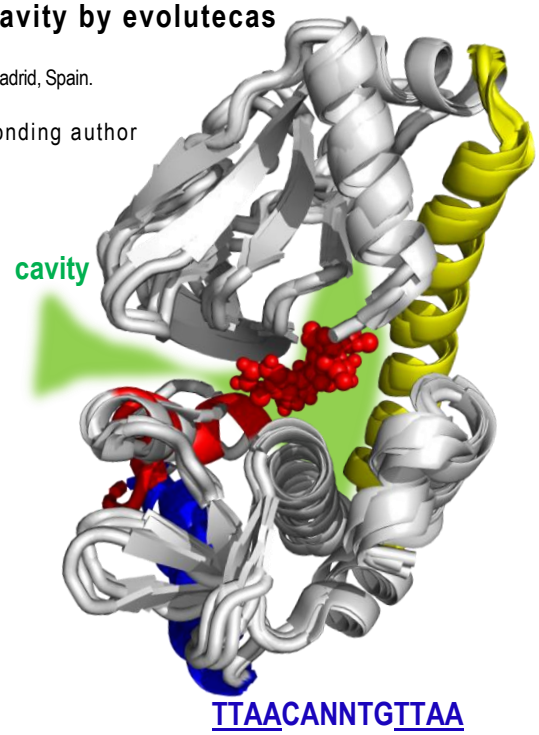
Targeting *Listeria* virulence-regulon glutathion-cavity by evolutecas

short title: anti-Listeria evolutecas

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



Abstract

Computational explorations are described generating small peptide and acetic-polymer evolutecas targeting one of the virulence regulons of *Listeria monocitogenes* (*Lm*). *Lm* are among the most important multiple-resistant food-borne infective bacteria whose antibiotic resistances raise worldwide concerns. The glutathion cavity of the *Lm* **p**ositive **r**egulatory **f**actor **A** (prfA) virulence regulon, was computationally targeted here because of the importance of the glutathion tri-peptide-dependent prfA activation of many virulent *Lm* genes. Computationally screening of alternative tri-,tetra-, and penta-mer peptides and innovative-acetic polymers and evolutecas generated from the top-screened molecules, were employed to optimize their fitting to glutathion-prfA cavities. Combining library-screening and children generation by co-evolution, several novel top-peptides and acetic-derived polymers were generated predicting low-toxicities (to reduce any known undesirable side-effects) and low nanoMolar consensus affinities (to maximize their specificities and accuracies). However, their possible inhibition of prfA virulence remains to be demonstrated. To continue these explorations, additional chemical synthesis and *in vitro* experimental validation, are strictly required to demonstrate any desirable interference with *Lm* virulence and/or biological functions.

Keywords: prfA; virulence regulon; Listeria monocitogenes; co-evolutionary docking; antibiotics; glutathion; acetic acid derivatives; computational restrictions

The prfA regulon of *Lm* virulence

Listeria monocitogenes (*Lm*) are intracellular Gram-positive bacteria causing increasing worldwide concerns of these food-borne infections. The increasing emergence of hypervirulent *Lm* strains with pan-antibiotic resistances further limits any previously antibiotic-treatment options. Unexplored targets may be one of strategies to propose any new *Lm* virulent antibiotics.

Among computationally possible unexplored *Lm* targets, virulence regulons such as prfA (**p**ositive **r**egulatory **f**actor **A**) has been chosen because it responds to the surrounding environment to induce the expression of hundreds of *Lm* virulence genes to generate biofilms, infection and antibiotic resistances.

The prfA protein monomer is coded by 237 amino acids that becomes active by dimerization to bind 14 bp palindromic DNA operators distributed among the promoters of many virulent genes (**Figure S3, blue letters: TTAA**CANNTG**TTAA,**). The tri-peptide glutathione (ECG in single letter amino acid code) allosterically stabilizes the prfA-dimer DNA-binding activity to coordinate the expression of many virulent genes.

Comparison of the 3D structures of prfA in the absence (2beo.pdb)[1] and presence (5lrr.pdb)[2] of ECG, predicted an ECG-binding cavity (**Figure S3, green cavity**). Each ECG tripeptide interacts with an hydrophobic cavity formed by ~ 60-70 (^{63}Y, ^{67}F), 110-125 (^{126}Y) and other residues[2]. Additionally, each prfA monomer contains an N-terminal domain (residues 1–108), a main long α-helix which interfaces with the other monomer (~ 109–137), and **h**elix-**t**urn-**h**elix DNA-binding HTH segment (170-196), similar to other prokaryotic DNA-binding proteins (flexible loop 170-177 + connecting loop 174-184 + α-helix 183-196)(**Figure S3, red cartoons**). The binding of ECG or the prfA G^{145}S mutation, induce α-helix conformations on the 170-177 prfA residues, contributing to the stabilization of the DNA-binding.

