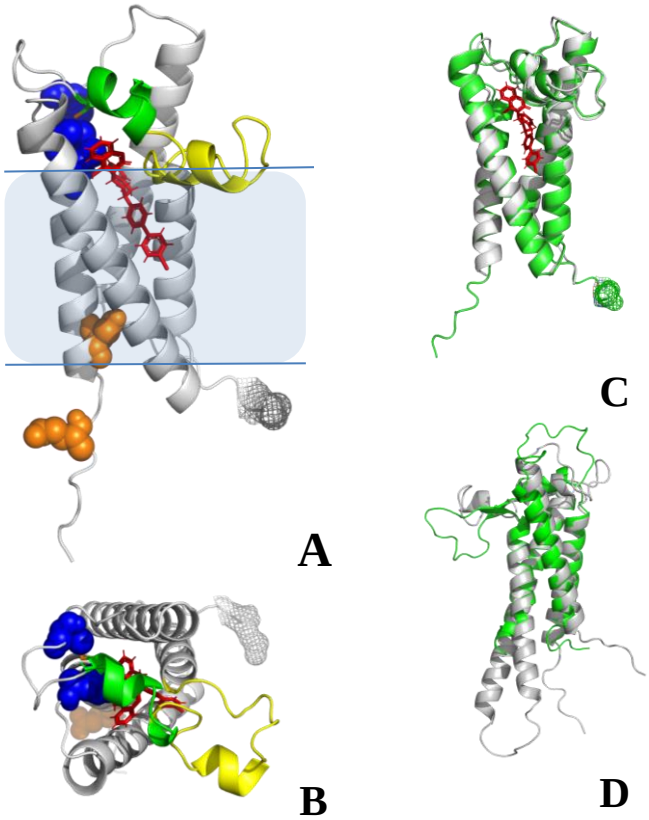

Figure 1
Rat VKORC1-Brodifacoum complex and alignments
A) Alphafold rat VKORC1-brodifacoum complex side view
B) Alphafold rat VKORC1-brodifacoum complex top view
Blue horizontal bars and background, cell membrane boundaries
Gray, VKORC1 cartoon of the amino acid carbons of the alphafold rat model.
Mesh amino acid, amino terminal MET1
Yellow cartoon, membrane anchor residues 60-80
Green cartoon, cap of the central cavity residues 52-59
Blue spheres, Cys43-Cys51 and Cys132-Cys135
Red sticks, brodifacoum
Orange, spanish mutations at residues 149 and 155
C) Alignment of alphafold rat VKORC1 (green) and human VKORC1-brodifacoum[9] models, RMSD=0.79 Å
D) Alignment of rat VKORC1 alphafold (green) and swiss[8] (gray) models, RMSD=1.71 Å

Generation of brodifacoum-children

To generate thousands of children, the DW optimized criteria will be described later in more detail. Briefly, DW randomly added small molecular modifications to a parent molecule generating a defaulted value of 128 raw-children molecules per generation. The raw-children were DW filtered by similarity to approved drugs or natural compounds. Additional fitting criteria were user-added to operate during each of the generation selections, such as docking-score (x4), molecular weight (x2), hydrophobicity logP (x1). The best 16 docking and fitting children were DW selected for the next generations. Each generation run stopped automatically when reaching their fitness plateau. A maximal number of 3 runs were repeated. Each run starts randomly from the parent. DW memorizes the children from any previous run to avoid duplicates. Memory requirements could be high, so that both heap memory (Java monitoring console) and RAM memories, were monitored in real time during evolution to reduce computer crashes[19, 20].

The fitted children were saved as *.dwar files for storage of the complete evolutionary data. The fitted children were additionally filtered by excluding predicted high and low toxic DW chemical properties (mutagenesis, tumorigenicity, reproductive interference, irritant, and nasty functions) using a home-designed macro. Final non-toxic children were saved as *.dwar and special *.sdf (vs3) files, maintaining all evolutionary information for further analysis[19, 20].

