Tailoring evolved-ligands to *Plasmodium* circumsporozoite-protein

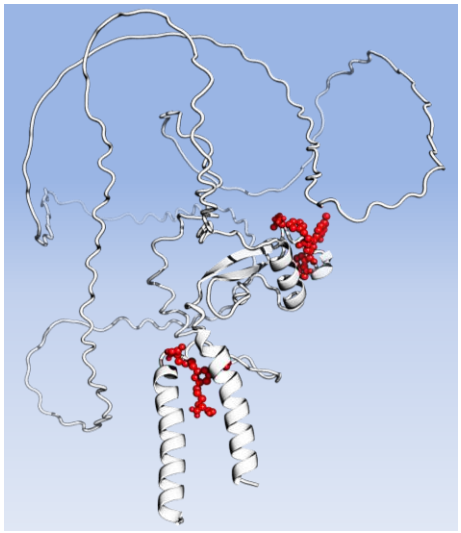
short title: co-evolutionary docking to *Plasmodium* CSP

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@inia.csic.es (JC)
* Corresponding author

Abstract

To prevent malaria deathly infections, the *Plasmodium* circumsporozoite major protein (CSP) have been targeted world-wide to develop most recent vaccines inducing anti-CSP antibodies. In contrast, drug-like anti-CSP to complement that anti-CSP tool-box, remain underdeveloped. Despite the tridimensional coat of disordered-repeats, computational predictions mimicking natural co-evolution tailored evolved ligands to adapt to most ordered CSP cavities. Tens of thousands of parent-generated raw-candidates selected hundreds of fitted-children conformers predicting low nanoMolar affinities, low toxicities, and cross-docking N-terminal signal peptide with C-terminal α-helices or docking C-terminal cavities. These repeat-independent drug-like predictions, could provide some proof-of-concept examples for basic *in vitro* experimentation.

Keywords: co-evolutionary docking; malaria: circumsporozoite, CSP, plasmodium, drugs



What considerations were made?

Mimickig accelerated natural evolution has been applied to generate small molecules computationally fitting protein cavities. Unlike screening huge molecular banks[1, 2] or predicting docking by protein sequence [3-5], co-evolutionary docking algorithms deeply penetrate into the vast drug-like chemical space[6, 7].

Non-toxic nanoMolar affinities drug-like candidates[8], have been predicted into different protein / ligand pair cavities. To briefly mention, new antibiotics fitting FtsZ of resistant *Staphilococcus*[9], alternative Abaucin-derivatives against *Acinetobacter* lipoprotein[10], anticoagulant non-human brodifacoum-derived raticide ligands[11], anti-monkeypox conformers to Tecovirimat-resistance mutants[12], anti-glycoprotein trimer inner-cavity of omicron coronavirus[13], anti-inflammatory coronavirus-coded protein[14], new docking to prokaryotic models for human potassium channels[15], or anti-fish rhabdovirus cross-docking their glycoprotein trimers[16].

Java-based ***D**ata**W**arrior **B**uild **E**volutionary **L**ibrary* (DWBEL)[2-5] fast algorithms have been employed here to target the most abundant and highly disordered **c**ircum**s**porozoite **p**rotein (CSP) of *Plasmodium falciparum*.

P. falciparum protozoan species causing malaria, are transmitted by *Anopheless* mosquitos and related species. Infections start after intradermal inoculation of 10-100 gliding elongated unicellular circumsporozoites[17]. Circumsporozoites traverse surrounding cells to reach mammalian blood vessels to circulate to mammalian organs. Some migratory circumsporozoites reach the liver, bind to the hepatocyte surfaces and intracellularly invade them. Inside the hepatocyte, the elongated circumsporozoites become rounded and divide to produce thousands of parasites (merozoites)[18]. Merozoites scape hepatocytes to penetrate blood erythrocytes[19] where they further replicate reaching clinical infection manifestations[20]. Malaria affects millions of people in tropical regions causing deathly diseases, specially in human infants.

The surface of **C**ircum**s**porozoites is densely coated by one **p**rotein (CSP) of 397 amino acids, accounting for 5–15 % of *Plasmodium sp* proteins. In the *P.falciparum* reference 3D7 isolate, the CSP contains an N-terminal domain that starts by the **s**ignal **p**eptide (SP) removed shortly after synthesis[18], followed by the disordered repeat (~ 45 % of the total sequence), and the C-terminal domain (**Table S2**). The N-terminal domain codes for a protease cleavage motif (RI). Flanked by the N- and C-terminal domains, 3D disordered repeats form dynamical ~ shield-like coatings, repeating 4 amino acids of varying total length among *Plasmodium sp* and isolates. In *P.falciparum*, the repeat domain extends from residues 129-273[16], starting with a "junctional region" ([101]NPDP + 3x [105]NANPNVDP repeats) and followed by 35x NANP repeats with one [197]NVDP insertion [21]. The C-terminal domain codes for a glycosyl-phosphatidylinositol (GPI) membrane anchor, followed by a conserved cell adhesive **t**hrombo**s**pondin **r**epeat (TSR) and ended by an hydrophobic α-helix.

Numerous studies using specific anti-CSP domain antibodies, monoclonal antibodies (mAbs), multiple mutations and/or inhibitors of protease cleavage, suggested that to trigger hepatocyte invasion, both N- and C-terminal domains bind to surface **h**eparan **s**ulfate **p**roteo**g**lycans (HSPG) [22, 23]. CSP highly conserved hydrophobic cavities with unknown functions on the conserved crystallographic C-terminal domain [24] may be implicated in HSPG-binding but that has yet to be proven. Flanked by those hydrophobic cavities, TSR residues [345]R and [347]L remain masked to Abs during circumsporozoite migration maintaining a non-adhesive conformation. N-terminal binding to HSPG induces proteolytic RI cleavage, unmasks TSR to an adhesive conformation and triggers hepatocyte invasion[24]. Homologous TSR-like motifs are known in ~ 200 proteins from several species (~ 40 human), but the TSR of *P.falciparum* has different amino acids and unique disulphide patterns[25].

Only immunodominant anti-N-terminal and/or anti-repeat Abs were protective in mice models, in contrast to anti-N-terminal or anti-SP Abs[26]. Comparison of ~ 200 mAbs induced by circumsporozoite immunization suggested that only those with anti-repeat affinities at low nanoMolar ranges would inhibit malaria infections[21, 27-30]. Among those mAbs, the mAb850 (targeting N-terminal junctional and repeats) induced spiral repeat conformations (conditional-folding?). The picoMolar affinities of mAb850[18, 19] inhibited *P.falciparum in vitro* and infection in mice, suggesting new epitopes for vaccine improvements. Anti-TSR Abs are very rare, most probably due to their masking before proteolytic RI cleavage[20]. The C-terminal shield-like protection and the absence of SP on the surface of circumsporozoites could further explain their low immunogenicity.

During recent years several anti-*Plasmodium* vaccines have been proposed to combat malaria (i.e, RTS,S and R21)[31, 32]. After numerous investigations[33] a successful CSP-based subunit vaccine (RTS,S) inducing < 40% protection, was developed a few years ago. The RTS,S vaccine included 19 NANP repeats and the C-terminal region III-TSR fused to an hepatitis antigen to increase its immunogenicity. Most recently, a modified R21 vaccine has improved protection levels to < 80% [34-36]. RTS,S and R21 are now the only malaria vaccines recommended by the World Health Organization (WHO) for prevention of *P. falciparum* malaria in children[37].