Although the binding affinity of ECG to prfA is ~ 1000 nM, their eukaryote intracellular concentrations in host eukaryote cells are ~ 10 mM, which could compensate the ECG low affinity. Nevertheless, the glutathion synthase enzyme of *Lm* generates additional ECG tripeptide molecules from any environmental Cystein and/or peptides-containing Cystein. The prfA 3D structures ± ECG are very similar having an overall carbon backbone rmsd difference of only 0.25-0.35 Å^2. However, when in the presence of ECG at the *Lm* intracellular space, the binding affinity of prfA to DNA increased to ~ 50 nM. In contrast, to Cystein-containing peptides, no-Cysteine-containing peptides induce prfA inhibition through steric blockade of the ECG binding cavity[3]. The strength of the prfA DNA binding, depends on the ratio between Cystein-containing (activating) *versus* no-Cystein (inhibitory) peptides. Such mechanism may provide a control of expression of unneeded virulence when outside a Cystein-containing environment. Thus, the ECG-mediated prfA conformational changes to activate DNA-binding can be inhibited by no-Cystein peptides such as LL (6hck.pdb) or EVFL (8cb7.pdb), which sterically blocked the ECG-binding cavity[3]. The hydrophobic side-chains of the ECG binding cavity would probably require small molecular weight and highly-hydrophobic inhibitors. However, since particular peptides, could also lock the prfA in an activated conformation similar to the G^{145}S mutation, experimental validation is critical to evaluate any possible inhibitions.

Therefore, the present computational search, is based on several theoretical working hypothesis. The candidates should:

i) contain no Cysteins.
ii) sterically block any ECG-binding cavities in a inactivating manner,
iii) have small molecular weight sizes to fit the ECG-binding cavity,
iv) predict higher docking affinities than ECG (i.e., low nanoMolar), and
v) require chemical synthesis for experimental validation

Compounds predicting the first four properties can be computationally predicted. However, the fifth requirement requires experimental validation.

To generate large numbers of such candidates, the following strategy was:

a) screening all tri-, tetra-, and penta-peptide possible libraries, except those peptides containing Cystein

b) screening of tri-acetic novel polymers mimicking peptide bonds but eliminating terminal Nitrogen, and

c) generating evolutecas from the top-parents by the "build evolutionary library" co-evolutions of ***D**ata**W**arrior **B**uild **E**volutionary **L**ibrary* (DWBEL)[2-5].

DWBEL algorithms were employed before as one of the best alternative to screening extralarge drug-like banks[4, 5] or predicting deep-learning docking models from protein sequences [6-8] or from recent RFdiffusion[9-13] and/or on recent development Boltz-2 alternatives including affinity and synthetic-pathway predictions[14, 15]. Based on our own preliminary RFdiffusion applications to snake venoms[16] and monkeypox[17], our attempts to generate RFdiffusion small peptides fitting the ECG cavity of *Lm* prfA were unsuccessful, most probably because their docking cavities were inner at the prfA molecule and not in the surface (not shown). The DWBEL co-evolution alternative was/is a powerful method to generate large numbers of small molecular weight candidates through mimicking natural evolution. DWBEL had computationally generated higher affinity gains of non-toxic drug-like alternatives when targeting different protein-ligand pairs. For instance, new antibiotics for resistant *Staphilococcus*[18], Abaucin-derivatives against *Acinetobacter*[19], non-human anticoagulant rodenticides[20], monkeypox Tecovirimat-resistant mutants[21], inner-cavity SARS omicron[22], inflammatory coronavirus ORF8 protein[23], prokaryotic models of human K$^+$ channels[24], inner-cavity of VHSV rhabdovirus[25], malaria circumsporozoite protein[21], RSV resistant mutations[26] anti-HIV-Vif A3G[27] and A3A hypermutational human metastasis[28].

Combining computational screening of peptide and acetic-polymer libraries with evoluteca generation to generate children novel compounds by co-evolutionary docking, were only the first computational steps to propose new predictions of candidates for prfA inhibition of virulence. Chemical synthesis and strict *in vitro* experimental validations remain to be performed before any possible applications to fine-tune the possible control any *Lm* virulence and/or biological functions.

Computational predictions

Comparisons among the crystallographic *Lm* prfA monomers of apo[2], and complexed with ECG[1], EVFL[3] and LL[3] , predicted that the inhibitory EVFL and LL no-Cystein peptides sterically blocked the conformational changes on the ~170-177 residues, otherwise occurring in the presence of ECG and/or Cystein peptides (**Figure S3, red cartoons**). Therefore, to compete with ECG-binding, screening for higher affinities was performed among tri-peptide libraries containing all except peptides containing Cysteins. One library of 6859 possible tri-peptides excluding Cysteins, was generated by home-designed Python scripts and ChemAtom 3D modeling. The ADV docking to apo prfA resulted in most tri-peptides predicting higher affinities than ECG (**Figure 1A, blue dashed-line**). The highly hydrophobic top-tri-peptides FWW, FWF, FFW predicted the highest ~ 1 nM affinities (lower docking-scores) (**Figure 1A, cyan circles**). These predictions confirmed that the hydrophobicity favoured the affinities of possible inhibitors to the ECG binding-cavity, as previously proposed from crystallographic studies[2] .