Computational programs employed

The "Build evolutionary library" and "Dock Structures into Docking Cavity" subprograms included into the DataWarrior (DW) [46] (dw550win.zip for Windows) were obtained (https://openmolecules.org/datawarrior/download.html) as described at DW and our previous work [19, 20]. MolSoft (ICM Molbrowser vs3.9Win64bit, https://www.molsoft.com/ download.html) was used for manipulating the *.sdf files, as described[19]. Origin (OriginPro 2022, 64 bit, Northampton, MA, USA) (https://www.originlab.com/) was used for calculations and drawings. The predicted structures 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 visualized in PyMol. A multithreading multi-core i9 (47 CPU) PCSpecialist (AMD Ryzen Threadripper 3960X) provided with 64 Gb of RAM (Corsair Vengeance DDR4 at 3200 MHz, 4 x 16 GB) (https://www.pcspecialist.es/) was used.

Results

To perform docking, the DW subprogram was chosen because it specifically targets a previously defined protein cavity rather than targeting wider grids like many other docking programs. Targeting only the docking-cavity favors the specific selection of the most fitting ligands provided that the docking-cavity was previously identified. In the present case, we took advantage of the recently accurate crystallographic identification of the brodifacoum binding-cavity of human VKORC1. The corresponding brodifacoum conformer coordinates at the human complex could be transferred to the aligned rat VKORC1 protein sequences after alphafold modeling.

The initial DW "Build Evolutionary library" subprogram generated numerous children docking to rat VKORC1-brodifacoum cavity from the whole brodifacoum molecule. The evolution from the whole brodifacoum molecule predicted more numerous and higher affinities (lower docking-scores) (**Figure S1, DWR6**), than two brodifacoum-derived fragments (coumarin and hydrophobic tail, respectively) (**Figure S1, DWR4, DWR5**). Therefore, the rest of the work was focused to the evolution of the parent brodifacoum molecule.

All next evolutionary work were performed by selecting the docking-score criteria with the highest relative-weight (x4) respect to other few criteria. Several brodifacoum-derived libraries were then generated by choosing different molecular weight target criteria with lower relative-weights (x2). To reduce the evolutionary selection of undesirable unspecific dockings, lower hydrophobicity (logP criteria <3) than brodifacoum were preferred, but selecting the lowest relative-weight (x1). Lowering the hydrophobicity should reduce hydrophobic-dependent toxicities, decrease potential unspecific off-targets, facilitate their aqueous handling and most probably, favor any resulting taste for rats if formulated in poison baits. Additional criteria to be optimized were the DW-dependent "create compound like: approved drugs or natural products", targeting different molecular weight preferences (user-dependent), and number of runs (user-dependent).

The here called raw-children were generated just after introducing random small molecular mutations and atom changes into the brodifacoum parent molecule. In the experiments performed here, the number of raw-children per run varied from 10044 to 26316 (**Table 1**). The percentage of raw-children fitting the evolutionary criteria mentioned above, varied from 7 to 15.9 %. Final non-toxic children percentages downsized those numbers from 8.8 to 66.6 %. Non-toxic children reflected the molecular weights and logP preferences selected by the user criteria (not shown). The downsized numbers of the final non-toxic children compared to the raw-children, suggested that to obtain the required few thousands of non-toxic children for considering additional constrains (i.e., more mutants, rat predators, ecological bacterial, animal or plant species, etc), the amounts of raw-children have to be maximized (**Table 1**).