Complementing vaccine-induced[31, 32] and/or therapeutic anti-CSP Abs[38], pre-erythrocytic circumsporozoite CSP could be targeted also by small drug-like molecules by traditional screening or by computational means. Difficulties are high for computational explorations of those possibilities, because of the CSP highly disorder 3D structure (only the C-terminal domain has been crystallized) and no binding-drugs or docking cavities have been proposed. A preliminary exploration of some CSP possible docking cavities and some of their corresponding initial week ligand candidates, have been performed here. By mimicking natural evolution, computational predictions tailored the initial ligands to CSP cavities to improve their affinities. Dozens of drug-like ligand conformers targeting either the N-terminal SP and C-terminal α-helices or the C-terminal cavities, independently of the repeat domain, could be predicted with nanoMolar affinity ranges and low toxicities. Despite those predictions being highly hypothetical, drug-like high-affinity small molecules targeting the most conserved sequences of CSP could add new tools to circumsporozoite basic research.

What were the results?

What are the properties of actual CSP models?

To search for drug-like ligand conformers CSP 3D models are required. However, only the conserved C-terminal amino acid 310-375 residues (αTSR) have been solved by crystallography[24]. The αTSR included one short α-helix (residues 312-324),an adhesive TSR (331-347), and a membrane anchor GPI motif (375). Full-length CSP alphafold predictions added an ending α-helix (376-395) and repeats predicting ~20-40 Å differences by α-carbon alignements and structures with low probabilities (**l**ocal **d**istance **d**ifference **t**ests, LDDT). Since full-length CSP models are not reliable, one out of 10 alphafold model was selected with a minimal similarity between two predictions. To best understand the disordered sequences, the full-length model downsized by computationally removing any disordered sequences, generated a minimum variability of ~ 1 Å when compared with the downsized UniprotKB model (AF-P19597-F1). Therefore, the most conserved 3D CSP among different models only included the crystallographic solved residues ended by a larger α-helix (**Table S2,** and **Figure 1A, gray cartoons**).

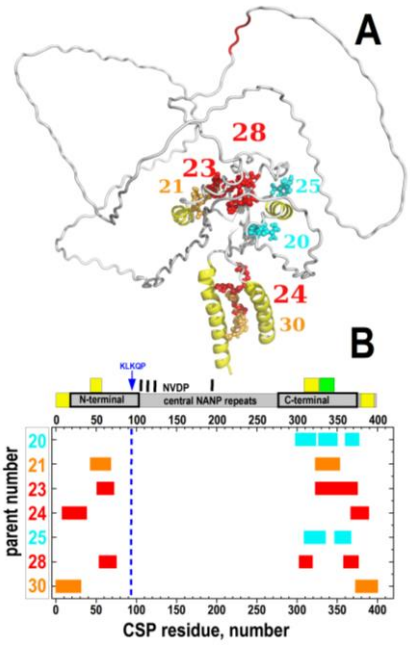


Figure 1
3D scheme of full-length CSP alphafold model and star-like ligands predicting the amino acids nearby when docked

Domains and α-helices were as displayed in **Table S2**.
A) CSP cartoons
B) amino acids nearby star-like ligands

Yellow helices (A) and rectangles (B), alphafold predicted larger N- and C-terminal α-helices (residues 1-27 and 376-395)

Blue vertical line and arrows (B), Cleavage R1 region ([93]KLKGP).

Green square (B), TSR-adhesive motif.
Red, 23L, 24L and 28L star-like ligands targeting 23C, 24C and 28C.
Orange, 21L and 30L star-like ligands targeting 21C and 30C.
Cyan, 20L and 25L star-like ligands targeting 20C and 25C.

Could star-like conformers predict CSP docking cavities?

Since no previous docking conformers have been described for CSP, an small library of 3-fold star-like ligands (L) of different sizes was designed for preliminary ADV docking, because they may best predict intramolecular cross-interactions[16]. Grids surrounding the whole molecules (ADV blind-docking) were employed because neither binding-pockets, nor docking-cavities (C), had been described at CSP. The only possible candidates but without any known function, were the hydrophobic pockets crystallographically located at CSP residues 310-375 (αTSR) [24]. Therefore, grids centered on the PyMol centerofmass and of 90x90x90 Å size were used to explore full-length CSP (including repeats).

Results targeting CSP with star-like ligands predicted seven docking-cavities defined by their best ligand conformers 20L, 21L, 23L, 24L, 25L, 28L, and 30L. Conformers 21L, 23L, 24L, 28L and 30L, cross-docked N- with C-terminal domains, while 20L and 25L only targeted C-terminal domains (**Figure 1AB**). There were no star-like conformers predicting docking-cavities within the repeats. As expected, the highest affinities to CSP of the 3-fold star-like conformers were only at low ~ μM ranges, however they helped to define starting parents and targeted cavities for co-evolution.

How higher-affinity conformers were generated?

DWBEL co-evolutions were employed to randomly generate new conformers and select those best-fitted to cavities. Each DWBEL co-evolution require two different inputs: **i)** a 2D parent such as one star-like ligand (L), to randomly derivate raw-children, and **ii)** one 3D cavity (C) on CSP, to evaluate fitting and affinities of each raw-children conformer (fitted-children). Additional preference criteria were adjusted to generate tens of thousands of raw-children to select a few thousands of non-toxic cavity-fitted-children 3D conformers (**Figure 2**). DWBEL fitted-children conformers were finally ADV re-docked to generate additional 3D conformers, explore wider cavities by blind-docking and rank their affinites for a first comparison of the results.

The results predicted that the highest ~ 3 -4 nM (n=2) affinities were those from top-children conformers derived from 24L (**Figure 2, red circles,** and **black-edged red circles).** Lower ~10-50 nM affinities were predicted for 28L and 23L (**Figure 2, red squares and triangles**). Still lower ~ 100 nM affinities were predicted for 25L and 30L (**Figure 2, cyan circles** and **orange circles**) and ~ 1000 nM for 20L and 21L (**Figure 2, cyan squares** and **orange squares**). The 24C, 28C and 23C were the dominant targeted CSP cavities, apparently corresponding to either cross-docked N- with C-terminal domain (24C), or C-terminal domains (28C, 23C). Confirmation of those targeting cavities were then performed.