All possible tetra-peptides derived from the top-tripeptide FWW were then generated adding all possible N- and/or C-terminal amino acids, except Cysteins. The corresponding ADV computational predictions could not identify any significant affinity improvements on the tetra-peptides (**Figure 1A, violet-cyan circles**). However, PyMol visualization of their docked complexes revealed that most of the top-tetramers coded for the same amino-terminal W (Tryptophan) and targeted the downward sites of the ECG-cavity. In contrast, the tri-peptides at their carboxy-terminal amino acids were variable (i.e., W,E,Y,G,D,A,I,Q) and targeted the upper sites of the ECG-cavity (not shown). These observations suggest that carboxy-terminal amino acids could be added to the tetra-peptides to improve their docking affinities. However, the addition of all possible amino acids, except Cystein to the C-terminal residues of WFWW, failed to generate significant affinity improvements (**Figure 1A, violet circles**). Representative top-penta-peptides were for instance: WFWWW, WFWWE, WFWWY, WFWWG, WFWWD, WFWWA, WFWWI, WFWWQ.

To explore for additional tetra-peptide-derivatives predicting affinity improvements, their C terminus was extended by small atom combinations. For that the WFFW top-tetra-peptide was conserved during DWBEL as parent and partial co-evolutions were only performed of its carboxy-terminus carbon atom to derive children (WFFW-). The prfA complexed WFFW-E cavity (containing one of the largest cavities targeted by the penta-peptides), was chosen as the fitting cavity for selection during co-evolution. Such partial co-evolutions, randomly generated 19023 raw-children of which 2011 were selected by the program and filters as non-toxic fitted-children WFFW-X derivatives. However, none of these co-evolved derivatives predicted significative affinity improvements when compared to the tri-, tetra- or penta-peptides (**Figure 1A, blue circles**). Despite conserving a high number of hydrophobic side-chains and multiple molecular variations at their C-terminus, no improvement of affinities were obtained.

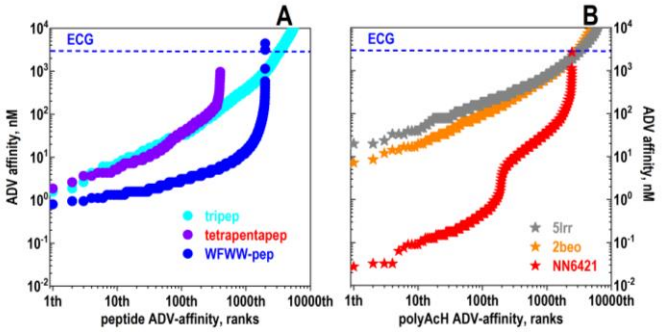


Figure 1
ADV-prfA docking rank profiles of non-Cystein peptides and tri-acetic derivatives

A) the 3-, 4- and 5-mer peptides were generated in single letter code by home-made Python scripts. The 4- and 5-mer computational predictions were mixed for simplification. The single amino acid letter code were converted to amino acid sdf files using a dictionary of amino acid smiles by the ChemAtom Converter web server (https://datascience.unm.edu/tomcat/biocomp/convert).

B) Tri-acetic acid derivatives were generated by home-made Python scripts using a dictionary of single tri-acetic polymers corresponding to the amino acid smiles mostly by eliminating the amino of the 20 amino acids, excluding those corresponding to Cysteins. In both peptides and acetic-derivatives, their 2D conformations were maintained by generating their DW mmff94s+ conformers before ADV docking. Most of the resulting 3D conformers were ADV docked to 2beo.pdb (apo-prfA) monomers, except when indicated to 5lrr.pdb (ECG-conformation without ECG). PyRx ADV docked affinities were finally rank-ordered.

Blue dashed horizontal line, ECG (glutathion) ADV binding affinity to 2beo.pdb
A) 3-4-5-mer peptides (tripep, tetrapentapep) and DWBEL WFWW-derivatives to apo prfA 2beo.pdb
B) 3-Acetic acid polymers and NN6421 DWBEL-derivatives to apo prfA 2beo.pdb and 5lrr.pdb±ECG
Cyan circles, all possible 3-mer peptides, except Cystein docked to prfA (6859 peptides).
Violet circles, all possible 4- and 5-mer peptides derived from FWW tripeptide (399 peptides) docked to prfA
Blue circles, DWBEL derivatives from WFWW- tetra-peptide NN364 docked to prfA.
Gray stars, all possible tri-acetic polymers docked to prfA 5lrr.pdb ± ECG (6859 polymers)
Orange stars, all possible tri-acetic polymers docked to apo prfA 2beo.pdb (6959 polymers)
Red stars, DWBEL derivatives from the top NN6421 tri-acetic polymer docked to apo prfA