Table 1
Characteristics of brodifacoum-children generated with different criteria

exp	Ch / Ge	<MW	R	Raw	Raw /R	Fitting children	Fitting %	Nontoxic children	Nontoxic %
***B15	16	<200	3	33513	11171	5344	15.9	3097	58.0
DWR10	8	<350	3	37432	12477	2749	7.3	1311	47.7
DWR11	8	<350	2	26487	13243	1885	7.1	904	48.0
B2	16	<350	2	44856	22428	6523	14.5	1137	17.4
B6	16	<400	3	78948	26316	11359	14.4	3298	29.0
***B13	16	<400	3	37828	12609	5489	14.5	3657	66.6
DWR9	8	<500	3	37205	12401	2611	7.0	712	27.3
***B10	16	<500	3	38997	12992	5568	14.3	3526	63.3
B7	16	<522	1	13168	13168	1861	14.1	164	8.8
***DWR8.3	8	<650	1	10044	10044	738	7.3	391	53.0

Initial parent: brodifacoum.sdf and rat VKORC1-brodifacoum-cavity.pdb
Common evolutionary criteria and relative weights(x): fitting to brodifacoum-cavity (x4); < MW (x2); logP<3 (x1)
Exp, name of the experiment and color-codes as in Figure 2
Ch / Ge, number of children (Ch) saved between generations of 128 children per generation (Ge)
<MW, children maximal molecular weight criteria during evolution
R, number of runs: number of times the evolution restarted from the initial parent to generate unique children
Raw, total number of randomly generated children,
Light blue vertical backgrounds, calculated values
Raw / R, number of raw-children generated per run, calculated by the formula: raw / R
Fitting children %, children fitting the brodifacoum cavity and their % calculated by the formula 100*fitting children/raw
Nontoxic children, non-toxic children and their % calculated by the formula: 100*non-toxic children / fitting children.
***, DW proposed general criteria= approved drugs. Unlabeled general criteria= natural products.

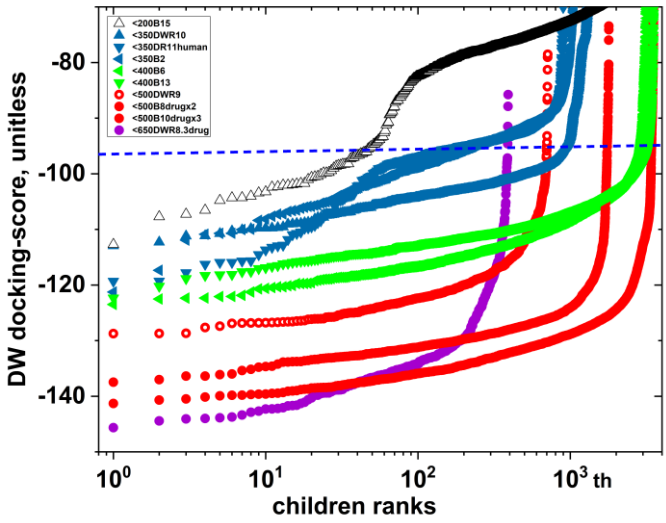

Figure 2
Rank profiles of brodifacoum-parent at different children Molecular Weight targets
DW Build Evolutionary Libraries were generated selecting different Molecular Weights (MW, g/mol) for target children, different runs and DW either with approved drugs or natural compounds criteria.
Brodifacoum molecular weight is 522 g/mol.
Horizontal dashed blue line, brodifacoum DW docking-score
Black triangles, MW <200 Dalton, 3 runs, ~drugs
Blue triangles, MW <350 Dalton, 2 & 3 runs, ~ natural, rat and human VKORC1
Green triangles, MW <400 Dalton, 3 runs, natural & drugs
Red spheres, MW <500 Dalton, 3,2,3 runs, natural &drugs
Purple spheres, MW <650 Dalton, 1 run, drugs