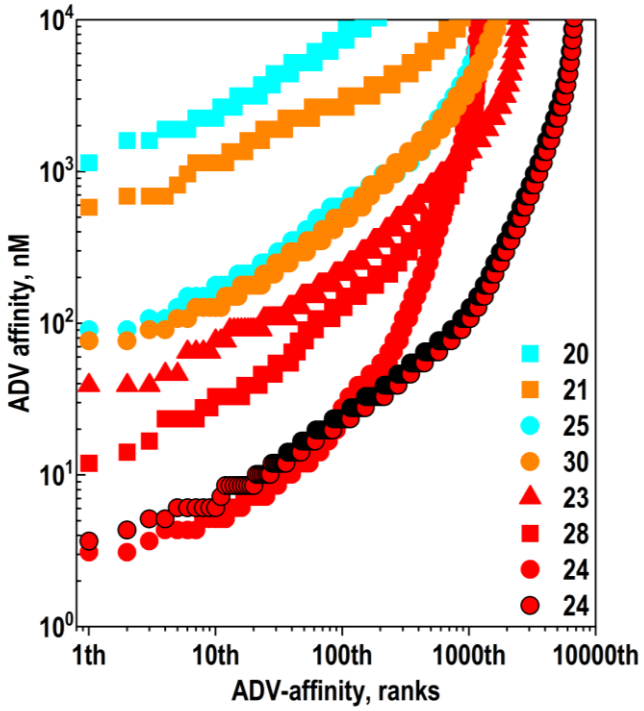


Figure 2
ADV-affinities of DWBEL-children conformers targeting CSP cavities

Pairs of ligands / cavities described in **Fig1AB** were employed for DWBEL co-evolutions. Criteria were designed to generate thousands of non-toxic children fitting CSP cavities. The ADV conformers were then ranked in ~nM affinities after blind-docking to CSP.
Cyan squares, 20L targeting CSP-20C (**Fig1AB**).**Orange squares**, 21L targeting CSP- 21C (**Fig 1AB**). **Cyan squares**, 25L targeting CSP-25C (**Fig1AB).Orange circles** , 30L targeting CSP-30C (**Fig1AB**). **Red triangles**, 23L targeting CSP-23C (**Fig1AB**). **Red squares**, 28L targeting CSP-28C (**Fig1AB**). **Red circles**, 1301 children 24L-derivatives targeting CSP-24C (**Fig1AB**). **Black-edged red circles**, 6970 children 24L-derivatives targeting CSP-24C (**Fig1AB**).

Do the ADV conformers targeted their initial DWBEL-cavities?

The top-children conformers (n=100) targeting the 24C predicted 100 % targeting to their initial DWBEL 24C. In contrast, 99 or 65 % from 28L- or 23L-derived top-children conformers, respectively, also targeted 24C instead of targeting their corresponding 28C or 23C (not shown), suggesting that 24C is the dominant cavity for full length CSP docking. Most probably the displacement of targeted cavities was most probably due to the random conformer generation, selection of only the best conformer per children and the wider ADV target space (blind-docking). In many cases, the displaced cavities were targeted with lower affinities than their corresponding top-children (not shown).

All the ADV targeting 24C predicted cross-docking of the two CSP longest α-helices, ~ N-terminal (SP) and ~ C-terminal**.** There were no other ligand conformers predicting similar cross-docking of SP and C-terminal α-helices, except 30L. However, the 30L-derived children conformers predicted ~100-fold lower affinities than 24L (**Figure 2, orange circles** and **red circles**).

To further explore the docking possibilities of 24C, DWBEL preference criteria were adjusted to generate larger numbers of 24L-derived fitted-children. Results showed that co-evolutions generating 1301 (2 runs) and 6970 (6 runs) fitted-children, predicted similar affinity rank profiles (**Figure 2, red circles** and **black-edged red circles,** respectively) and similar maximal affinites at the low nanoMolar ranges (**Supplementary Materials / 1301CSP.dwar and 6970CSP.dwar**).

Most of the top-children conformers were molecules centered around a benzene ring displaying ~ 3-fold star-like structures, with different atom small variations. In the 1301 co-evolution, most top-children conformers predicted 7 ring scaffolds while 4-6 ring scaffolds were minor (**Figure 3A**). In contrast, in the 6970 co-evolution, most top-children conformers predicted 5 ring scaffolds and 6 rings were minor (**Figure 3B**). Despite their different number of ring scaffolds, top-

children similarly cross-docked N- (SP) and C-terminal α-helices (**Figure 3CD**). These results were confirmed by identification of the predicted interactions with nearby CSP amino acids (**Table S3**).

These results may suggest that cross-docking N-(SP) and C-terminal α-helices with low nanoMolar affinities, could be enough to interfere with CSP during protein synthesis. Despite most protective human Abs targeting repeats, some examples with protective activity targeting both N-terminal and repeat sequences have been also described [30], However, since the cross-docking between N-(SP) and C-terminal α-helices may be CSP model-dependent and/or require SP to be present on CSP, further studies were performed.

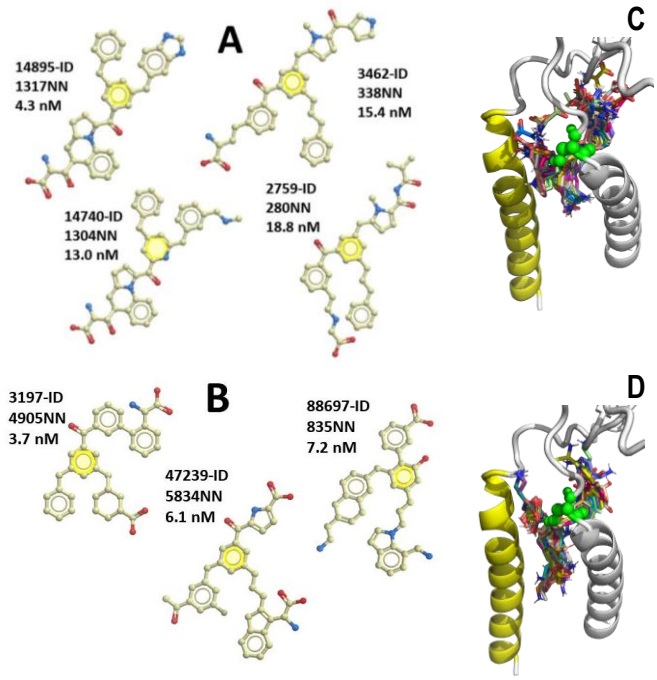


Figure 3
24L-derived 2D top-children (A,B) and their 24C-docked 3D conformers (C,D)
A) 233 ADV top-children scaffolds selected from 1301 children (n=2) and including children with 4-6 rings (**Figure 3A, red circles**)
B) 84 ADV top-children scaffolds selected from 6970 children (**Figure 3A, black edged red circles**).
A,B, legends, DWBEL generation order (ID), ADV order (NN) and ADV affinity in ~nM drawn in MolSoft.
Red spheres, Oxygens. **Blue spheres**, Nitrogens. **Light green spheres/sticks**, Carbons and bonds
Yellow rings, central hexagonal rings (cyclohexane or benzenes) of top-children conformers
C,D, 19 or 84 top-children conformers ADV blind-docked to CSP
Yellow α-helix, 1-25 CSP SP (left α-helix) **Gray α-helix**, 376-396 (right α-helix)
Green spheres, residue 375 GPI-anchor motif.
Multicolor sticks, 24L-derived top-children docked to CSP-24C. 3D docked complexes of top-children conformers were supplied (**Supporting Materials / 19topCSP.pse, 84topCSP.pse**).

Is the cross-docking of α-helices CSP model-dependent?

Although the main 3D differences among CSP models were mostly due to their disordered repeats, there were also some differences between the two N- and C-terminal α-helices. Therefore, to explore for possible model-dependence of the dockings targeting 24C, several full-length-alphafold models, were ADV blind-docked to 24L-derived CSP top-children conformers selected among those predicting < 20 nM affinities. Results showed ~ 10-50-fold ranked affinities among different CSP models (**Figure 4AB**). The CSP0 initial model predicted the highest affinities, because it was employed to DWBEL generate the 24L-derived top-children.

It could be concluded that the disordered repeats and/or the relative position variations of N- and C-terminal α-helices, interfered with 24C docking-affinities. The CSP disordered repeats may have unknown functions with stretches without any folding (intrinsically disordered) or with conditional folding [41]. Intrinsically disordered domains may be highly dynamic changing their conformations to favour antibody escape to protect the C-terminal CSP functions[42 43]. Docking recognition of intrinsically disordered repeats may be challenging, as shown for some viral nucleoproteins (i.e, 51 % intrinsically disordered nucleocapsid of SARS-CoV-2)[44-46]. On the other hand, some of the CSP disordered repeats may be conditional, getting some fold after additional protein interactions such as those spiral repeat conformations induced by mAb850 binding[18, 19]. Alphafold cannot yet predict such flexibilities[47] since their predicted models are only accurate at the crystalized αTSR domains[24]. In contrast to intrinsically disordered or conditional folding repeats (difficult to target by docking because possible multiple, transient and/or non-specific interactions), the well known CSP folded motifs may be more reliable to target for docking (i.e., hydrophobic cavities described by crystallographic analysis[24]).