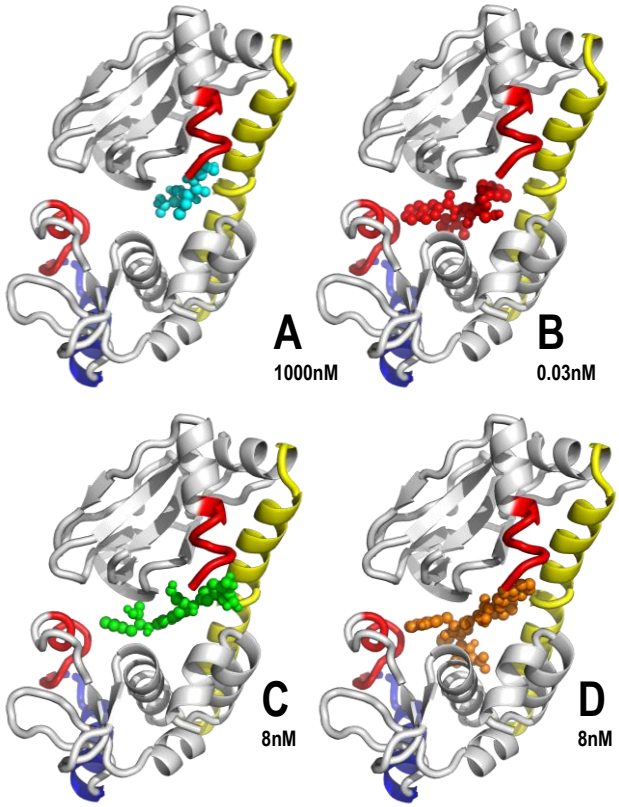


Figure 2
Visualization of docking cavities targeted by ECG (A), NN331 DWBEL-derivative (B),FWW (C) and WFWWE (D) to the apo prfA conformation

ADV affinities were approximated for C and D by their mean ± 1 nM (n=2), those for A and B were ± 500 nM and 0.01 nM (n=2), respectively.
A) ECG. main cavity binding/docking ECG
B) NN331 DWBEL-derived top-children from NN6421.
C) tri-peptide FWW (NN1424).
D) penta-peptide WFWWE (NN 364)

To explore additional alternatives to computationally improve docking affinities to the ECG-binding cavity, innovations were developed by generating "tri-acetic polymers". The so called tri-acetic polymers were similarly designed by constructing a dictionary of " no amino" acid-likes without terminal Nitrogen but contained the same 19 amino acid side-chains (except Cystein). Their targeting to the cavities of the prfA monomers, predicted similar affinities whether the cavities were derived from the ±ECG models or the apo prfA (**Figure 1B, gray** and **orange stars,** respectively), but did not improve the ~ 1-10 nM affinities obtained by the corresponding tri-peptides (**Figure 1A, violet-cyan circles**). However, the PyMol visualization of top-tri-acetic polymers in complex with prfA, identified additional nearby amino acids targeting distal sites at the ECG-binding cavity (**Figure S1B**). Those observations suggested that co-evolution could help to generate derivatives predicting higher activities.

Therefore, the top-NN6421 tri-acetic polymer ("WSW"-build with the corresponding side-chains) was selected as a parent molecule and prfA cavity to generate their DWBEL co-evolution children. The DWBEL program randomly generated 41622 raw-children of which 2437 were selected as being non-toxic and cavity-fitted-children. The rank-ordered ADV affinites of the generated children predicted top non-toxic fitted-children increasing 100-1000 fold their affinities (**Figure 1B, red stars**). Several top-children predicted maximal affinities ~ 0.03 nM (**Figure 1B red stars**), >100-fold higher than those obtained by the tri-, tetra-, penta-peptides and/or their peptide-derivatives (**Figure 1A, cyan, violet, blue**). Because these novel acetic-polymer derivatives did not contained any Cystein side-chains and occupied longer prfA cavities (**Figure 2C,D**) than ECG (**Figure 2A**), these compounds may be sterically blocking ECG-binding. Possible *in vitro* side-effects may also be reduced, because such high affinities may allow their use at low concentrations. Nevertheless compared with the peptides, tri-acetic derivatives may be more difficult to synthesize. Chemical synthesis and strict *in vitro* experimental validation remains to be done for any possible biological function studies.

Top-peptides and top-acetic-polymers predicted higher affinities to prfA than the ECG affinity (~ 500-fold or ~ 30000 higher affinities, respectively) (**Figure 1AB**, **blue horizontal dashed line**). The comparison of 4 Å nearby amino acids of some representative top-peptides and top-acetic-polymer derivative conformers, predicted that all of them always included the two prfA main sites of the ECG-binding cavity (~ 60-70 and 110-125 residues) (**Figure 3, ECG vertical cyan rectangles**). Additional prfA residues were targeted by both top-peptides or top-acetic-polymers such as the amino-terminal ~10 first amino acids, 140-165 or around the 230 carboxy-terminal residue (**Figure 3, different colors**).

These top-molecules are only some of the many other possible solutions that were identified among those predicting higher prfA affinities than ECG, suggesting they could successfully compete for binding to prfA. However, although easier to synthesize, top-peptides contain a large proportion of highly hydrophobic side-chains that may be difficult to handle and/or cause excesive unspecific bindings to other proteins. Given the existence of a vast chemical space with many other alternatives, perhaps and despite their chemical synthetic difficulties, the acetic-polymers predicting lower hydrophobicities with higher affinities, may better fit any experimental steps. Only *in vitro* experimental validation could demonstrate any *Lm* virulence inhibition.