To maximize the number of raw-children, the number of children surviving each generation of 128 random compounds and the number of runs per experiment were increased. Increasing from 8 to 16 the surviving children, ~ doubled the number of raw-children. Variations in the percentages of non-toxic children were not significant (**Table 1**). Nevertheless, numbers >16 were not attempted because of the risk of excessively increasing memory with low-fitting children. On the other hand, increasing from 1 to 3 runs, proportionally increased the number of raw-children. Because of the random nature of each generation, quality and quantity of unique outputs and evolution changed from one run to another, increasing the diversity of the children using the highest number of runs. However, the number of runs was limited because of excessive memory resources and longer generation times. Thus, although most of the runs the RAM memory usage varied from 5-15 Gb, 50-60 Gb were also required, for instance by molecular weights >500 g/mol, difficult to dock cavities (unpublished work) or excessive number of runs. Since children with > 500 g/mol would be more unspecific and more difficult-to-be-synthesized, higher molecular weights were avoided. Most probably, more numbers of unique children could be generated using additional experiments of 3 runs each, but those were not attempted.

The children ranks from the evolutionary data predicted with different variables as described on **Table 1** were then compared (**Figure 2**). Those children predicting affinity ranges higher than brodifacoum (lower binding-scores under the **Figure 2, horizontal dashed blue line**), were more numerous when preferentially targeting molecular weights between <400 to <500 g/mol. There were no high differences in the profiles when comparing children obtained by either the approved drugs (***) and natural products (**Figure 3, green triangles & red spheres**). On the other hand, only 1.9 and 0.8 % of the children molecular structures were repeated among different experiments (**B6/B13,** <400 g/mol and **DWR9/B8/B10,** <500 g/mol, respectively). Therefore, we pooled all the children predicted when using <400 (**B6 & B13**) and < 500 (**DWR9, B8, B10**) g/mol. After pooling, children predicting > -90 docking-scores (low affinities), <250 g/mol molecular weights (low specificity), and duplicates were removed (library A). Library A contained 12037 brodifacoum children, predicting -90 to -141 DW docking-scores, 295 to 516 g/mol, and -2.1 to 7.4 logP.

Next, we asked what percentage of the children of library A predicted also high affinities to human VKORC1 (an example of off-target effects). Only 10.2 % of the children of library A predicted affinities greater in humans compared to rats (**Figure S2, blue spheres and blue horizontal dashed line**). These results suggested that at least those children should be excluded for further work because of the high risk of being toxic to humans. In this work, we have arbitrarily chosen a rat - human differential threshold of -25 DW docking-scores to be strict in reducing the possibilities of any undesirable anticoagulant human interferences. The -25 threshold eliminated 93.7 % of the children of library A, resulting in library B. Library B contained 747 children with maximal VKORC1 rat docking-scores (**Figure S2, red points**) and minimal affinities to human VKORC1 (**Figure S2, blue spheres at -25 docking-score**).

To additionally explore the children affinities to rat spanish mutants, their amino acid changes were introduced one-by-one into the sequences of wild-type rat VKORC1. The mutated sequences were then 3D modeled in alphafold. Then, the brodifacoum conformer from the crystal structure of the human VKORC1 complex, was transferred to each of the mutated rat VKORC1 to provide binding cavities to DW dockings. DW dockings were then performed with children of the library C to each of the mutated 3D rat models containing the transferred brodifacoum as required by DW. After docking all the mutants to the children of library B, the DW docking-scores including those from wild-type rat and human VKORC1 were laydown into the same *.sdf file to be compared. Multiple DW filters could be then applied using different thresholds for each of the VKORC1 docking-scores. To show a prove-of-concept example of this methodology, DW filters using stringent docking-score thresholds of < -100 were applied to retain higher-affinities to rat wild-type and spanish mutants (see the **150B17.dwar** file included as supplementary material). The same threshold was applied to skip higher-affinities to human VKORC1. In this prove-of-concept example, the use of the above mentioned filters downsized the 747 children library B to a 150 children library C (20 % of library B) (**Figure 3 and 150B17.dwar** file). Any other threshold values could be flexibly applied to obtain any desirable numbers of selected children. To adjust the final number of children candidates will be most probably required for instance, after adding to the Table at the dwar file any new computational data (i.e., VKORC1 off-target species) or experimental in vitro or *in vivo* data (i.e., binding to VKORC1 recombinant proteins or bait supplied to rats).