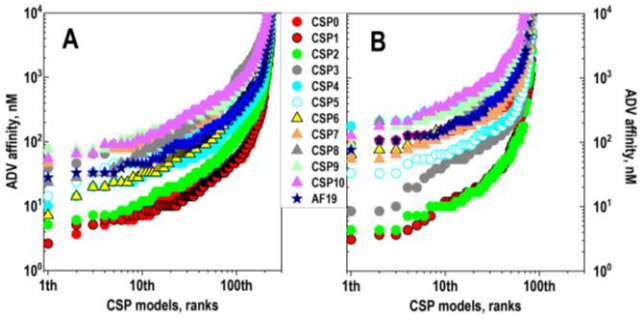


Figure 4
Top-children conformer ADV-affinities were CSP model-dependent
ADV affinity ranks of 24L-derived top-children predicting < 20 nM affinities targeting alphafold CSP models.
A) 233 top-children conformers from 1301 children (from **Figure 3B, red circles**)
B) 84 top-children conformers from 6970 children (from **Figure 3B black-edged red circles**)
Circles and triangles of different colors, alphafold-predicted models
Red circles, CSP0 initial model .**Blue stars**, UniProtKB proposed model

Did SP removal alter CSP docking?

One of the concerns about targeting 24C was due to the temporary SP. SPs are ~ 20 mer peptides located at the N-terminal nascent chains of proteins to be secreted. Since SPs are co-translationally removed after the proteins are translocated to the membranes of the endoplasmic reticulum, the SP moiety of 24C should be absent on the surface of migratory circumsporozoites[18]. Perhaps the 24C could be shortly targeted during its intracellular biosynthesis, whether inside hepatocyte and/or erythrocyte cells. Examples of co-translational translocation inhibitors targeting SPs specifically are beginning to appear as new drug targets[48], therefore, it may be possible that some top-children conformers targeting 24C could inhibit CSP during its biosynthesis.

The possible effect of SP removal on the 24C affinites were also explored, since it could be possible that removal of SP do not change the 24C conformer affinities. For that, residues 1-27 (SP) were computationally deleted from CSP (CSP-SP) and the resulting affinities compared to CSP (CSP+SP). The results predicted ~ 15-fold reduction of rank profiles for CSP-SP (n=2) (compare **red open small + large circles** and **red solid small + large circles** at **Figure 5**).

To maximize the probabilities to find any top-children with higher affinities to CSP-SP, larger numbers of fitted-children (~10000, 16 runs) were generated from the 1268NN top-child (~ 3 nM affinity). The results predicted higher affinities but still ~ 3-4-fold lower than those from CSP+SP (compare **red partially-open circles** and **red solid small + large circles** at **Figure 5**). Most important was the observation that the SP removal, displaced all the 1268NN-derived children conformers from their initial targeted 24C (**Figure 5, right-bottom cartoons**) to ~ 23C / 28C (**Figure 5, right-up cartoons**). Therefore, none of the top-children conformers initially targeting 24C (n=100) survived docking to 24C after SP removal. Any N- and C-terminal cross-docking was eliminated in CSP-SP.

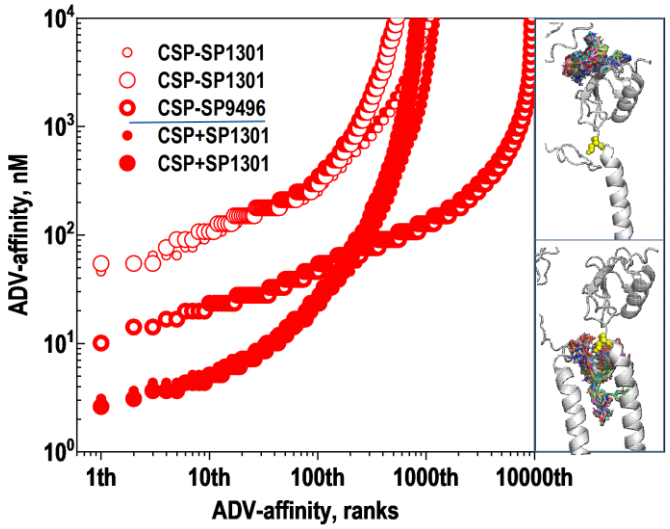


Figure 5
SP removal reduced ADV-docking affinity ranks (left) and displaced targeted cavities (right)
ADV affinity ranks were from CSP-SP (n=2) and CSP+SP (n=2) children.
Red open small + large circles, 1301 children derived from 24L targeting CSP-SP (n=2)
Red half-open circles, 9496 children derived from top-child 1268NN (~3 nM) targeting CSP-SP.
Red solid small + large circles, 1301 children derived from 24L targeting CSP+SP (n=2).
Right-up cartoons, 24L top-children ADV docked to 24C CSP-SP (C-terminal docking).
Right-down cartoons, 24L top-children ADV docked to 24C CSP+SP (cross docking N- and C-terminal)

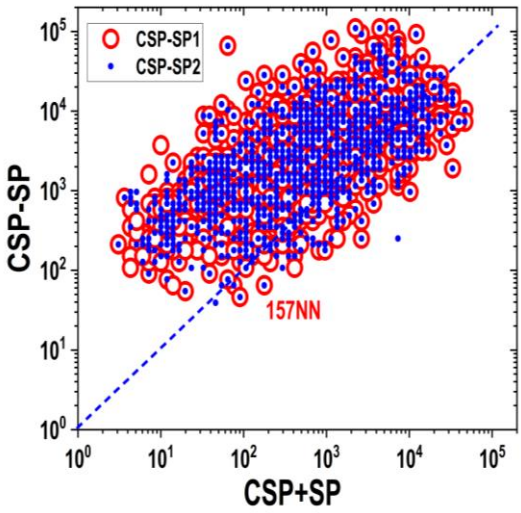


Figure 6
SP removal reduced most ADV-docking affinities and displaced their targeted cavities, except for one top-children conformer (157NN)
CSP-SP (n=2) and CSP+SP children derived by DW-BEL from 24L (data from **Figure 6**).
Red open circles, 1301 children derived from 24L targeting CSP-SP1
Blue circles, 1301 children derived from 24L targeting CSP-SP2 (duplicate)
157NN, 3-ring top-child conformer corresponding to DWBEL1354-ID

Could nanoMolar affinities be tailored to CSPmin cavities?

Top-children conformers targeting CSP independently of both SP and repeat shields would be most desirable drug-like predictions. Among the top-children maintaining high affinities in the absence and presence of SP (**Figure 6**) there was a unique 3-ring conformer. The 157NN was unique because it predicted 3 identical conformers ± SP (**Figure 6, 157NN**), targeting the same C-terminal conserved cavities (CSPmin) with maximal ~ 76 nM affinities (n=3). No other similar 3-ring children were detected on any of the previous co-evolutions, even those generating 9496 fitted-children. The 175 NN conformer targeted amino acid residues located at the conserved crystallographic C-terminal pockets[24] (**Figure 7A, red sticks** and **blue sticks** and **Figure 7B**).