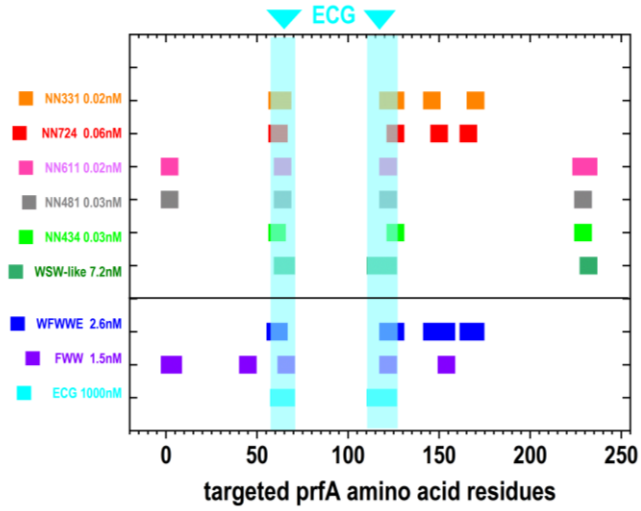


Figure 3
prfA Amino acids targeted by ECG and top-peptides and top-acetic-polymers
Nearby amino acid positions within 4 Å of docked top-molecules were computationally predicted by home-designed PyMol/Python *.py scripts and drawn in Origin templates.
X-axis, amino acid position numbers of the prfA amino acid monomer (left amino to right carboxy positions).
Lower part of the Figure, ECG and representative top-tri-(FWW) and penta-(WFWWE) peptides.
Upper part of the Figure, WSW-like tri-acetic parent and selected representatives of its top-DWBEL co-evolved children ranked-ordered by their NN numbers. Their approximated affinities were calculated by ADV docking.
Cyan vertical rectangles, prfA amino acids targeted by ECG docking.
Violet rectangles, prfA amino acids targeted by docking of the tri-peptide FWW.
Blue rectangles, prfA amino acids targeted by docking of the penta-peptide W**FWW**E (derived from FWW).
Other color rectangles, prfA amino acids targeted by top-acetic-polymer derivatives (DWBEL from WSW-like tri-acetic-polymer)

Computational limitations

These predictions are strictly limited to computational docking made by including numerous simplifying constrains (i.e., fixed amino acid side-chains in ligands and protein target, absence of water interactions which may interfere the docking processes, exclusively rely on approximated maximal affinites using only one or two consensus docking programs, targeting one of the possible alternative cavities, etc). Those predictions require experimental validation of possible biological functions. Small peptides may be the easiest to synthesize for *in vitro* experimental tests. Despite limitations, the predicted prfA inhibitor candidates have not been described before and may also provide examples for exploration and development of similar or related computational strategies. However, any drug-like inhibitors of *Lm* virulence strictly depend on additional *in vitro* and *in vivo* experimental validation.

Conclusions

New drug-like small molecular weight compounds predicting no-toxicities together with low-nanoMolar affinities to the ECG-binding cavity of *Lm* prfA monomeric models of one of the regulons controlling virulence have been computationally generated. ADV docking screening and DWBEL co-evolution generation were combined to identify two kinds of small molecular candidates: peptides and acetic-polymers. Before any *in vitro* experimental validation tests, peptide and/or combinatorial chemical synthesis, are required to obtain small amounts to test these candidates.

Supplementary information

Table S1
Computational software and hardware

name	Version/year	Main use and references	url
DataWarrior	6.4.0 Windows	Evolutionary docking[34] Commercial Alipheron Hyperspace	https://openmolecules.org/ datawarrior/download.html)
Toxicity & nasty NTN macro	2023	Eliminate residual toxic / nastic fitted-children after co-evolution[19]	
Toxicity Risks	2023 Updated DataWarrior	Minimize toxic / nasty raw- children to increase specificity[29]	https://openmolecules.org/ datawarrior/download.html)
Babel & AutoDockVina	Home-adapted PyRx 098/1.0	Mmff94s force-field minimization & 2D conservation	https://pyrx.sourceforge.io/
2D geometry conservation	2023	DW mmff94s+ force-field minimization [57]	
ChemAxon converter	server	Convert single amino acid sequences in 3D molecules	(https://datascience.unm.edu/tomcat/biocomp/convert).
MolSoft	3.9 Win64bit	Easiest manipulations of sdf files	https://www.molsoft.com/download.html
PyMol	2.5.7.	Visualization & PyMol-Python scripts	https://www.pymol.org/this work
Clone Manager	9th professional	Protein alignment	
OriginPro	20222024	Calculations and Figures Macros to handle large numbers of data	https://www.originlab.com/this work
AMD Ryzen i9 computer	4 DDR4 x 32[23] Gb memory	47 CPU Computational hardware	https://www.pcspecialist.es/