In the downsized highly-specific library C example, DW cluster analysis identified, at least, 4 representative scaffolds which contained children predicting 3, 4 or 5 rings (**Figure 4**). A graphic representation of all the 150 children of library C docked to rat VKORC1, predicted their docking targeted the upper part of the VKORC1 cavity. Most of the children occupied a similar cavity than the brodifacoum (**Figure 5**). The VKORC1 amino acids targeted by the docked 4 representative scaffolds, predicted maximal numbers of Hydrogen and hydrophobic bonds (**Table S2**).

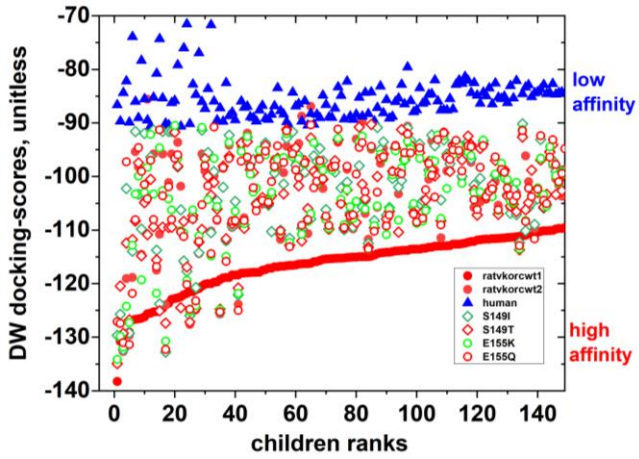

Figure 3
Rank profiles of 150 children of library C
The children were obtained by filtering the 747 children library C for VKORC1 high-affinities (low binding-scores) to rat wild-type and spanish mutants and low-affinity to human
Blue triangles, low affinity children to human VKORC1
Red spheres, high affinity children to wild-type rat VKORC1
Green & red open diamonds, high affinity children to rat VKORC1 S149 resistant mutants
Green & red open spheres, high affinity children to rat VKORC1 E155 resistant mutants

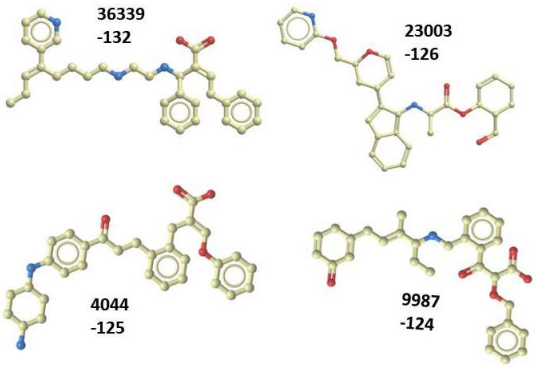

Figure 4
Children representative 2D molecules of library C clusters
Red, Oxygens. **Blue**, Nitrogens

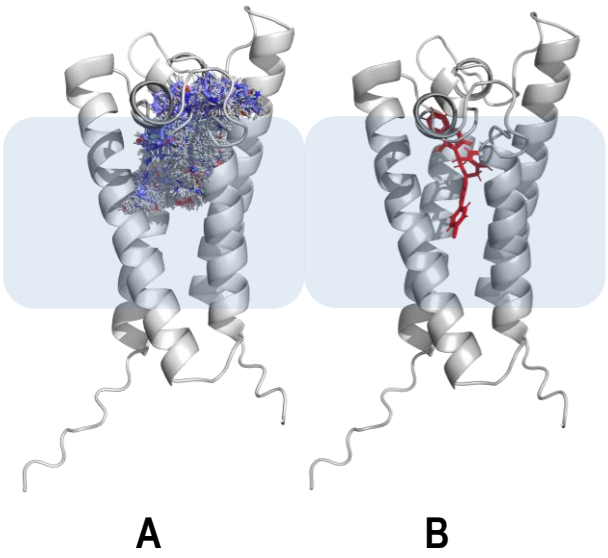

Figure 6
Mapping of 150-children library C (A) compared to the crystallographic brodifacoum parent (B, red) DW docked to wild-type rat VKORC1
Blue backgrounds, approximated location of cellular membranes

Discussion

This work describes a method to generate and explore a large portion of the chemical space by designing on-demand libraries using random evolution of the brodifacoum rodenticide anticoagulant. For potential rodenticide purposes, the new libraries of brodifacoum-children molecules were computationally selected by best fitting to the corresponding rat VKORC1 cavity. This work was mostly based in the recent 3D crystallographic elucidation of human VKORC1-brodifacoum complexes, in the Data Warrior (DW) Build Evolutionary Library subprogram and in our previous publications[19, 20].

In particular, brodifacoum children were generated by targeting the alphafold modeled rat VKORC1 cavity, with newly generated molecules with similar molecular weights (400-500 g/mol) and higher specificities (lower hydrophobicity). Molecular weight and hydrophobicity targeted values were so chosen to theoretically reduce the possible unespecificity of the generated children. Any other criteria could be used to generate alternative libraries since the chemical space to be explored for brodifacoum alternatives remains enormous. The elimination of those children which also predicted high affinities to human VKORC1, was one example of how further off-targets could be eliminated from the new molecules. Therefore, the objective of this work was to explore brodifacoum children with more specificity to reduce off-target effects (in this case only applied to humans) but conserving docking to the rat and their spanish mutants. In particular, the restricted docking-cavity of DW favored the generation of enriched numbers of random brodifacoum-derived children which automatically were selected for best docking-scores (evolutionary docking). Around 12000 children brodifacoum alternatives with potential anticoagulant activities could be generated. Such high numbers of potential brodifacoum-derived anticoagulant rodenticides would had been impossible to predict by any other computational screening programs (AutoDockVina, DEEPScreen, seeSAR, etc).

For a prove-of concept example to show the flexibility of this approach, the numbers of potentially useful children molecules were downsized using stringent thresholds to maximize affinities to rat wild-type and spanish mutants and minimize human off-targets. Despite using this evolutionary and filtering powerful approach to identify numerous highly specific anticoagulant candidates, most of the corresponding vast chemical/chemotype space [21, 22] fitting the VKORC1 cavity still remains unexplored. A similar large number of anticoagulant candidates would have been impossible to retrieve by more traditional screening of the largest public chemical libraries available (Mcule, chemSpace, Zinc, PubChem, Chembl, etc). The method employed here generated large focused libraries which could be easily adapted to any additional variations by selecting other threshold values and/or starting from other parents or fitting other criteria (i.e, other targeted molecular weights, hydrophobicity, number of Nitrogen, number of Oxygens, etc)

The main known limitations of the library predictions mentioned above include for example, having explored only a brodifacoum-limited VKORC1 docking-cavity, elimination of water molecule contributions to dockings, and/or using particular values for the thresholds. Additionally, any exploration of the vast chemical/chemotype space possibilities, including the ones reported here, will always be limited. Furthermore, predictions did not included unexpected *in vivo* physiological variables. On the other hand, once a few candidates would be defined after excluding off-targeting other biological species, chemically synthesis will be necessary for any *in vitro or in vivo* validation. In this aspect and on due course, the identification of possible chemical synthesis pathways would be facilitated by the developing modern retrosynthesis predictions[18], actually including those provided by the IBM RXN server (https://rxn.res.ibm.com/). All the above mentioned obstacles still limit the accuracy of the predictions made here.