The 157NN top-child could be an exceptional 3-ring conformer targeting CSP independently of SP and of the repeat shields of the C-terminal domains. Downsizing the CSP to the 309-374 residue limits targeted by 157NN (CSPmin), efforts to tailor 157NN-derivatives were attempted to explore the possible increase of their affinities. Those alternatives included, increasing the number of 24C fitted-children (n=4028) , limiting to 3-ring the DWBEL generating criteria, expanding to < 700 g/ml alternative molecular weights, and using alternative artificially-derived parents. However, despite increasing the number of 157NN-derivatives, only a few new conformers predicted similar affinities (~ 76 nM) with similar amino acid contacts (**Table S4, 157NN-derived 384NN conformer**). However, additional DWBEL iterations targeting CSPmin to sequentially tailor 157NN and 384NN-derivatives, predicted a set of new top-children could with affinities of ~ 27 nM (**Supplementary Materials / 1770CSPmin.dwar and 18CSPmin.pse**). The new-derivatives increased to 7 the rings per conformer (**Figure 7C**), and induced additional amino acid contacts including new Hydrogen bonds (**Table S4**).). Additional iterative "tailoring" could be performed to explore higher affinities to CSPmin cavities, but that was beyond this scope.

Conclusions:

Computational generations of drug-like non-toxic nanoMolar affinity 3-fold star-shaped conformers targeting hypothetic cavities have been explored by co-evolutionary docking to *P.falciparum* CSP alphafold models. Most probably due to its disordered repeat conformations, drug-like docking compounds had been rarely proposed to target CSP before. Many of the newly predicted top-conformers cross-docked N-(PS-dependent) and C-terminal domains or only the C-terminal domain. While contrast to the immunodominant protective anti-repeat CSP antibodies, these drug-like candidates may constitute proof-of-concept examples for *in vitro* basic research, to apply to more elaborated CSP models (i.e, alphafold3), or to interfere with circumsporozoite hepatocyte invasion. Among their most important limitations, these candidates suffer from: **a)** CSP model-dependence, **b)** limited numbers of parent molecules, **c)** possible induction of resistant mutations, and **d)** rigid amino acid side-chain cavities. Although some of these limitations may generate practical issues, the hundreds of conformers generated may favor the possibilities for alternative solutions. Further penetration efforts into the enormous chemical space would be required to continue possible anti-CSP drug-like explorations[36, 39].

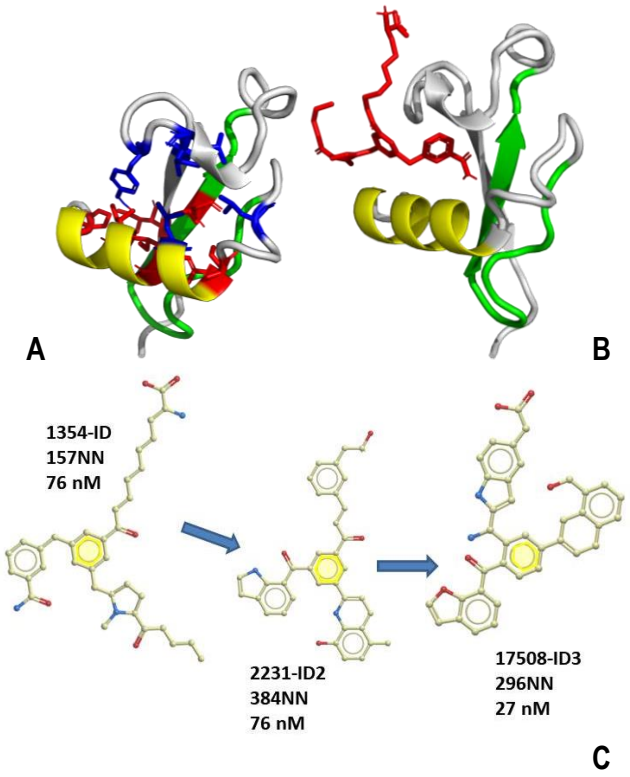


Figure 7
How 157NN-derivatives were tailored to low affinities?
CSP-SP and CSP+SP co-evolutions identified the same 24L-derived child conformer 157NN. The 175 NN targeted CSPmin 309-374 residues including the crystallographic pockets [24] (**cavity-1**: ^{311}P, ^{316}I,^{319}Y, ^{323}I, ^{342}I, ^{344}V, ^{368}I, and **cavity-2**: ^{320}L, ^{327}L, ^{358}L, ^{360}Y, ^{364}I).
AB) Grey cartoons, CSPmin 309-374. **Yellow helix,** 313-322 α-helices. **Green cartoons**, 331-347 TSR.
A) Hydrophobic cavities: 1 (**red sticks**) and 2 (**blue sticks**).
B) 157NN docked to CSPmin: Red sticks, 157NN- DWBEL1354-ID
C) Tailored top-children scaffolds, 157NN and 384NN-derived
Yellow benzenes, central ring of ~3-fold star-like conformers
Red spheres, Oxygens. **Blue spheres**, Nitrogens. **Light green spheres/sticks**, Carbons and bonds..

How the computational methods have been applied?

How 3D models of *Plasmodium falciparum* CSP were selected?

The 397 amino acid full-length **c**ircum**s**porozoite **p**rotein (CSP) sequence P19597 CSP_PLAFO (UniprotKB) *P.falciparum* (isolate NF54) was alphafold modeled[45] (https://golgi.sandbox.google.com/) .The most representative of 10 predicted alphafold models (CSP) was selected for docking (**Table S1, Figure 2A**). The disordered repeats expanding to ~ 45% of full-length CSP were different among alphafold-predicted models including the one proposed at UniprotKB (AF-P19597-F1). The disordered repeats accounted for α-carbon 3D alignment differences among models of ~20-40 Å. Deletion of disordered repeats and other sequences, reduced to ~ 1.05 Å the alignements differences with the corresponding downsized model proposed by UniprotKB. The downsized CSP expanded 95 C-terminal residues from 301-396 (**Table S2**), including the crystallized αTSR construct (310-374, 3VDJ.pdb)[24] . During this research, further downsizing to minimal CSP (CSPmin) was performed by following the stretch of only 65 residues (309-374) targeted by the 157NN conformer, independent of both SP and disordered repeats.

How the parent ligands and cavities were predicted on CSP?

To apply co-evolution algorithms, two inputs were required, **i)** one parent 2D molecule (to generate raw-children) and **ii)** a cavity protein (to select best-fitting children). Since there were no previous ligands, nor target cavities described for *P.falciparum* CSP, a previously designed home-library of 3-fold star-like small 2D molecules, with different central atoms, rings and arms between 1-6 carbons was used for ADV blind-docking[16] (**Supplementary Materials / StarLikeLigands.sdf)**. Parent ligands (L) and their cavities (C) were defined on the CSP model to start DWBEL co-evolutions.

How the co-evolution conformers were generated?