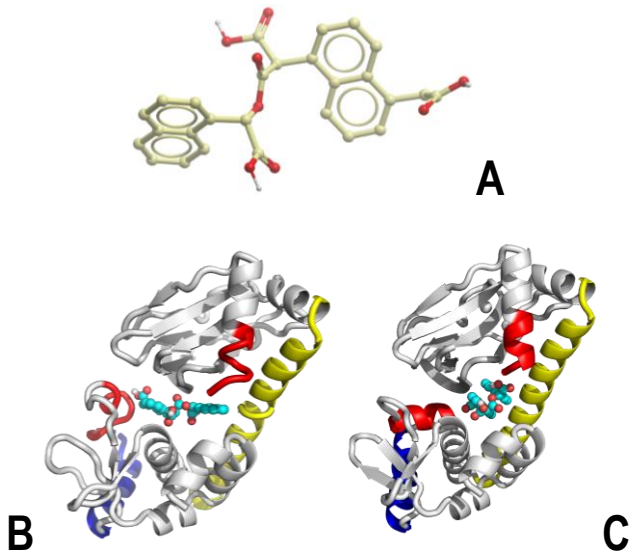


Figure S1
Derivative top-acetic-polymer NN6421 (A) docked to apo prfA (B) and prfA±ECG (C)
The top-tri-acetic-polymer NN6421 was docked to prfA monomers in the apo (2beo.pdb) and the ±ECG (5lrr without ECG). Comparison of NN6421 complexed to prfA in the apo (**B**) with the ±ECG (**C**) conformations could explain the sterically inhibition of NN6421 binding by the ECG induced 170-177 α-helix (**C, red α-helix**) . In contrast, the NN6421 conformer remains compact when docking to prfA monomers in their ECG-induced/empty cavities (Figure S1C). .
Grey cartoons, carbons of the prfA monomer.
Yellow cartoons, main α-helices 108-137 residues.
Blue cartoons, partial DNA-binding HTH 183-196 residues.
Red cartoons, prfA 170-177 residues changing after ECG binding.
Different color spheres, derivative NN6421.

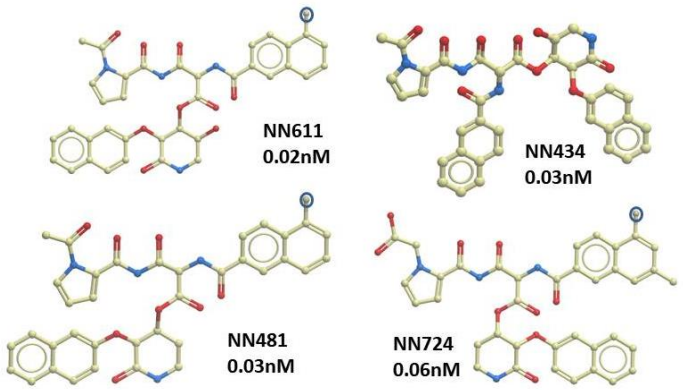
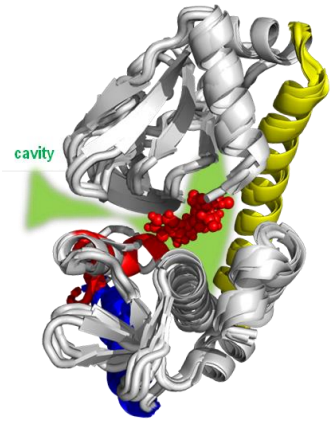


Figure S2
NN6421-derivative top-children
Examples of the similar 2D structures of some of the most representative top-NN6421 DWBEL derivatives which may deserve chemical synthesis for strictly required experimental validation steps.
Red atoms, Oxygen. **Blue atoms**, Nitrogen, **Greenish atoms**, carbon. **Encircled greenish atoms**, Fluor.



TTAACANNTGTTAA

=GraphycalAbstract

Figure S3
PyMol merged *Lm* prfA monomers: 2beo (apo-prfA), 5lrr (prfA+ECG), 8cb7 (prfA+EVFL) and 6hck (prfA+LL).

Grey cartoons, 2beo.pdb. 5lrr.pdb, 8cb7, 6hck carbon backbones
Yellow cartoons, 108-137 main α-helices.
Blue cartoons, 183-196 DNA-binding HTH α-helices.
Green background, possible alternative docking cavities around the crystallographic ECG-binding cavity (upper site).
Red spheres, merged EVFL, LL inhibitory peptides and ECG. To note the steric contacts with the ECG-induced red α-helix.
Red cartoons, 5lrr ECG-induced α-helix (~170-177 residues), that includes stabilization of DNA-binding affinity. in contrast, the corresponding sequences in the presence of inhibitory peptides and in the apo-prfA showed variable conformations.
Blue letters, operator DNA of 14-bp palindromic consensus sequence of *Lm* virulence genes.

Computational methods

PrfA 3D models

The amino acid sequences of *Lm* prfA were extracted from 5lrr.pdb (in complex with the ECG glutathion). Other prfA structure files were the apo-prfA (2beo.pdb), prfA+ECG cofactor (5lrr.pdb), prfA+EVFL inhibitory non-Cystein tetrapeptide (8cb7.pdb) and prfA+LL inhibitory non-Cystein dipeptide (6hck.pdb).