The described results, identified new chemotypes predicting low nanomolar (evidences to be described) brodifacoum-children with high specificity while conserving its targeting to the rat VKORC1-cavity. The results remain to be confirmed after other off-target impacts would be minimized by further computational filters and their subsequent chemical synthesis [23] would allow experimental validation. Possible off-target ecologically important species may include not only known predators but also soil bacteria, fungus and plants, since the VKORC1 enzyme is widely distributed among biological species. Therefore, further work will include selection of those molecules with minimal affinity to predator species, chemical synthesis and experimental validation..

Supporting information

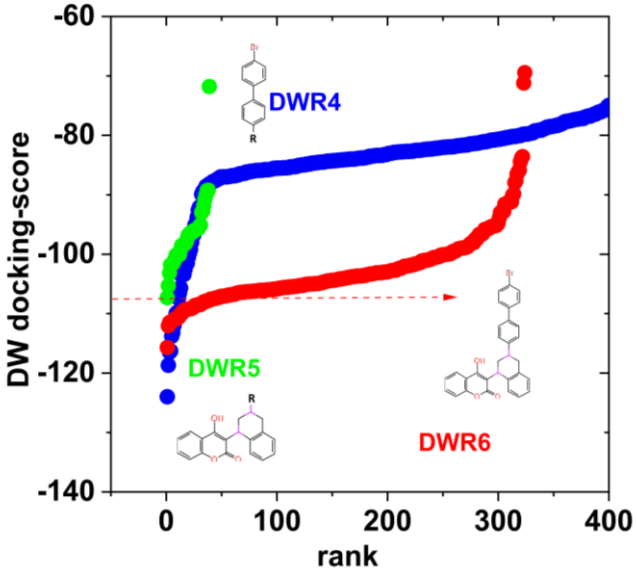

Figure S1
Rank profiles of children from brodifacoum and two fragments
The represented children were filtered by known toxicities. The 2D drawings were the structures of the parents used for each of the evolutionary dockings. Criteria used for evolution: 50 generations, 1 run; fitting to the rat VKORC1 brodifacoum-cavity (x4); molecular weight < 650g/mol (x2); logP<3 (x1).
Horizontal dashed red line, DW docking-score of brodifacoum.

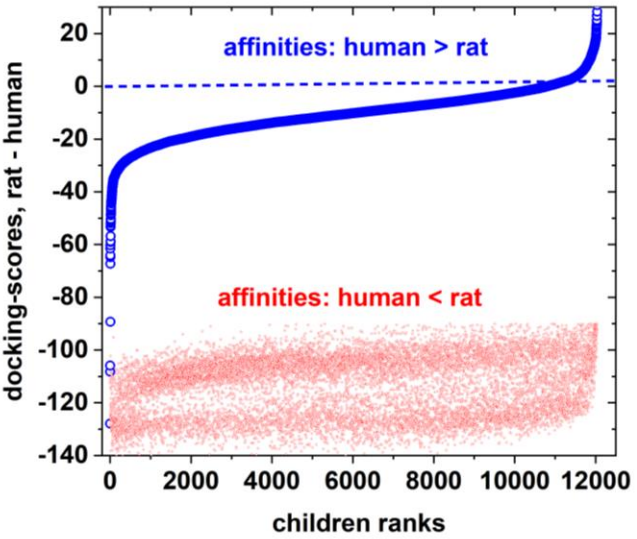

Figure S2
Differential rank profile of rat - human VKORC1 of the brodifacoum-children library A
Low docking-scores corresponded to high affinities and *viceversa*.
Blue open circles, differential docking-scores calculated by the formula rat - human
Red points, docking-scores of the rat VKORC1

Table S1
Resume of DW docking-scores of representative clustered VKORC1 chemotypes

ID	wt	S149I	S149T	E155K	E155Q	human
36339	-132	-135	-138	-138	-137	-99
23003	-126	-113	-111	-112	-121	-73
404	-125	-109	-115	-115	-115	-99
9987	-124	-124	-114	-117	-120	-97

Each of the ID numbers were automatically assigned by the DW during evolution

Table S2
Amino acid residues of RAT VKORC1 predicting interactions with representative leads

position	Aa	Aa	brodifacoum	36339	23003	9987	4044
19	G	Gly					
22	L	Leu					
23	S	Ser					
26	A	Ala					
30	K	Lys					
54	V	Val					
55	F	Phe					H
58	R	Arg					
59	W	Trp					
60	G	Gly					H
61	R	Arg					
62	G	Gly					
63	F	Phe					
77	N	Asn					
78	G	Gly					
79	S	Ser					
80	D	Asp	H				
81	S	Ser			H		
82	I	Ile					
83	F	Phe					
84	G	Gly			H		
85	C	Cys					
87	F	Phe					
88	Y	Tyr					
109	S	Ser					
113	S	Ser					
114	V	Val					
115	A	Ala					
116	G	Gly					
117	S	Ser					
118	L	Leu					
120	L	Leu					
124	L	Leu					
134	V	Val					
135	C	Cys					
138	T	Thr				H	
139	Y	Tyr					
142	N	Asn					

Aa, Amino acid residues of the docking-cavity of wild-type rat VKORC1.
Colored rectangles, amino acid residues predicted as ligands by LigPlus
H, predicted Hydrogen bonds by LigPlus.

Funding: This research was not externally funded

Competing interests
The author declares no competing interests

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

References

[1]Nakayama, S.M.M., *et al.* A review: poisoning by anticoagulant rodenticides in non-target animals globally.*J Vet Med Sci.* 2019, 81: 298-313 http://dx.doi.org/10.1292/jvms.17-0717.

[2]Nakayama, S.M.M., *et al.* Avian interspecific differences in VKOR activity and inhibition: Insights from amino acid sequence and mRNA expression ratio of VKORC1 and VKORC1L1.*Comp Biochem Physiol C Toxicol Pharmacol.* 2020, 228: 108635 S1532-0456(19)30351-5 [pii], http://dx.doi.org/10.1016/j.cbpc.2019.108635.

[3]Fisher, P., *et al.* Anticoagulant Rodenticides, Islands and Animal Welfare Accountancy.*Animals (Basel).* 2019, 9: ani9110919 [pii], http://dx.doi.org/10.3390/ani9110919.

[4]Tasheva, M. Anticoagulant rodenticides.*WHO IPCS Environmental Health Criteria.* 1995, 175: 124p ISBN9241571756, ISSN 0250-863X.

[5]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.

[6]Chetot, T., *et al.* Vitamin K antagonist rodenticides display different teratogenic activity.*Reprod Toxicol.* 2020, 93: 131-136 S0890-6238(20)30016-2 [pii], http://dx.doi.org/10.1016/j.reprotox.2020.02.003.

[7]Lettoof, D.C., *et al.* Toxic time bombs: Frequent detection of anticoagulant rodenticides in urban reptiles at multiple trophic levels.*Sci Total Environ.* 2020, 724: 138218 S0048-9697(20)31731-9 [pii], http://dx.doi.org/10.1016/j.scitotenv.2020.138218.

[8]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.

[9]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.

[10]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.

[11]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.

[12]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.

[13]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.

[14]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

[15]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.

[16]Ferencz, L. and Muntean, D.L. Identification of new superwarfarin-type rodenticides by structural similarity. The docking of ligands on the vitamin K epoxide reductase enzyme's active site.*Acta Universitatis Sapientiae Agriculture Environment.* 2015, 7: 108-122 http://dx.doi.org/10.115/ausae-2015-0010.

[17]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.

[18]Schwaller, P., *et al.* Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy.*Chem Sci.* 2020, 11: 3316-3325 http://dx.doi.org/10.1039/c9sc05704h.

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

[20]Coll, J.M. New star-shaped ligands generated by evolutionary fitting the Omicron spike inner-cavity.*ChemRxiv.* 2023: https://doi.org/10.26434/chemrxiv-2023-v8gqj.

[21]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.

[22]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.

[23]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.