Thousands of unique children candidates were randomly generated from the previously identified parent ligands by using **D**ata**W**arrior-**B**uild **E**volutionary **L**ibrary (DWBEL) co-evolution algorithms. Once supplied with one input parent / cavity, co-evolutions were performed with the same preference criteria, relative maximal importance for docking-score of 4 (x4), molecular weight <= 600 g/mol (x2), hydrophobicity LogP <=4 (x1) and Toxicity risk <=1 (x4). The DWBEL co-evolution generation iterations randomly added/inserted small molecular variations into the 2D parents to originate tens of thousands of consecutively numbered raw-children (ID). The ID number, therefore, corresponds to the number of raw-children generated before finding the fitted-children conformer. Using the optimal mmff94s+ force-field algorithm[49] to generate the best conformers, each raw-children conformer was evaluated for fitness to the defined criteria. Therefore, DWBEL selected the non-toxic best-fitted children conformers as output. To prepare for ADV docking, the non-toxic fitted-children were further filtered with a macro designed to exclude any survivor molecules with remaining preferences for mutagenesis, tumorigenicity, reproductive interference, irritant molecular signatures, and/or nasty functions by screening the presence of hundreds of those motifs[12, 15, 16, 50] . The non-toxic fitted-and filtered children were ordered from low to high DWBEL docking-scores (high to low affinities, NN numbers) and finally saved as *.sdf 3D files. The final children *.sdf included their corresponding mmff94s+ minimization conformers[37] required to preserve 2D geometries for optimal PyMol visualization (using the split_states PyMol command)[9] and/or to increase the accuracy and reproducibility of ADV docking[32].

How the affinity and targeted cavities were confirmed?

As in our most recent work [16], the AutoDockVina (ADV) [46] and OpenBabel using mmff94s force-field algorithms (PyRx-0.98/1.0 packages) were employed because of their high accuracies and world-wide ongoing improvements[10, 11, 28, 29]. ADV was employed to: **i)** initial identification of CSP ligands (L) and cavities (C), **ii)** quantify ADV-conformer affinities in ~ nM, **iii)** identify nearby amino acids in ADV docked complexes and **iv)** confirm the CSP cavities targeted by ADV-conformers. A wide grid of 90x90x90 Å centered to the PyMol / centerofmass, surrounding most of the CSP model molecules (~ blind-docking) was employed. The output ADV-conformer docking-scores in - Kcal/mol[51, 52-54] were converted to ~ nM affinities by applying the formula, 10^9-(exp(Kcal/mol/0.592)). To identify the CSP amino acids nearby 4 Å distance of each ADV-conformer, a Python script was designed to be run in PyMol-opened *.pdb or *.pdbqt files. After preliminary ADV tests, to estimate the percentage of cavities targeted, the following defined CSP cavities were selected: **Bc** (bottom cavity between the 2 largest α-helices): 347K and/or 364I and/or 366K, **Sc** (side of the CSP): 66K and/or 63Y and/or 368I and **Tc** (top of the TSR): 296T and/or 366S and/or 411Y. The approximated percentages of those cavities targeted by thousands of ADV-conformers were calculated by the use of the PyMol/Python script (nearby11.py) [16]. The script calculated cavity percentages by the formula, 100 * number of ADV-conformers in each cavity / total numbers of docked ADV-conformers. The resulting approximations accounted for ~75-120 % of the total numbers of ADV-conformers because of some overlapping between the 3 amino acid cavity definitions.

What supplementary information is provided?

Table S1

What computational software, improvements and hardware have been used ?

name	version	Main use and references	url
DataWarrior	Updated 5.5.0 Windows/Linus	Evolutionary docking[34] Commercial ChemSpace	https://openmolecules.org/ datawarrior/download.html)
Toxicity & nasty macro	2023	Eliminate residual toxic / nastic fitted-children after co-evolution[10]	
Toxicity Risks	2023 Updated DataWarrior	Minimize toxic / nasty raw- children during co-evolution increasing specificity[8]	https://openmolecules.org/ datawarrior/download.html)
Babel & AutoDockVina	Home-adapted PyRx 098/1.0	Mmff94s force-field minimization & 2D conservation	https://pyrx.sourceforge.io/
ADV consensus	2023 2024	First attempted for Anti-bacterial [9, 13] grid conformer comparisons	http://dx.doi.org/10.26434/chemrxiv-2023-ld9d3 This work
2D geometry conservation	2023	DW saving-SD files corrected by mmff94s+ force field minimization [57]	
MolSoft	3.9 Win64bit	Easiest manipulations of sdf files 2D drawing	https://www.molsoft.com/download.html
PyMol	2.5.7.	Visualization of molecules PyMol-Python scripts to detect nearby atoms	https://www.pymol.org/ this work
Discovery Studio	21.1.1.0.20298	Visualization of 2D molecules Structure/geometry fixing	https://discover.3ds.com/discovery-studio-visualizer-download
OriginPro	2022 2024	Calculations and Figures Macros to handle large numbers of data	https://www.originlab.com/ this work
Home-made pseudoligands	2023 2024	Pseudoligand parents for DWBEL co-evolutions[14, 16]	
3-fold star-like Ligands	2024	Initial ligands to identify cavities by ADV blind-dockings[16]	
LigPlot	2.2.8.	Prediction of amino acid hydrogen bonds of docked conformers[9]	https://www.ebi.ac.uk/thornton--rv/software/LigPlus/applicence.html
AMD Ryzen i9 computer	4 DDR4 x 32 Gb memory	47 CPU Computational hardware	https://www.pcspecialist.es/

Table S2

What amino acid numbering was followed for alphafold-modeled *P.falciparum* CSP?

Name	explanation	conformation	CSP	reference
SP	Signal peptide	Alphafold-predicted α-helix	1-26	[18]
N-terminal:	Before repeats	disordered	27-104	[55]
helix	helix	Alphafold-predicted α-helix	45-55	[55]
HSPG-binding	Hepatocyte surface	HSPG-binding	85-92	alphafold
R1	protease target	^{93}KLKQP	93-97	[23]
Central repeats:	NPDP+NANP+NVDP	Conserved seq. disordered 3D	105-273	[22, 23]
N-junction	Junctional epitope	^{101}NPDP	101-104	[21]
NVDP	NVDP	3x^{105}NANPNVDP	105-128	alphafold
NANP	NANP	35 contiguous NANP	129-172	alphafold
NVDP	NVDP	Intermediate NVDP	197-200	alphafold
C-terminal:	After repeats	Partial crystal and alphafold	274-375	[55]
linker	Linker from repeats	disordered	283-309	[24]
RIIIhelix	Short helix	α-helix crystal (3VDJ.pdb)	312-324	[24]
RII (TSR)	Cell adhesive	~ TSR domain (3VDJ.pdb)	331-347	[24]
CS-flap	Protective segment	~ CS-flap (3VDJ.pdb)	348-363	[24]
GPI	Membrane anchor	Glyco-phosphatidylinositol	375	[56]
TM	Transmembrane ?	Alphafold-predicted α-helix	376-395	alphafold

The *Plasmodium falciparum* reference 3D7-strain circumsporozoite protein (PF3D7_0304600), XM_001351086.1 mRNA was submitted to alphafold. One of the most representative predicted models was chosen for this study.
Amino acid 310-375 residues (RIII-RII) aligned to 10 *Plasmodium* species showed high conservation of their ^{338}C-^{369}C and ^{342}C-^{374}C disulphide bonds and 37.3% of their amino acid sequences. Most of these domains were mapped to **Supplementary Materials/GraphycalAbstract**

Table S3

What CSP amino acids were nearby top-children representative conformer scaffolds?

			NN						
Dom	n	**Aa**	**1317**	**1304**	**338**	**280**	**4905**	**5834**	**835**
SP	15	PHE						15F	
	16	VAL	16V	16V	16V		16V	16V	16V
	19	LEU	19L	19L	19L	19L	19L	19L	19L
	20	PHE	20F	20F	20F	20F	20F	20F	20F
	22	GLU	**22E**	22E	22E	22E	22E	22E	22E
	23	TYR	23Y	23Y	23Y	23Y	23Y	23Y	23Y
	25	**CYS**			25C	25C			**25C**
	26	TYR	26Y	**26Y**	26Y	26Y	26Y	26Y	26Y
	29	SER	29S	29S	29S				
	30	SER	30S		30S				
	32	**THR**	32T	32T					
TSR	**337**	**THR**	337T	**337T**	337T	337T			
	338	**CYS**		338C					
	372	**GLU**			**372E**	372E			**372E**
GSI	375	SER	375S		375S	375S	**375S**	**375S**	375S
helix	377	VAL	377V		377V	377V	377V	377V	377V
	378	PHE	378F	378F	378F	378F	378F	378F	378F
	381	VAL	381V	381V	381V	381V	381V	381V	381V
	382	**ASN**	382N	382N	382N	382N	382N	382N	382N
	384	**SER**						384S	
	385	ILE	385I	385I	385I	385I	385I	385I	385I
	388	ILE		388I		388I	388I	388I	388I
	389	**MET**					389M	389M	

The CSP amino acids to 4 Å top-children ADV-conformers, identified by a Python/PyMol script
Column numbers, NN from 24L derived children
Red numbers-letters, Hydrogen bonds identified by LigPlus.
Yellow background, amino acids targeted by top-children and by initial star-ligands

Table S4

What CSPmin amino acids were nearby top-children conformer scaffolds?

			NN					
			157	**384**	**296**	**410**	**558**	**1423**
Dom	n	**Aa**	**24L**	**157L**	**384L**	**384L**	**384L**	**384L**
helix	**313**	**ASP**	313D	**313D**	313D	313D	313D	313D
	314	LYS						
	317	LYS	317K	317K	317K	317K	317K	317K
	320	LEU	**320L**	320L	320L	320L	320L	320L
	321	ASN	321N	**321N**	321N	321N	**321N**	321N
	323	**ILE**	323I					
	324	GLN	324Q	324Q	324Q	324Q	324Q	324Q
	327	LEU	327L	327L	327L	327L		327L
CS-	353	LYS					**353K**	
flap	355	LYS	355K					
	356	ASP			356D	356D	356D	356D
	357	GLU	357E	357E	357E	357E	357E	357E
	358	LEU	358L	358L	358L	358L	**358L**	358L
	359	TYR	359D	359D	**359D**	359D	359D	**359D**
	360	**TYR**	360Y	360Y	360Y	360Y	**360Y**	360Y
	361	ALA		361A	**361A**	**361A**	361A	**361A**
	362	ASN	362N	362N		362N		
	363	ASP			363D	363D	363D	
	364	**ILE**	364I	364I	364I	364I	364I	364I
	366	LYS		**366K**				

The CSPmin amino acids to 4 Å top-children ADV-conformers, identified by a Python /PyMol script.
Column numbers, NN from L derived children: 24L, 24L-derived, 157L, 157-derived, 296, 384L-derived.
Red numbers-letters, Hydrogen bonds identified by LigPlus.
Yellow background, amino acids targeted by top-children and by initial star-ligands

What Supplementary Materials can be downloaded?

- GraphycalAbstract.pse The mRNA of *Plasmodium falciparum* reference 3D7-strain circumsporozoite protein (PF3D7_0304600), XM_001351086.1 was submitted to alphafold. One of the most representative predicted models was selected for this study (**Gray cartoons**). Amino acid 310-375 residues aligned to 10 CSP models *Plasmodium* species showed high conservation of their ^{338}C-^{369}C and ^{342}C-^{374}C disulphide bonds (**Table S2**). **Red spheres,** Representative examples of top-children targeting two repeat-independent CSP cavities: cross-docking N- (SP) with C-terminal α-helices and C-terminal domain.

- StarLikeLigands.sdf. Contains 3-fold star-like ligand molecules manually designed in MolSoft by 2D drawing different central atoms, rings and sizes including 3-6 carbon arms ended by amino and carboxy structures (alanines). To conserve their 2D geometries during ADV docking, optimal conformers were generated by the DW / mmff94s+ force-field algorithm[49].

- 1301CSP.dwar. This *.dwar DW table contains 1301 children generated from the 24L parent targeting CSP by 2 runs of DWBEL (DW-ID) and their corresponding ADV affinities (ADV-NN). The table is provided with threshold slider-filters to select for (309-374 residues)threshold combinations (https://openmolecules.org/datawarrior/download.html)

- **6970CSP.dwar.** This *.dwar DW table contains 6970 children generated from the 24L parent targeting CSP by 6 runs of DWBEL (DW-ID) and their corresponding ADV affinities (ADV-NN). The table is provided with threshold slider-filters to select for threshold combinations (https://openmolecules.org/datawarrior/download.html)

- **1770CSPmin.dwar.** This *.dwar DW table contains 1770 children targeting CSPmin (309-374 residues) by DWBEL (DW-ID) derived from the 384NN parent (derived from the 157NN parent) and their corresponding ADV affinities (ADV-NN). The table is provided with threshold slider-filters to select for threshold combinations (https://openmolecules.org/datawarrior/download.html)

- **19top.pse.** Contains top-children ADV conformers docked to full-length CSP from the 1301 DWBEL 1301 (2 runs). To view the docked individual children click on the NN number to the right of the PyMol scene after opening the *.pse file in one of the latest PyMol 2023-24 versions.

- **84top.pse.** Contains top-children ADV conformers docked to full-length CSP from the 6970 DWBEL (6 runs). To view the docked individual children click on the NN number to the right of the PyMol scene after opening the *.pse file in one of the latest PyMol 2023-24 versions.

- **18top.pse.** Contains top-children ADV conformers docked to CSPmin (309-374 residues) from the 1770 DWBEL (3 runs). To view the docked individual children click on the NN number to the right of the PyMol scene after opening the *.pse file in one of the latest PyMol 2023-24 versions.

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 computational work and drafted the manuscript.

Acknowledgements
Thanks are specially due to Dr. J.M. Andreu of the CSIC at Madrid (Spain) for his initial news about the alphafold3 version and N.Behrnd of Idorsia Pharmaceuticals Ltd. at Allschwil (Switzerland) for help with DW macros and 2D DWBEL-ADV conformer issues.

What references were consulted?

[1]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.

[2]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.

[3]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.

[4]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.

[5]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.

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

[7]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.

[8]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.

[9]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

[10]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.

[11]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.

[12]Coll, J.M. New ligands predicted to Monkeypox viral membrane proteins generated by co-evolutionary docking *ChemRxiv.* 2024: http://dx.doi.org/10.26434/chemrxiv-2024-bc9cx.

[13]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

[14]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.

[15]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.

[16]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.

[17]Frischknecht, F. and Matuschewski, K. Plasmodium Sporozoite Biology.*Cold Spring Harb Perspect Med.* 2017, 7: http://dx.doi.org/10.1101/cshperspect.a025478.

[18]Singer, M. and Frischknecht, F. Fluorescent tagging of Plasmodium circumsporozoite protein allows imaging of sporozoite formation but blocks egress from oocysts.*Cell Microbiol.* 2021, 23: e13321 http://dx.doi.org/10.1111/cmi.13321.

[19]Zhao, J., *et al.* A Comprehensive Analysis of Plasmodium Circumsporozoite Protein Binding to Hepatocytes.*PLoS One.* 2016, 11: e0161607 http://dx.doi.org/10.1371/journal.pone.0161607.

[20]Cowman, A.F. and Crabb, B.S. Invasion of red blood cells by malaria parasites.*Cell.* 2006, 124: 755-66 http://dx.doi.org/10.1016/j.cell.2006.02.006.

[21]Oyen, D., *et al.* Structure and mechanism of monoclonal antibody binding to the junctional epitope of Plasmodium falciparum circumsporozoite protein.*PLoS Pathog.* 2020, 16: e1008373 http://dx.doi.org/10.1371/journal.ppat.1008373.

[22]Coppi, A., *et al.* The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host.*J Exp Med.* 2011, 208: 341-56 http://dx.doi.org/10.1084/jem.20101488.

[23]Coppi, A., *et al.* The Plasmodium circumsporozoite protein is proteolytically processed during cell invasion.*J Exp Med.* 2005, 201: 27-33 http://dx.doi.org/10.1084/jem.20040989.

[24]Doud, M.B., *et al.* Unexpected fold in the circumsporozoite protein target of malaria vaccines.*Proc Natl Acad Sci U S A.* 2012, 109: 7817-22 http://dx.doi.org/10.1073/pnas.1205737109.

[25]Tan, K., *et al.* Crystal structure of the TSP-1 type 1 repeats: a novel layered fold and its biological implication.*J Cell Biol.* 2002, 159: 373-82 http://dx.doi.org/10.1083/jcb.200206062.

[26]Scally, S.W., *et al.* Rare PfCSP C-terminal antibodies induced by live sporozoite vaccination are ineffective against malaria infection.*J Exp Med.* 2018, 215: 63-75 http://dx.doi.org/10.1084/jem.20170869.

[27]Oyen, D., *et al.* Structural basis for antibody recognition of the NANP repeats in Plasmodium falciparum circumsporozoite protein.*Proc Natl Acad Sci U S A.* 2017, 114: E10438-E10445 http://dx.doi.org/10.1073/pnas.1715812114.

[28]Oyen, D., *et al.* Cryo-EM structure of P. falciparum circumsporozoite protein with a vaccine-elicited antibody is stabilized by somatically mutated inter-Fab contacts.*Sci Adv.* 2018, 4: eaau8529 http://dx.doi.org/10.1126/sciadv.aau8529.

[29]Pholcharee, T., *et al.* Diverse Antibody Responses to Conserved Structural Motifs in Plasmodium falciparum Circumsporozoite Protein.*J Mol Biol.* 2020, 432: 1048-1063 http://dx.doi.org/10.1016/j.jmb.2019.12.029.

[30]Tan, J., *et al.* A public antibody lineage that potently inhibits malaria infection through dual binding to the circumsporozoite protein.*Nat Med.* 2018, 24: 401-407 http://dx.doi.org/10.1038/nm.4513.

[31]Hill, S.R., *et al.* Antimalarial Medications. 2024: https://www.ncbi.nlm.nih.gov/books/NBK470158/ [bookaccession].

[32]Ngou, O., *et al.* R21/Matrix-M malaria vaccine: a vital tool in the arsenal against malaria, not a silver bullet.*Lancet Infect Dis.* 2024: http://dx.doi.org/10.1016/S1473-3099(24)00010-0.

[33]Kumar, K.A., *et al.* The circumsporozoite protein is an immunodominant protective antigen in irradiated sporozoites.*Nature.* 2006, 444: 937-40 http://dx.doi.org/10.1038/nature05361.

[34]Asante, K.P., *et al.* Feasibility, safety, and impact of the RTS,S/AS01(E) malaria vaccine when implemented through national immunisation programmes: evaluation of cluster-randomised introduction of the vaccine in Ghana, Kenya, and Malawi.*Lancet.* 2024: http://dx.doi.org/10.1016/S0140-6736(24)00004-7.

[35]Ahmad, E. and Ahmed, S. The Road to a Malaria-Free World: The Promise of R21/Matrix-M Vaccine.*Asia Pac J Public Health.* 2024, 36: 281-282 http://dx.doi.org/10.1177/10105395241233633.

[36]Schmit, N., *et al.* The public health impact and cost-effectiveness of the R21/Matrix-M malaria vaccine: a mathematical modelling study.*Lancet Infect Dis.* 2024: http://dx.doi.org/10.1016/S1473-3099(23)00816-2.

[37] WHO prequalifies a second malaria vaccine, a significant milestone in prevention of the disease.*Saudi Med J.* 2024, 45: 219 45/2/219 [pii].

[38]Kayentao, K., *et al.* Subcutaneous Administration of a Monoclonal Antibody to Prevent Malaria.*N Engl J Med.* 2024, 390: 1549-1559 http://dx.doi.org/10.1056/NEJMoa2312775.

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

[40]Tang, S., *et al.* Accelerating AutoDock Vina with GPUs.*Molecules.* 2022, 27: http://dx.doi.org/10.3390/molecules27093041.

[41]Yegambaram, K., *et al.* Protein domain definition should allow for conditional disorder.*Protein Sci.* 2013, 22: 1502-18 http://dx.doi.org/10.1002/pro.2336.

[42]Wang, L.T., *et al.* A Potent Anti-Malarial Human Monoclonal Antibody Targets Circumsporozoite Protein Minor Repeats and Neutralizes Sporozoites in the Liver.*Immunity.* 2020, 53: 733-744 e8 http://dx.doi.org/10.1016/j.immuni.2020.08.014.

[43]Heller, G.T., *et al.* Picosecond Dynamics of a Small Molecule in Its Bound State with an Intrinsically Disordered Protein.*J Am Chem Soc.* 2024, 146: 2319-2324 http://dx.doi.org/10.1021/jacs.3c11614.

[44]Pontoriero, L., *et al.* NMR Reveals Specific Tracts within the Intrinsically Disordered Regions of the SARS-CoV-2 Nucleocapsid Protein Involved in RNA Encountering.*Biomolecules.* 2022, 12: http://dx.doi.org/10.3390/biom12070929.

[45]Schiavina, M., *et al.* The Role of Disordered Regions in Orchestrating the Properties of Multidomain Proteins: The SARS-CoV-2 Nucleocapsid Protein and Its Interaction with Enoxaparin.*Biomolecules.* 2022, 12: http://dx.doi.org/10.3390/biom12091302.

[46]Quaglia, F., *et al.* SARS-CoV-2 variants preferentially emerge at intrinsically disordered protein sites helping immune evasion.*FEBS J.* 2022, 289: 4240-4250 http://dx.doi.org/10.1111/febs.16379.

[47]Sacquin-Mora, S. and Prevost, C. When Order Meets Disorder: Modeling and Function of the Protein Interface in Fuzzy Complexes.*Biomolecules.* 2021, 11: http://dx.doi.org/10.3390/biom11101529.

[48]Lumangtad, L.A. and Bell, T.W. The signal peptide as a new target for drug design.*Bioorg Med Chem Lett.* 2020, 30: 127115 http://dx.doi.org/10.1016/j.bmcl.2020.127115.

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

[50]Coll, J.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:

[51]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.

[52]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.

[53]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.

[54]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.

[55]Geens, R., *et al.* Biophysical characterization of the Plasmodium falciparum circumsporozoite protein's N-terminal domain.*Protein Sci.* 2024, 33: e4852 http://dx.doi.org/10.1002/pro.4852.

[56]Sauer, L.M., *et al.* FT-GPI, a highly sensitive and accurate predictor of GPI-anchored proteins, reveals the composition and evolution of the GPI proteome in Plasmodium species.*Malar J.* 2023, 22: 27 http://dx.doi.org/10.1186/s12936-022-04430-0.