Generation of peptide and acetic-polymer libraries

One text library of all possible 6859 tri-peptides, was generated by a home-made Python script from an smiles 20 natural amino acid dictionary combining all possible orders of three amino acids, except Cystein[3]. The initial single letter code tri-peptide text file was converted to one sdf file coding for atom tri-peptide structures by the Converter ChemAxon server (https://datascience.unm.edu/tomcat/biocomp/convert). To conserve their 2D geometries during ADV docking, 3D-conformers were generated for each tri-peptide by the DW mmff94s+ force-field. All possible tetramers conserving the top FWW tri-peptide were generated adding all possible N- and/or C-terminal amino acids except C and following similar procedures than for the tri-peptides. Additionally, all possible pentamers were also generated from the top WFWW tetramer. Tri-acetic polymers were computationally generated by derivation of the single amino acid smiles dictionary except Cysteins mainly by eliminating the amino-terminal peptide bond (tri-acetic derivatives) using appropriated home-made Python scripts.

AutoDockVina docking

Before docking, any libraries in sdf files were converted to 3D-conformers applying the DW mmff94s+ force-field algorithm[30] to strictly preserve their 2D geometries during minimization, pdbqt generation and ADV docking.

Computational affinity predictions and prfA targeted cavities were performed by **A**uto**D**ock**V**ina (ADV)[46] docking using the OpenBabel mmff94s[13, 17] force-field algorithm included into the home-tuned PyRx-0.98/1.0 packages. ADV was chosen because of their limited but relatively high accuracies compared to other docking programs experiencing world-wide ongoing improvements[10, 11, 28, 29]. In particular ADV docking was employed here to:

i) Select the best parents and prfA targeted cavities by blind-docking (targeting grids of 30x30x30 Å, surrounding most of the prfA molecule),

ii) Quantify the approximated affinities of the best-conformers in ~ nM. The ADV predicts score outputs estimated in Kcal/mol[31, 32-34]. Kcal/mol estimations were converted to ~ nM affinities by applying the formula, 10^9-(exp$^{(-Kcal/mol/0.592)}$), and

iii) identify nearby amino acids to ADV docked 3D conformer-prfA complexes to confirm their PyMol visual targeted prfA cavities.

Co-evolutionary random-generation and selection of fitted-children molecules

Co-evolutionary docking was performed by **D**ata**W**arrior-**B**uild **E**volutionary **L**ibrary (DWBEL). DWBEL co-evolutions randomly generated molecular atom variations during 3 consecutive-independent runs. For each run, selection of children were using fixed criteria values and relative weights (x1-x4) such as minimal docking-score (4 value) and maximal relative weight (x4), optimal molecular weight <= 600 g/mol (x2), low-medium hydrophobicity LogP <=4 (x1), and minimal Toxicity risk <=1 (x4). From the parent molecule, DWBEL co-evolutions randomly added/inserted small atom variations to originate tens of thousands of consecutively numbered raw-children IDs. The raw children selected for best cavity-fitted-children and absence of know toxicity motifs, were finally filtered by a the NTN DW macro to exclude hundreds of children molecules containing any remaining signatures of mutagenesis, tumorigenicity, reproductive interference, irritant molecular signatures, and/or nasty motifs[21, 24, 25, 27, 35]. The resulting non-toxic fitted children were provided with NN numbers from the highest to lower DW affinities, their mmff94s+ conformers generated and all the generated data rows (one children per row) saved to an unique sdf vs3 file. The final children were ADV docked and rank-reordered by their predicted ADV nM affinities.

Funding: The work was carried out without any external financial contribution
Competing interests: The author declares no competing interests
Authors' contributions: JC designed, performed and analyzed the work and drafted the manuscript.
Acknowledgements: Thanks are due to Dr.Juan Arques at CSIC-INIA for suggesting this work. Dr. Jo for help with the acetic-polymers.

References

[1] Eiting, M., *et al.* The mutation G145S in PrfA, a key virulence regulator of Listeria monocytogenes, increases DNA-binding affinity by stabilizing the HTH motif.*Mol Microbiol.* 2005, 56: 433-46 http://dx.doi.org/10.1111/j.1365-2958.2005.04561.x.
[2] Hall, M., *et al.* Structural basis for glutathione-mediated activation of the virulence regulatory protein PrfA in Listeria.*Proc Natl Acad Sci U S A.* 2016, 113: 14733-14738 http://dx.doi.org/10.1073/pnas.1614028114.
[3] Krypotou, E., *et al.* Control of Bacterial Virulence through the Peptide Signature of the Habitat.*Cell Rep.* 2019, 26: 1815-1827 e5 http://dx.doi.org/10.1016/j.celrep.2019.01.073.
[4] Awad, I.E., *et al.* High-throughput virtual screening of drug databanks for potential inhibitors of SARS-CoV-2 spike glycoprotein.*J Biomol Struct Dyn.* 2020: 1-14 http://dx.doi.org/10.1080/07391102.2020.1835721.
[5] Dallakyan, S. and Olson, A.J. Small-molecule library screening by docking with PyRx.*Methods Mol Biol.* 2015, 1263: 243-50 http://dx.doi.org/10.1007/978-1-4939-2269-7_19.
[6] 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.
[7] 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.
[8] Lee, I. and Nam, H. Sequence-based prediction of protein binding regions and drug-target interactions.*J Cheminform.* 2022, 14: 5 http://dx.doi.org/10.1186/s13321-022-00584-w.
[9] Baek, M., *et al.* Accurate prediction of protein structures and interactions using a three-track neural network.*Science.* 2021, 373: 871-876 http://dx.doi.org/10.1126/science.abj8754.
[10] de Haas, R.J., *et al.* Rapid and automated design of two-component protein nanomaterials using ProteinMPNN.*Proc Natl Acad Sci U S A.* 2024, 121: e2314646121 http://dx.doi.org/10.1073/pnas.2314646121.
[11] Sumida, K.H., *et al.* Improving Protein Expression, Stability, and Function with ProteinMPNN.*J Am Chem Soc.* 2024, 146: 2054-2061 http://dx.doi.org/10.1021/jacs.3c10941.
[12] Kshirsagar, M., *et al.* Rapid and accurate prediction of protein homo-oligomer symmetry using Seq2Symm.*Nat Commun.* 2025, 16: 2017 http://dx.doi.org/10.1038/s41467-025-57148-3.
[13] Lisanza, S.L., *et al.* Multistate and functional protein design using RoseTTAFold sequence space diffusion.*Nat Biotechnol.* 2024: http://dx.doi.org/10.1038/s41587-024-02395-w.
[14] Cretu, M., *et al.* SYNFLOWNET: Design of diverse and novel molecules with synthesis constrains.*arXiv.* 2024, 2405.01155:
[15] Passaro, S., *et al.* Boltz-2: Towards Accurate and Efficient Binding Affinity Prediction.*preprint.* 2025: http://jeremywohlwend.com/assets/boltz2.pdf.
[16] Coll, J.M. Diffusion generation of peptides crossdocking three finger-loops of snake-venoms.*ChemRxiv.* 2025: https://chemrxiv.org/engage/chemrxiv/article-details/680f3ac950018ac7c5842ceb.
[17] Coll, J. Poxvirus targeted by RFdiffusion peptide-binders.*BioRxiv.* 2025: https://doi.org/10.1101/2025.05.14.654163.
[18] Coll, J.M. Evolutionary-docking targeting bacterial FtsZ.*ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/6405c36fcc600523a3bcb679: http://dx.doi.org/10.26434/chemrxiv-2023-ld9d3
[19] Coll, J. Could Acinetobacter baumannii Lol-abaucin docking be improved?*ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/649aa71aba3e99daef1d1756: http://dx.doi.org/10.26434/chemrxiv-2023-962ht.
[20] Coll, J.M. Anticoagulant rodenticide novel candidates predicted by evolutionary docking.*ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/6479b8cfbe16ad5c57577cce: http://dx.doi.org/10.26434/chemrxiv-2023-gh4xl-v2.
[21] Coll, J.M. Tailoring evolved-ligands to Plasmodium circumsporozoite-protein *ChemRxiv.* 2024: http://dx.doi.org/10.26434/chemrxiv-2024-g394q.
[22] 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: http://dx.doi.org/10.26434/chemrxiv-2023-v8gqj
[23] Bello-Perez, M. and Coll, J.M. Ligands docking to ORF8 by co-evolution. Could they reduce the inflammation levels induced by SARS-COV-2 infections?*ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/655f2f3629a13c4d47ced2aa: http://dx.doi.org/10.26434/chemrxiv-2023-6r8s7.
[24] Coll, J.M. Low-toxicity nanoMolar scaffolds with hundreds of variants generated by computational co-evolution into prokaryotic potassium channel cavities, .*ChemRxiv.* 2024: https://doi.org/10.26434/chemrxiv-2024-vlxbt.
[25] Coll, J.M. Star-shaped conformers generated by co-evolutionary docking predict cross-fitting glycoprotein trimer pre-fusion interfaces on VHSV fish rhabdovirus.*ChemRxiv.* 2024: https://doi.org/10.26434/chemrxiv-2024-xs262.
[26] Coll, J.M. Star-like docking to F mutations of respiratory syncytial virus *ChemRxiv.* 2024: http://dx.doi.org/10.26434/chemrxiv-2024-br178.
[27] Coll, J.M. Could HIV-Vif complex docking ligands protect human-A3G? .*ChemRxiv.* 2024: http://dx.doi.org/10.26434/chemrxiv-2024-l0z77.
[28] Coll, J.M. A3A-ligand evolution against anti-hypermutational human metastasis.*ChemRxiv.* 2025: https://chemrxiv.org/engage/chemrxiv/article-details/679c91606dde43c908146ea4.
[29] Coll, J.M. Exploring non-toxic co-evolutionary docking.*ChemRxiv.* 2023, https://chemrxiv.org/engage/chemrxiv/article-details/6512b162ade1178b2424c325: http://dx.doi.org/10.26434/chemrxiv-2023-rr5b0.
[30] 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.
[31] 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.
[32] Huey, R., *et al.* A semiempirical free energy force field with charge-based desolvation.*J Comput Chem.* 2007, 28: 1145-52 http://dx.doi.org/10.1002/jcc.20634.
[33] 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.
[34] 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.
[35] Coll, J.M. Generation of ligands to F13L mutations isolated in MonkeyPoxVirus-infected patients resistant to Tecovirimat-treatment.*ChemRxiv.* 2024, http://dx.doi.org/10.26434/chemrxiv-2024-bc9cx: