**Original Research Article**

**Structure Based-Design of Novel Analogues Carbonitriles Pyrimidine Targeting the Cysteine Protease *falcipain 2* of *Plasmodium falciparum* (*pf*FP2) in Trophozoïte Stage with ADME Properties Favorable**

.

|  |
| --- |
| **ABSTRACT**  **Aim and Objective:** Structure-based drug design (SBDD) of new antimalarials at a moment resistance of the most causative agent, *Plasmodium falciparum*against the most valuable artemisinin combined therapy (ACT) is more than urgent. Carbonitriles pyrimidine derivatives (CNP) has emerged as potential inhibitors of the cysteine protease falcipain 2 of *Plasmodium falciparum* (*pf*FP2), therefore we report here a pharmacophore based virtual screening of the CNP chemical subspace yielding new CNP analogs with high predicted inhibitory potency against *pf*FP2.  **Methods:** A quantitative structure activity relationships (QSAR) complexation model has been developed from a series of fifteen carbonitriles pyrimidine derivatives to establish a linear correlation between the calculated Gibbs free energies (GFE: ΔΔGcom) of *pf*FP2-CNP complex formation and the experimental half-maximal enzymatic inhibition concentration ().The predictive power of the QSAR model was then validated with the generation of a 3D-QSAR-PH4 pharmacophore (PH4) model as CNP chemical subspace (exemplified as a virtual combinatorial library of more than 83,300 CNP analogs) explorer for novel predicted more potent CNP analogs. Finally the best PH4 hits were evaluated with the initial QSAR model for predicted potency () and pharmacokinetic profile.  **Results:** The QSAR model linear correlation equation:p= -0.1025 x ∆∆Gcom + 7.2867, R2 = 0.94, the subsequentPH4 model linear correlation between experiment and PH4-estimated IC50: p= 0.9366 x p+ 0.2849, R2 = 0.91 documents the high predictive power of this approach. Finally the screening of the virtual library of CNP analogs yielded 52 orally bioavailable candidates the best reaching a predicted potency () of 14 pM and displaying favorable pharmacokinetic profile.  **Conclusion:** The combined use of one descriptor complexation QSAR model and 3D-QSAR Pharmacophore model performs well in the identification of novel CNP analogs against *pf*FP2 and the handful of top predicted analogs are worth undergoing synthesis and biological evaluation.  **Keywords:** Malaria; CNP; FP2, QSAR; pharmacophore; virtual screening,ADMET |

INTRODUCTION

Malaria is an endemic disease that occurs mainly in tropical areas and is caused by unicellular eukaryotic parasites of the genus plasmodium. Among the five known malaria parasites infecting humans (plasmodium falciparum, plasmodium vivax, plasmodium malariae, plasmodium ovale and plasmodium knowlesi), plasmodium falciparum is the most virulent**[[1]](#endnote-2)**.This disease continues to be one of the main causes of death in the world and according to the WHO Malaria report 2022, there were an estimated 247 million clinical cases and 619.000 deaths due to malaria. Africa is paying a heavy price and bears a disproportionate share of the global malaria burden with 95% of malaria cases and 96% of deaths due to this disease. Children under the age of 5 are the most vulnerable group to malaria; in 2021, they accounted for almost 80% of all malaria deaths in the WHO Africa region. Children under the age of 5 are the most vulnerable group to malaria; in 2021, they accounted for almost 80% of all malaria deaths in the WHO Africa region**[[2]](#endnote-3)**. An obstacle in the treatment of malaria is the spread of resistance to most drugs, including Chloroquine, which was one of the pillars of antimalarial treatment**[[3]](#endnote-4)**.Current efforts to control malaria and prevent its spread to new regions have been hampered by the emergence of new resistance to existing available antimalarial drugs, in particular Artemisinin, our last line of defense**[[4]](#endnote-5),[[5]](#endnote-6)**including Artemisinin-based Combination Therapies (ACTs)**[[6]](#endnote-7),[[7]](#endnote-8),[[8]](#endnote-9)**. Given that resistance covers a wide range of antimalarial drugs and that it spreads in populations “at risk of parasitic infection" around the world and that no protective vaccine is available**[[9]](#endnote-10)**, there is therefore an urgent and imperative need to identify and develop new classes of antimalarial drugs with a view to a new antimalarial therapy**[[10]](#endnote-11)**.The development of new inexpensive antimalarial drugs bioavailable by oral route, overcoming drug resistance is an urgent necessity. This is subject to the constraint of finding a new attractive potential target while proceeding to a rational drug design approach and filtering a large diversified library of components to finally obtain an almost perfect pharmacokinetic profile and a multi-target compound**[[11]](#endnote-12)**. The parasite survival depends on the digestion of 3/4 of the hemoglobin contained in infected erythrocytes for supplying the amino acids needed for the parasite growth**[[12]](#endnote-13),[[13]](#endnote-14),[[14]](#endnote-15)**. Four parasitic proteases families well known namely aspartic proteases, cysteine proteases, aminopeptidases and metalloproteases, involved in this metabolic process, are expressed during the erythrocytic stage of the parasite life cycle and their inhibition has proven to be central to avoid parasite proliferation**,[[15]](#endnote-16)**. Hemoglobin degradation is blocked by cysteine protease inhibitors (I), causing characteristic morphological abnormality in which the food vacuole is full of non-degraded hemoglobin blocking by this way parasite development**[[16]](#endnote-17),[[17]](#endnote-18)**.Falcipains have drawn great interest due to their central role in the life cycle of Pf through hemoglobin degradation**[[18]](#endnote-19)**. Falcipain-2 (FP2) and Falcipain-3 (FP3) are key papain-family (C1) Clan CA trophozoite cysteine proteases localized in the Digestive Food Vacuole (DFV) that cleave host native or denatured human hemoglobin. FP2, the most expressed and best studied enzyme among falcipains, is a promising target for novel antimalarial drugs development**[[19]](#endnote-20)**. Falcipain 2 and related plasmodial cysteine proteases are thus logical targets for antiparasitic chemotherapy, and therefore, we were interested in the development of inhibitors of these enzymes as antiparasitic**[[20]](#endnote-21)**. The active site of FP2 is large and contains four pockets, S1, S’1, S2 and S3, each pocket accommodating one substituent, P1, P’1, P2 and P3 of substrate. Recent studies have shown the selective specificity of the pockets of falcipain2 (FP2 II). According to these studies, the S1 pocket has a strong affinity for compounds bearing a nitrile group at P1**[[21]](#endnote-22)** and the S2 pocket has a marked preference for compounds having a hydrophobic group at P2**[[22]](#endnote-23)**.The main classes of FP2 inhibitors bearing the most popular pharmacophore of cysteine protease inhibitors, such as vinyl sulfones, halomethyl ketones, and aldehydes. Furthermore, many other chemotypes have been identified as FP2 inhibitors, such as isoquinolines, thiosemicarbazones, and chalcones**[[23]](#endnote-24)**. Also Different non peptide heteroarylnitrile derivatives were studied as potential falcipain inhibitors and therefore potential antiparasitic lead compounds, with the 5-substituted-2-cyanopyrimidine chemical class well know under the name of carbonitriles pyrimidine (CNP) emerging as the most potent and promising lead series. However, the P1 side chain of the carbonitriles pyrimidine family comprising the 2-cyanopyrimidine nucleus offers a great possibility of substitution in the S1 pocket. Pyrimidines are organic molecules similar to benzene but with two nitrogen atoms (N) in positions 1 and 3. In this family of active molecules, thanks to a process of sequential optimization of derivation taking into account the different positions present in the initial scaffold of the pyrimidine nucleus, inhibitors of FP2 having an activity against parasites cultured in nanomolars and subnanomolars have been identified**[[24]](#endnote-25)**.The introduction of protonable amines (cyano/carbonitriles) into the molecules led to a clear improvement of up to 1000 times the activity against the cultivated parasites without noticeable alteration of the other trends in terms of structure-activity relationship. The possibilities in terms of structural diversity have a strong advantage because they are very useful in the fight against drug resistance of certain pathogens and constitute a reliable means for designing powerful, bioavailable and favorable oral pharmacokinetic antimalarial. Having no crystallographic structure of FP2-CNP complexes available in the literature, we therefore used the results of a previous study on azadipeptides nitrile (ADPN) peptide inhibitors of FP2**[[25]](#endnote-26)**, by *in situ* modification to explore the active site of FP2 with CNPs.The structural information obtained with the ADPNs, guided the simulations carried out with the CNPs and opened a gate to the structure-based design of novel potent antimalarial agents. In this work we start from training set of CNP to construct a one descriptor (Gibbs free energy, GFE upon FP2-Inhibitor complex formation) QSAR model of FP2 inhibition, correlating GFE with experimental activity . Subsequently, a 3D-QSAR pharmacophore protocol was used to prepare a four-feature pharmacophore (PH4) model from FP2 active site bound CNP conformations.Moreover, the computed enzyme – ligand interaction energy map Eint correlates well with ; thus allowing us to reach its breakdown to each active site residue contribution. From this last structural information we were able to select P1 and P2 suitable fragments as building blocks for a Virtual Library (VL) of *pf*FP2 inhibitors. In order to prevent any toxicity issue and access good pharmacokinetic profile analogues, the VL was focused, prior to any screening, to those compounds the ADME of which has 0 property descriptors that fall outside the range of values determined for 95% of known drugs out of 24 selected descriptors computed by the QikProp**[[26]](#endnote-27)**.The predictability of the obtained QSAR models of inhibitor-enzyme binding cross-checked with a PH4 3D-QSAR pharmacophore model was used to screen the VL. The best Hit Fits from the PH4-based virtual screening of the VL have been in silico MM-PB evaluated to yield a predicted inhibitory activity reaching the picomolar range for the most potent analogues.

material and methods

**Training and validation set**

The chemical structures of the class of 5-substituted-2-cyanopyrimidines well known as carbonitriles pyrimidine (CNP) and their experimental biological activities in the training and validation sets used in this work have been taken from the literature. Their activities cover a sufficiently wide range of activity to allow the construction of a reliable QSAR model. The TS containing 12 CNP ligands and the VS3 CNP taken from the reference.

**Model building and calculation of binding affinity**

No crystallographic structures of the FP2-CNP complexes exist. The FP2:CNP complexes were built by *in situ*modify-cation of the high-resolution crystallographic structure of the reference complex FP2:E64 (PDB code 3BPF at a resolution of 2.90 Å) using the Insight-II 2005 Molecular Modeling program**[[27]](#endnote-28)**.The structures of the proteins and their comp-lexes were considered at a pH of 7 with neutral N- and C-terminal residues and all charged "protonable" and "ioni-zable" residues. No water molecules from the crystallogra-phic structure were kept in the model. An exhaustive confor-mational search on all the single bonds of the newly created residues coupled with an energy minimization of the inhibi-tor and the active site of the protein was necessary to identify the lowest energy bound conformation of the modified inhi-bitor. The low-energy structure of the complex thus obtained is carefully optimized by minimizing the entire complex. In practice, the *insitu* modifications generate variations in the torsion angles and the bond angles of the ligand substituents. Then, in order to avoid steric bumps and to take into account the flexibility of the lateral chains of the residues of the active site of the ligand receptor, a local minimization is carried out (within a radius of 5 Angstroms around the current modification), followed by a global minimization of the receptor-ligand complex to obtain stable structure. The complete description of the computation of relative ligand binding affinity (ΔΔGcom) has been reported earlier**[[28]](#endnote-29)**.

ΔΔ*G*com=Δ*G*com(I)−Δ*G*com(Iref)=ΔΔ*H*MM–ΔΔ*TS*vib+ΔΔ*G*sol(1)

ΔΔ*H*MMdescribes the relative enthalpic contribution to the GFE change related to the intermolecular interactions in E:I complex derived by molecular mechanics (MM), ΔΔ*G*soland ΔΔ*TS*vibrepresent, respectively, the relative solvation GFE and simplified relative vibrational entropy.

**Molecular mechanics**

Simulations of the inhibitors and their complexes models FP2-CNPx were carried out with all-atom representation using atomic and charge parameters of the class II consistent force field CFF**[[29]](#endnote-30)**. The dielectric constant of was set to 4 for all molecular mechanics calculations taking account of the dielectric shielding effect in proteins. Minimizations of the E:I complexes, free E and I were carried out by relaxing the structures gradually, starting with added hydrogen atoms, then with residue side chain heavy atoms and by the protein backbone relaxation. In all the geometry optimiza-tions, a sufficient number of steepest descent and conjugate gradient iterative cycles were used with the convergence criterion for the average gradient set to 0.01 kcal/molÅ−1.

**Conformational search**

Conformational research is a method for calculating the rela-tive energy associated with the conformation of a molecule**[[30]](#endnote-31)**. Its aim is therefore to find the minimum possible and to calculate Boltzmann's population, which gives us infor-mation on the population of occupied levels at a given temperature**[[31]](#endnote-32)**. This method has been described earlier.

**SolvationGibbs free energy**

The process of protein ligand interaction takes place in the biological medium in a solvent. It contributes through hydrogen bonding and solvation phenomena to the bonding process. The electrostatic component of the solvation GFE which includes the effect of ionic strength via solving the nonlinear Poisson-Boltzmann (PB) equation**[[32]](#endnote-33),[[33]](#endnote-34)**was computed by the DelPhi module of Discovery Studio**[[34]](#endnote-35)**. This method has been described fully earlier.

**Interaction energy**

The MM interaction energy (Eint) calculation protocol available in Discovery Studio 2.5 was used to compute the non-bonded interactions (van der Waals and electrostatic interatomic potential terms) between two sets of atoms in the E:I complexes was described earlier.

**Pharmacophore generation**

The Discovery Studio's Catalyst Hypogen algorithm pro-gram**[[35]](#endnote-36)**allowed us, based on the models of the various EI complexes used, to generate the hypothesesto construct a 3D-QSAR pharmacophore as described previously.

**ADME-related properties**

Properties that determine the pharmacokinetics profile of a compound, besides octanol/water partitioning coefficient, aqueous solubility, blood/brain partition coefficient, Caco-2 cell permeability, serum protein binding, number of likely metabolic reactions and other eighteen descriptors related to adsorption, distribution, metabolism and excretion (ADME properties) of the inhibitors were computed by the QikProp program based on the methods of Jorgensen**[[36]](#endnote-37),[[37]](#endnote-38),[[38]](#endnote-39)** as described previously.

**Virtual library generation**

The virtual library (VL) generation was performed as described earlier.

**ADME based library focusing**

The orientation of the virtual library has been made thanks to numerous selection criteria as described above []. Twenty-four pharmacokinetics-related molecular descript-tors available in QikProp, which characterize a wide spectrum of molecular properties as described in the footnote of Table 12, were used. Optimum ranges of these 24 descriptors were defined in terms of upper and lower bounds according to QikProp. Among them predicted drug-likeness (#stars, footnote of Table 12) was used to retain drug-like CNP analogs in the focused VL.

**Pharmacophore based library searching**

PH4 based library screen process was described earlier.

***In silico* screening**

The conformer with the best mapping on the PH4 in each cluster of the focused library subset was selected for *in silico* screening by the complexation QSAR model. The relative GFE of E:I complex formation in water ΔΔ*G*comwas computed for each selected new CNP analogue and then used for prediction of FP2 inhibitory potencies () of the focused VL by inserting this parameter into the *pf*FP2 receptor target-specific scoring function:

= *a*·ΔΔ*G*com+*b* (2).

results

**Training and validation set**

The training set of 12 CNPs and validation set of another 3 analogs (Table 1) were selected from a homogeneous series of Pf-FP2 inhibitors with known experimentally determined inhibitory activities originating from a single laboratory. The whole series was obtained by variations at two positions R1 and R2 on the backbone of carbonitriles pyrimidine as show in Table 1. The experimental biological activities (0.001 ≤ ≤ 0.609 µM) cover a sufficiently wide concentration range for building of a reliable QSAR model. The ratio between the sizes of training and validation sets remains a critical point of correct classification but is limited by the count of the set of homologous compounds available from the literature**[[39]](#endnote-40)**.

**QSAR model**

The relative GFE expressed during the formation of the (E:I) complex was calculated for the FP2-CNPx complexes as described in method a in section 2-2. Table 2 presents the GFE and its components. ΔΔGcom reflects the mutual affinity between the enzyme and the inhibitor. Since it is calculated with an approximate approach, the relevance of the binding model is evaluated by a linear regression, equation (2) analysis which led to a linear correlation with the experimental activity data **20**. A correlation equation obtained for the GFE ΔΔGcom, equation (3) is presented in Table 3 with the relevant statistical data. The relatively high values of the regression coefficient R2 = 97% and the Fischer F-test of the correlation involving ΔΔGcom, indicate a strong relationship between the binding model and the experimental inhibitory power of the CNP series as indicated in Figure 2.

Table 1: Training set (CNP1-12) and validation set (CNPV 1-3) of *Pf*FP2 inhibitors used in the preparation of QSAR models of inhibitor binding. The R1(Bottom) and R2(Top) groups are numbered in the first part of the Table as #R ≡ group index

|  |
| --- |
|  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Training set | CNP1 | CNP2 | CNP3 | CNP4 | CNP5 | CNP6 |
| #R1−#R2 | 1-9 | 4-10 | 5-9 | 6-9 | 4-11 | 4-12 |
| (nM) | 60 | 9 | 40 | 79 | 24 | 2 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Training set | CNP7 | CNP8 | CNP9 | CNP10 | CNP11 | CNP12 |
| #R1−#R2 | 4-13 | 3-9 | 4-14 | 4-15 | 7-9 | 8-9 |
| (nM) | 12 | 9 | 1 | 3 | 609 | 196 |

|  |  |  |  |
| --- | --- | --- | --- |
| Validation set | CNPV1 | CNPV2 | CNPV3 |
| #R1−#R2 | 2-9 | 4-9 | 4-16 |
| (nM) | 24 | 3 | 7 |

Table 2: Complexation Gibbs free energy (binding affinity) and its components for the training set of FP2 inhibitors CNP 1-12 and validation set CNPV 1-3

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Training set a | *M*W b | ΔΔ*H*MM c | ΔΔ*G*sol d | ΔΔ*TS*vib e | ΔΔ*G*com f | g |
| CNP1\* | 286 | 0 | 0 | 0 | 0 | 7.22 |
| CNP2 | 365 | -1.15 | -6.62 | -1.37 | -6.40 | 8.04 |
| CNP3 | 365 | -0.15 | -2.08 | -0.79 | -1.44 | 7.39 |
| CNP4 | 320 | -1.48 | -2.99 | -0.18 | 1.69 | 7.10 |
| CNP5 | 328 | -1.57 | 1.90 | 1.33 | -1.00 | 7.61 |
| CNP6 | 331 | -9.13 | -4.54 | 0.23 | -13.90 | 8.69 |
| CNP7 | 365 | -7.71 | 3.55 | -0.04 | -4.12 | 7.92 |
| CNP8 | 355 | -15.36 | 7.94 | -1.18 | -6.24 | 8.04 |
| CNP9 | 355 | -11.49 | -6.58 | -0.11 | -17.96 | 9.00 |
| CNP10 | 355 | -14.35 | -0.73 | -1.71 | -13.37 | 8.52 |
| CNP11 | 335 | -1.12 | 11.97 | 1.66 | 9.19 | 6.21 |
| CNP12 | 339 | -0.50 | 4.93 | 0.22 | 4.21 | 6.70 |
| Validation set | *M*W | ΔΔ*H*MM | ΔΔ*G*sol | ΔΔ*TS*vib | ΔΔ*G*com | / |
| CNPV1 | 315 | -4.47 | 4.61 | -0.68 | 0.82 | 0.945 |
| CNPV2 | 320 | -3.05 | -4.74 | -1.72 | -6.07 | 0.928 |
| CNPV3 | 354 | -5.22 | -4.15 | -5.08 | -4.29 | 0.947 |

aFor the chemical structures of the training set of inhibitors see Table 1; b𝑀𝑤 is the molecular mass of the inhibitor; cΔΔ𝐻MMis the relative enthalpic contribution to the Gibbs free energy change related to the protease-inhibitor complex formation derived by molecular mechanics (MM): ΔΔ𝐻MM≅ [𝐸MM{PR:CNPx}−𝐸MM{CNPx}]−[𝐸MM{PR:CNP1}−𝐸MM{CNP1}], CNP1 is the reference inhibitor; dΔΔ𝐺solv is the relative solvation Gibbs free energy contribution to the Gibbs free energy change related to protease-inhibitor complex formation: ΔΔ𝐺solv=[𝐺solv{PR:CNPx}−𝐺solv{CNPx}]−[𝐺sol{PR:CNP1}−𝐺sol{CNP1}]; e−ΔΔ𝑇𝑆vib is the relative entro-pic contribution of the inhibitor to the Gibbs free energy related to protease-inhibitor complex formation: ΔΔ𝑇𝑆vib=[𝑇𝑆vib{CNPx}PR−𝑇𝑆vib{CNPx}]−[𝑇𝑆vib{CNP1}PR− 𝑇𝑆vib{CNP1}]; fΔΔ𝐺comp is the relative Gibbs free energy change related to the enzyme-inhibitor complex formation: ΔΔ𝐺comp ≅ ΔΔ𝐻MM+ΔΔ𝐺solv−ΔΔ𝑇𝑆vib. GIC50expis the experimental inhibition constant obtained from [20].

|  |
| --- |
|  |
| Figure 1: Plot of correlation equation between and relative complexation GFE ΔΔGcom of the training set, all in kcal.mol−1. The validation set data points are shown in red color |

Table 3: Regression analysis of computed binding affinities ∆∆Gcom, its enthalpic component ∆∆HMM, and their observed activities = -log10.**20**of (CNPs)

|  |  |
| --- | --- |
| **Statistical Data of Linear Regression** | |
| = -0.1025 × ∆∆Gcom + 7.2867 (3) | ∆∆Gcom |
| Number of compound n | 12 |
| Square correlation coefficient regression R2 | 0.97 |
| LOO cross-validated Square correlation R2xv | 0.97 |
| Standard error of regression σ | 0.141 |
| Statistical significance of regression, Fisher F-test | 363.4 |
| Level of statistical significance α | > 95 % |
| Range of activities [nM] | 1 – 609 |

**Binding mode and interaction energy**

**Binding mode**

The active conformation of the most active of the pyrimidine carbonitriles CNP9 from this QSAR is revealed in Figure 2. However, the side chain P1 of the carbonitriles pyrimidine family comprising the 2-cyanopyrimidine core offers a great possibility of substitution in the pocket S1. The grafting of the halogens, in particular the Br in position 5 of the nucleus, promotes an improvement in the interactions with the residues Trp43, Asn173 and His174. Also, an intensity of hydrophobic interactions is noticed in the S2 pocket, particularly with the residues Ser149, Leu172, Ala175 and Asp-234. Figure 3 presents in 2D and 3D the mode of binding to the active site of the best CNP9 supported by a hydrogen bond with Gly83.

**Interaction Energy**

Subsequently, the distribution of the energy of inhibitory enzymatic intermolecular interaction to the key residues of the various pockets of the active site is listed in Table 4 with their correlation as a function of the experimental activity () plotted in figure 3. A comparative indivi-dual contribution between the most active CNP9 and the least active CNP11 confirming the observed trend of experimental activities by pockets is presented in Table 5.

|  |
| --- |
|  |
|  |
| Figure 2: (top) 2D schematic interaction diagram of the most potent inhibitor CNP9 (Table 1) at pfFP2 active-site; (bottom) 3D structure of the active-site with bound inhibitor CNP9. |

|  |
| --- |
|  |
|  |
| Figure 3: Plot of the correlation between the residue of the active site Leu172 and the contribution of S2 to the intermolecular interaction energy of *pf*FP2-CNPx |

|  |
| --- |
|  |
| Figure 4: Molecular mechanics intermolecular interaction energy Eint breakdown (in kcal.mol-1) to residue contributions for CNP9 and CNP11, Table 2.2 |

Table 4: Active site residue contribution to Eint(in kcal/mol)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Pockets** | **CNP** | | | | | | | | | | | | |
|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
|  | 7.2 | 8.1 | 7.4 | 7.1 | 7.6 | 8.7 | 7.9 | 8.1 | **9** | 8.5 | **6.2** | 6.7 |
| **S1** | Gnl36 | -0.1 | -0.1 | -0.1 | 0.0 | -0.1 | -0.1 | -0.2 | -0.1 | **-0.1** | -0.1 | **-0.1** | -0.1 |
| Gly40 | -0.7 | -1.1 | -0.7 | -0.8 | -1.2 | -1.1 | -1.7 | -0.8 | **-0.3** | -0.2 | **-0.7** | -0.7 |
| Ser41 | -0.2 | -0.3 | -0.2 | -0.3 | -0.3 | -0.3 | -0.4 | -0.3 | **-0.2** | -0.1 | **-0.2** | -0.2 |
| Cys42 | -0.7 | -0.5 | -0.7 | -0.7 | 3.8 | -0.6 | -1.0 | -0.7 | **-0.6** | -1.3 | **-0.7** | -0.7 |
| Trp43 | -1.4 | -1.4 | -1.5 | -1.5 | -1.4 | -1.3 | -1.4 | -1.5 | **-1.8** | -2.2 | **-1.5** | -1.5 |
| Cys80 | -0.3 | -0.3 | -0.3 | -0.3 | -0.4 | -0.4 | -0.4 | -0.3 | **-0.1** | -0.1 | **-0.2** | -0.3 |
| Asn81 | -1.3 | -1.4 | -1.2 | -1.3 | -1.4 | -1.4 | -1.3 | -1.3 | **-0.7** | -0.7 | **-1.2** | -1.3 |
| **Total** | **-4.6** | **-5.1** | **-4.7** | **-4.8** | **-0.9** | **-5.1** | **-6.4** | **-4.8** | **-3.8** | **-4.7** | **-4.5** | **-4.7** |
| **S2** | Leu84 | -2.6 | -3.0 | -3.1 | -3.3 | -2.9 | -3.0 | -3.0 | -3.0 | **-3.2** | -2.8 | **-3.0** | -2.7 |
| Ile85 | -0.1 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | **-0.3** | -0.4 | **-0.1** | -0.1 |
| Ser149 | -0.1 | -0.3 | -0.2 | -0.2 | -0.3 | -0.2 | -0.2 | -0.3 | **-0.5** | -0.7 | **-0.1** | -0.1 |
| Leu172 | -3.8 | -4.4 | -4.3 | -4.5 | -4.4 | -4.8 | -4.4 | -4.3 | **-5.3** | -4.2 | **-3.1** | -3.2 |
| Ala175 | -0.2 | -0.4 | -0.3 | -0.3 | -0.3 | -0.4 | -0.3 | -0.3 | **-0.9** | -0.8 | **-0.2** | -0.2 |
| Asp-234 | -1.3 | -1.1 | -1.4 | -1.0 | -0.9 | -1.0 | -1.2 | -1.0 | **-1.3** | -1.3 | **-1.1** | -0.7 |
| **Total** | **-8.1** | **-9.3** | **-9.4** | **-9.5** | **-8.9** | **-9.5** | **-9.2** | **-9.1** | **-11.6** | **-10.2** | **-7.7** | **-7.1** |
| **S’1** | Val150 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | **-0.3** | -0.5 | **-0.1** | -0.1 |
| Val152 | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 | **-0.1** | -0.1 | **-0.1** | -0.1 |
| Asn173 | -3.2 | -3.9 | -3.3 | -3.4 | -3.0 | -4.0 | -4.5 | -3.4 | **-3.3** | -3.5 | **-3.0** | -3.1 |
| Hisd174 | -1.0 | -1.5 | -1.1 | -1.1 | -1.5 | -1.2 | -1.3 | -1.1 | **-1.7** | -1.5 | **-1.0** | -1.0 |
| Trp206 | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 | -0.1 | **-0.1** | -0.1 | **0.0** | -0.1 |
| **Total** | **-4.5** | **-5.8** | **-4.8** | **-4.9** | **-5.0** | **-5.6** | **-6.2** | **-4.9** | **-5.5** | **-5.6** | **-4.2** | **-4.3** |
| **S3** | Lys+76 | 0.2 | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | 0.2 | 0.2 | **0.2** | 0.2 | **0.1** | 0.0 |
| Asn77 | -0.6 | -0.5 | -0.6 | -0.5 | -0.5 | -0.5 | -0.5 | -0.5 | **-0.4** | -0.5 | **-0.7** | -0.6 |
| Tyr78 | -1.4 | -1.1 | -1.5 | -1.5 | -1.5 | -1.5 | -1.5 | -1.5 | **-1.6** | -1.8 | **-1.5** | -1.5 |
| Gly82 | -2.5 | -2.6 | -2.5 | -2.4 | -2.6 | -2.5 | -2.9 | -2.5 | **-1.9** | -2.7 | **-2.6** | -2.7 |
| Gly83 | -4.0 | -4.1 | -4.1 | -4.2 | -4.1 | -4.2 | -4.1 | -4.1 | **-4.3** | -4.1 | **-3.9** | -3.9 |
| **Total** | **-8.3** | **-8.1** | **-8.5** | **-8.5** | **-8.7** | **-8.5** | **-8.8** | **-8.5** | **-8.0** | **-8.9** | **-8.6** | **-8.6** |

From the analysis of Table 4, as it has been noted for the azadipeptides nitrile (ADPN) in previous study, here in the case of pyrimidine carbonitriles (CNP) a relevant stabilizing drop of S2 pocket residues contributions to Enzyme – Inhibitor interaction energy (Eint) from the least active TS CNP11 to the best active one CNP9 by almost 4 kcal.mol-1i.e. 50%, close to the 45% increase of their observed inhibitory potency from 6.21 to 9 respectively (see Table 5 and Figure 3 Top). Therefore the preeminent role of the hydrophobic S2 pocket of *Pf*FP2 active site in the design of potent molecules against *Pf*FP2 is confirmed again as previously reported**[[40]](#endnote-41)** devoting a centralrole of Leu172 in that S2 pocket (Figures 3 Bottom and 4).

Table 5: Variation of the overall interaction energy per pocket of the active site between the most active and the least active of the inhibitors

|  |  |  |  |
| --- | --- | --- | --- |
| Pockets | Eint(kcal/mol) | | ΔEint (kcal/mol) |
| CNP9 ( | CNP11 ( |
| S1 | -3.83 | -4.50 | 0.67 |
| **S2** | **-11.50** | **-7.68** | **-3.82** |
| S’1 | -5.46 | -4.18 | -1.28 |
| S3 | -8.02 | -8.59 | 0.57 |

**QSAR Pharmacophore Model**

The pharmacophore model of CNP was developed according to the same method used for ADPNs study QSAR 3D Pharmacophore. It was generated from the active conformation of CNPA compounds in the active site of FP2. These compounds covering a wide range of experimental activity (0.001-0.609 nM). The 3D-QSAR PH4 generation was carried out in three steps: the constructive, the subtractive and the optimization step. During the constructive phasethe most active compounds of the training set, that which ≤ 1.3 × 0.001 = 0.0013 µM are automatically selected as "lead" or "serial head". Thus the most active compound of the training set, CNP9 whose inhibitory activity ( = 0.001 µM) is the only one to fulfill this condition served to generate the active centers of the initial PH4. The subtractive phase, in which the inactive compounds, that is to say those whose inhibitory activity fulfills the condition (> 0.001 × 103.50M = 3.16225M) serving to identify the active centers to be deleted from the initial PH4, does not present any inactive molecule. The results of the 10 best hypotheses are presented in Table 6 showing the cost, the RMSD and the correlation coefficient between predicted and experimental activities.

The CNPs series presents the costs of the 10 respective penalties of PH4 in the range between 68.812 (Hypo1) to 84.894 (Hypo10). The relatively small difference between the costs of the extreme hypotheses clearly reflects the homogeneity and the consistency of the training set used to produce them. The difference of gap ∆ = 280.644 between the fixed coast (32.305) and the null coast (312.949) greater than 70 is the supreme indicator of the quality and the predictive character of the PH4. This indicates the probability that the correlation between the activity values estimated by the PH4 and those experimental is real at more than 90%. The other standard statistical indicators such as the root mean square deviation (RMSD) of the various hypotheses, between 2.323 and 2.586 and the coefficient of determination (R2) between 0.942 and 0.911 (see Table 6) make it possible to retain the first hypothesis of PH4 for the screening of the virtual library of CNP analogues.Figure 6 presents the geometric characteristics of Hypothesis 1 (Hypo 1) of the FP2 inhibition pharmacophore.

|  |  |
| --- | --- |
|  |  |
|  |  |
|  | |
| Figure 5: (top) Coordinates, (middle) featuresand mappingof the FP2 inhibitor pharmacophore with the best training set CNP9 (yellow). The correlation plot of experimental vs. predicted inhibitory activity is displayed at the bottom. The features are colored blue for hydrophobic aliphatic (HYD\_AL), green for hydrogen-bond acceptor (HBA), and cyan for hydrophobic (HYD). The arrows represent the projection for donor and acceptor features. | |

Table 6: Output parameters of 10 generated PH4 hypotheses for FP2 inhibitors after CatScramble validation procedure.

|  |  |  |  |
| --- | --- | --- | --- |
| **Hypothesis** | **RMSD a** | **R2 b** | **Total cost c** |
| Hypo1 | 2.323 | 0.942 | 68.812 |
| Hypo2 | 2.592 | 0.927 | 72.134 |
| Hypo3 | 2.528 | 0.931 | 74.696 |
| Hypo4 | 2.490 | 0.933 | 74.868 |
| Hypo5 | 2.494 | 0.933 | 75.536 |
| Hypo6 | 2.524 | 0.931 | 76.024 |
| Hypo7 | 2.726 | 0.919 | 80.335 |
| Hypo8 | 2.756 | 0.918 | 81.630 |
| Hypo9 | 2.728 | 0.919 | 82.198 |
| Hypo10 | 2.856 | 0.911 | 84.894 |
| Fixed Cost | 0.0 | 0.0 | 32.305 |
| Null Cost | 6.949 | 0.0 | 312.949 |

aroot mean square deviation (RMSD); bsquared correlation coefficient; coverall cost parameter of the PH4

The best-selected hypothesis Hypo1 represents with a probability of 91% a PH4 model with a similar level of predictive power as the complexation GFE enzyme-inhibitor binding QSAR model. The resulting regression equation expresses as a function of , estimated by Hypo1: = 0.9366 x + 0.2849 are listed in table 7 (n = 12; R2 = 0.91; R2xv = 0.91; F-test = 114.34; σ = 0.551; α > 95%) and its graph is presented in Figure 5 above. Once again, the predictive power of the PH4 model, like that of the QSAR, is confirmed and both present interesting predictive powers. The information obtained from the RQSA and PH4 inhibition models relating to the hierarchy of mechanisms governing the activity of CNPs will be useful for filling the S1 and S2 pockets.

Table 7: Regression analysis ofand computed of CNPs toward *Pf*FP2

|  |  |
| --- | --- |
| Statistical Data of Linear Regression for Hypo 1 | |
| **= 0.9366 x  + 0.2849** |  |
| Number of compoundn | 12 |
| Square correlation coefficient regression R2 | 0.91 |
| LOO cross-validated Square correlation R2xv | 0.91 |
| Standard error of regression σ | 0.551 |
| Statistical significance of regression, Fisher F-test | 114.34 |
| Level of statistical significance α | > 95 % |
| Range of activities [nM] 1- 609 | |

**Virtual screening**

*In silico* screening of a virtual (combinatorial) library can lead to hit identification as it was shown in previous works on inhibitors design**[[41]](#endnote-42)**. From the different groups listed in Table 8, we have created a virtual combinatorial library by substitution in positions R1, R2, R3 and R4 on the pyrimidine nucleus and its side chain, the size of which is R1 x R2 x R3 x R4 = 14 x 17 x 14 x 25 = 83.300 CNPA analogs (CNPs analogs). This virtual combinatorial library has been filtered in order to retain the analogue which can be administered orally and which have a very good predictive pharmacokinetic profile. Following this filtering, a reduced virtual combinatory library of 32.856 analogues was obtained which represents 39.44% of its initial population. It was then screened from the generated PH4 3D-RQSA model of Hypo1. Only 127 analogues having a better alignment with the PH4 model were subjected to a last evaluation of their theoretical experimental activity or predicted with the complexation method ( is calculated from the correlation equation (3) Table 3). Finally, 52 analogues with better scores are selected and presented in table 8.

We have prepared histograms of frequency of appearance of groups R1, R2, R3 and R4 for the 52 CNPA. On the analysis of these histograms (Figure 6), it emerges that the fragments of greater occurrences (value in parentheses) are : 1(49), 2(1), 3(1), 4(1) in R1 for filling the S3 pocket ; 15(32), 31(20) in R2 and 38(14), 39(10), 40(12), 41(9) at R3 for filling the pocket S1 and finally preferentially the fragments 46(5), 49(4), 50(6), 51(4), 52(4), 53(5), 54(6), 55(6) and 56(6) in R4 for filling the pocket S2. The frequency of the occurrences of the groups R1, R2, R3 and R4 of the 52 best analogs following the superposition at PH4 is presented in following figure.

Table 8: Substituents used to generate the VL of CNP analogs



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | \*R1 |  | R2 |  | R3 |
| 1 | –C6H11 | 15 | –CH3 | 32 | -C3H7O |
| 2 | –C6H10OH | 16 | –OH | 33 | –CONH2 |
| 3 | –Me(C6H11) | 17 | –CH2F | 34 | –CH2CO2 |
| 4 | 1,2-diMe(C6H11) | 18 | –CH2Cl | 35 | –C2H5CONH2 |
| 5 | –Phe | 19 | –NO | 36 | –OC3H7 |
| 6 | –Me(phe) | 20 | –NO2 | 37 | –OC2H5 |
| 7 | Piperidine | 21 | –CH2I | 38 | –C4H9 |
| 8 | 2-Me(piperidine) | 22 | –CH2Br | 39 | –C4H9OH |
| 9 | 3-Me(piperidine) | 23 | –F | 40 | –C5H11 |
| 10 | 4-Me(piperidine) | 24 | –Cl | 41 | –C5H11OH |
| 11 | Pyridine | 25 | –Br | 42 | –C5H11F |
| 12 | 2-Me(pyridine) | 26 | –I | 43 | –C5H11Cl |
| 13 | 3-Me(pyridine) | 27 | –CCl3 | 44 | –C5H11Br |
| 14 | 4-Me(pyridine) | 28 | –CF3 | 45 | –C5H11I |
|  |  | 29 | –CBr3 |  |  |
|  |  | 30 | –CI3 |  |  |
|  |  | 31 | –H |  |  |
| R4 | | | | | |
| 46 | 2,4-diMe(C6H12)carboxamide | | | | |
| 47 | 2,3-diMe(C6H12)carboxamide | | | | |
| 48 | 4-Et-2-Me(C6H12)carboxamide | | | | |
| 49 | But(C6H12) | | | | |
| 50 | 1-But-2-Me(C6H12) | | | | |
| 51 | 1-But-3-Me(C6H12) | | | | |
| 52 | 1-But-4-Me(C6H12) | | | | |
| 53 | 2-But(piperidine) | | | | |
| 54 | 2-But-3-Me(piperidine) | | | | |
| 55 | 2-But-4-Me(piperidine) | | | | |
| 56 | 2-But-5-Me(piperidine) | | | | |

\*Fragments 1 to 14 have also been used for R4

Table 9: complexation Gibbs free energy and its components and predictive activities of the best new analogs.

| **Analogs**  **Designed** | **Mw a** | **ΔΔHMM b** | **ΔΔGsol c** | **ΔΔTSvib d** | **ΔΔGcom e** | **f** |
| --- | --- | --- | --- | --- | --- | --- |
| CNP9 | 332 | 0.00 | 0.00 | 0.00 | 0.00 | 1000g |
| 1-15-36-46 | 486 | -5.52 | -1.68 | 11.13 | -18.33 | 680 |
| 2-15-37-46 | 488 | -3.84 | 0.37 | 10.11 | -13.58 | 2000 |
| 3-15-37-46 | 486 | -4.06 | -1.59 | 11.99 | -17.64 | 800 |
| 1-15-36-47 | 486 | -0.16 | -1.08 | 13.02 | -14.26 | 1700 |
| 4-15-37-46 | 500 | -3.20 | 14.47 | 13.12 | -1.85 | 33300 |
| 1-15-36-48 | 500 | -24.92 | 8.33 | 11.83 | -28.42 | 63 |
| 1-31-36-46 | 486 | -22.31 | 8.48 | 09.92 | -23.75 | 190 |
| 1-15-38-48 | 498 | -25.17 | 8.08 | 15.52 | -32.61 | 23 |
| 1-31-38-48 | 484 | -22.13 | 8.22 | 14.49 | -28.40 | 63 |
| 1-15-38-49 | 469 | -14.68 | 6.17 | 18.34 | -26.85 | 92 |
| 1-15-38-50 | 483 | -15.78 | 5.29 | 18.56 | -29.05 | 54 |
| 1-31-38-50 | 469 | -16.30 | 5.93 | 14.01 | -24.38 | 160 |
| 1-15-38-51 | 483 | -14.61 | 6.76 | 16.30 | -24.15 | 170 |
| 1-15-38-52 | 483 | -14.66 | 6.25 | 19.09 | -27.50 | 78 |
| 1-15-38-53 | 470 | -14.19 | 7.13 | 16.99 | -24.05 | 170 |
| 1-15-38-54 | 484 | -19.13 | 6.51 | 18.32 | -30.94 | 34 |
| 1-31-38-54 | 470 | -16.19 | 6.93 | 14.49 | -23.75 | 190 |
| 1-15-38-55 | 484 | -18.22 | 6.29 | 18.29 | -30.22 | 41 |
| 1-31-38-55 | 470 | -18.54 | 6.30 | 15.35 | -27.59 | 76 |
| 1-15-38-56 | 484 | -18.65 | 6.39 | 18.33 | -30.59 | 37 |
| 1-31-38-56 | 470 | -14.66 | 6.91 | 15.44 | -23.19 | 210 |
| 1-15-39-49 | 485 | -11.74 | 7.25 | 18.44 | -22.93 | 230 |
| 1-15-39-50 | 499 | -13.87 | 6.03 | 17.89 | -25.73 | 110 |
| 1-15-39-51 | 499 | -14.11 | 7.26 | 19.80 | -26.65 | 95 |
| 1-15-39-52 | 499 | -11.10 | 7.63 | 20.11 | -23.58 | 190 |
| 1-15-39-53 | 486 | -13.96 | 7.72 | 18.49 | -24.73 | 150 |
| 1-15-39-54 | 500 | -18.23 | 6.53 | 19.32 | -31.02 | 34 |
| 1-31-39-54 | 486 | -17.04 | 7.40 | 13.92 | -23.56 | 190 |
| 1-15-39-55 | 500 | -18.77 | 8.04 | 16.64 | -27.37 | 80 |
| 1-15-39-56 | 500 | -15.82 | 6.96 | 19.65 | -28.51 | 61 |
| 1-31-39-56 | 486 | -16.12 | 8.08 | 14.26 | -22.30 | 260 |
| 1-15-40-49 | 483 | -13.99 | 6.15 | 19.37 | -27.21 | 83 |
| 1-15-40-50 | 497 | -18.27 | 5.57 | 17.65 | -30.35 | 40 |
| 1-31-40-50 | 483 | -15.60 | 5.77 | 16.40 | -26.23 | 100 |
| 1-15-40-51 | 497 | -12.31 | 6.50 | 18.65 | -24.46 | 160 |
| 1-15-40-52 | 497 | -13.23 | 5.96 | 19.72 | -26.99 | 88 |
| 1-15-40-53 | 484 | -14.29 | 7.20 | 18.71 | -25.80 | 170 |
| 1-15-40-54 | 498 | -20.06 | 5.56 | 17.61 | -32.11 | 26 |
| 1-31-40-54 | 484 | -18.04 | 6.02 | 17.26 | -29.28 | 51 |
| 1-15-40-55 | 498 | -20.72 | 6.75 | 20.77 | -34.74 | 14 |
| 1-31-40-55 | 484 | -19.33 | 6.63 | 15.99 | -28.69 | 59 |
| 1-15-40-56 | 498 | -18.82 | 6.62 | 19.77 | -31.97 | 27 |
| 1-31-40-56 | 484 | -15.02 | 6.84 | 18.30 | -26.48 | 99 |
| 1-15-41-49 | 499 | -14.37 | 7.10 | 20.57 | -27.84 | 23 |
| 1-31-41-50 | 499 | -13.38 | 5.43 | 17.73 | -25.68 | 120 |
| 1-31-41-51 | 499 | -9.67 | 6.79 | 19.14 | -22.02 | 280 |
| 1-31-41-52 | 499 | -12.63 | 7.37 | 19.75 | -25.01 | 140 |
| 1-15-41-53 | 500 | -17.86 | 7.59 | 22.00 | -32.27 | 25 |
| 1-31-41-53 | 486 | -13.53 | 7.65 | 17.66 | -23.54 | 19 |
| 1-31-41-54 | 500 | -16.32 | 6.48 | 19.06 | -28.90 | 56 |
| 1-31-41-55 | 500 | -14.09 | 7.66 | 17.78 | -24.21 | 170 |
| 1-31-41-56 | 500 | -14.46 | 7.25 | 18.35 | -25.56 | 120 |

aMW (g/mol) is molecular mass of the inhibitor; b∆∆HMM(kcal/mol) is the relative enthalpic contribution to the Gibbs free energy change related to the FP2-CNP complex formation ∆∆Gcom(for details see footnote of Table 2); c∆∆Gsol(kcal/mol) is the relative solvation Gibbs free energy contribution to ∆∆Gcom; d∆∆TSvib (kcal/mol) is the relative entropic (vibrational) contribution to ∆∆Gcom; e∆∆Gcom(kcal/mol) is the relative Gibbs free energy change related to the enzyme-inhibitor FP2-CNP complex formation ∆∆Gcom–∆∆HMM+∆∆Gsol+∆∆TSvib; f(pM) is the predicted half-maximal inhibitory concentration of CNPx towards *Pf*FP2 calculated from ∆∆Gcomusing correlation Equation (B), Table 3; gis given for the reference inhibitor CNP9 instead of (pM)[].

**Novel CNPs analogs**

The predictive activities of the best 52 new analogs of CNPA calculated from the correlation equation (3) are better and more powerful than that of the most active compound CNP9 of the training set () proposed by Cotereon *et al.*[] are presented in table 9.

The best CNP analogs (CNPA)with their predicted activity value ( ) in brackets are: 1-15-38-48 (23 pM); 1-15-38-54 (34 pM); 1-15-38-56 (37 pM); 1-15-39-54 (34 pM); 1-15-40-50 (40pM); 1-15-40-54 (26 pM); 1-15-40-55 (14 pM); 1-15-40-56 (27 pM); 1-15-41-49 (23 pM); 1-15-41-53 (25 pM); 1-31-41-53 (19 pM). The most active of the CNPAs, namely 1-15-40-55 (14 pM), has a predicted potency () of approximately 70 times greater than that of the best CNP9 inhibitor of the training set.

|  |  |
| --- | --- |
|  |  |
|  |  |
| Figure 6: Histograms of frequency of occurrence of the individual R-groups in the 52 best analogues selected and mapped to the four characteristics of the ph4 hypothesis of the hypo1 pharmacophore | |

**Pharmacokinetic profile of the best analogs**

The pharmacokinetic profile of the best designed CNPs has also been calculated and compared with those of the drugs used for the treatment of malaria alone or in combination with Artemisinin (CTA) or in clinical trials (Table 10).

The value of the drug likelihood descriptor #stars indicates that CNPs analogs have 24 descriptors that comply with those of 95% of drugs and have a better profile than the majority of ACTs. These results also present the percentage of absorption by oral route (%HOA) of the best analogues which is greater than or equal to 80% and also obey the WHO recommendation namely %HOA ⩾ 80%. Therefore, the best analogues identified in this work are potential drug candidates.

**DISCUSSION**

The study of the binding mode of *Pf*FP2:CNP from its RQSA model of complexation with a single descriptor and the 3D-QSAR PH4 pharmacophore model generated allowed us to access major structural information on the molecular complementarities of the enzyme and the inhibitor. The visual analysis and the calculation of the interactions between *Pf*FP2 and CNPs in the active site of the enzyme guided us in our efforts to design a virtual combinatorial library of new CNP analogues with four substitutions on the CNP scaffold at positions R1, R2, R3 and R4. A resulting targeted library filtered by a set of descriptors linked to ADME and screened by mapping the analogues to the PH4 pharmacophore, allowed the selection of a subset of libraries of CNPS bioavailable orally.

|  |
| --- |
|  |
| Figure 7: Presentation of some designed analogs of CNPA |

|  |
| --- |
|  |
|  |
| Figure 8 (Top) Close up of virtual hit 1-15-40-55 at the active-site of FP2. (Bottom) Connolly surface of the active-site of *Pf*FP2 with bound predicted most active CNP inhibitor 1-15-40-55. The binding site surface is colored according to residue hydrophobicity: red—hydrophobic, blue—hydrophilic, and white—intermediate. |

**QSAR model**

The robustness of this QSAR model with a descriptor is evaluated through the components of the GFE (ΔΔGcom), namely the contribution of the ΔΔHMM enthalpy, the ΔΔGsolv solvation and the loss of vibrational entropy ΔΔTSvib during the binding of the CNPs. The enthalpy contribution to GFE, then taking into account the effect of the solvent in order to get closer to the biological medium maintains the level of strong relationship between the experimental data and the simulation results. Finally, the likelihood of the model is increased by the loss of the inhibitory vibrational entropy TSvib to explain approximately 97% of the variations of the by that of the GFE. This last contribution is one of the most reliable indicators of the predictive power of the QSAR model as reported by Freire *et al.***[[42]](#endnote-43)**. Consequently, the correlation equation (3) and the calculated quantities ΔΔGcomp can be used respectively to predict the inhibitory potencies with respect to pfFP2 for new CNP analogues, since they share the same binding mode as the compounds of the corresponding validation set. The quality of the fixation model is also confirmed by the ratio between the calculated inhibitory activities and the values of the experimental activities and documents the considerable predictive power of the RQSA model of complexation (/, being calculated by equation 3, Table 2 which is close to 1 for the CNPs of the validation set see Table 2)

**Binding mode of inhibitors**

In addition to the robustness of the QSAR model, an analysis of the interactions between FP2 and carbonitriles pyrimidine allowed us to reveal key interactions justifying the affinity of the FP2:CNPx complexes. As shown in Figure 2 in 2D and in 3D, the mode of binding to the active site of the best CNP9 is supported by a hydrogen bond with Gly83 and hydrophobic contacts. CNPs unlike ADPNs have a pyrimidine core in P1 offering a multiple possibility of substitution. Since the fragments at P3 are identical for all the inhibitors, the halogenation of the C5 carbon of the pyrimidine nucleus at P1, in particular by Br, has the effect of stabilizing the complex, favoring an improvement in the interactions with the Trp43, Asn173 and His174 residues. This allows a better orientation of the hydrophobic fragments P2 in the S2 pocket facilitating interactions particularly with the residues with Ser149, Leu172, Ala175 and Asp234 which results in an increase in biological activity in accordance with the work of Löser and al.. Just like the ADPNs, the pyrimidine carbonitriles exhibit the same similarities of attachment to the active site of FP2. For the design of new analogs, we calculated the interaction energy (Eint) between each residue of the active site (Table 4) in order to verify if other interactions not displayed must be taken into account in the description of the mode of binding of the CNPs to the active site of the FP2. The structure of these compounds with a pyrimidine core highlighted the structural characteristics of the binding affinity and paved the way for the design of FP2 inhibitors taking advantage for the filling of the S1, S’1, S2 and S3 pockets. In this way, the key residues of these pockets contribute to the overall FP2–CNPx interaction energy. But specifically, the filling of the S2 pocket illustrates the overall impact of the inhibition of FP2 to more than 70% as illustrated in Figure 3 and also as reported previously**,**. In contrast, Figure 4 displaying the comparative contribution of each residue of the active site to the Eint for the most active CNP9 and the least active CNP11 of the training set confirms the observed trend of the but cannot justify the large difference in their inhibitory potencies. However, a recent work has succeeded in justifying the observed jump of 37.5% in the experimental biological activity between methylphosphonic arginine and the Leucyl aminopeptidase inhibitors (*Pf*A-M17) derived from hydroxamic acid by the contribution of the residues of the enzymatic active site to the Eint at a level of 35%**[[43]](#endnote-44)**. Therefore, the essential structural information in the design of new powerful analogues of CNP will not only be derived from the Eint but also from a more predictive descriptor. The analogs will be selected by virtual screening from a diversified virtual library of analogues with the hydrophobic contact S2 as the central structural requirement displayed by the Pharmacophore model of inhibition of FP2 provided by the single descriptor QSAR model (GFE) (Table 3) (Figure 2).

**Analysis**of new inhibitors from *in Silico* screening

The PH4-based screening of the virtual combinatorial library of CNP analogues has resulted in the identification of new compounds with better predicted activities. These anologues display solid theoretical values reaching 14 picomolars for the most powerful of the analogs ((1-15-40-55; =0.014 nM), which is approximately 70 times higher than that of the best CNP9 inhibitor of the training set. The representation of the best designed analogue in 3D interaction with FP2 (on the left) and the Connolly surface of the binding site (on the right) in Figure 8 shows that the methylpropane of the training set at P3 is replaced by cyclohexane (R1=1) ensuring great stability in the S3 pocket. Since the pyrimidine nucleus is conserved at P1, two substitutions have been made at positions 4 and 5. The bromine (Br) and hydrogen (H) atoms have been replaced respectively by methyl (R2 =15) and pentyl (R3 = 40) inducing an increase in the hydrophobic interactions at S1 with the Trp43 residue in particular. The lipophilic pocket S2 contains a larger 2-But-4-Me(piperidine) (R4= 55) in place of the cyclohexane fragment, which leads to an increase in the hydrophobic contact equivalent to better stabilization and greater affinity. Just like the nitrile azadipeptides, the results of the carbonitriles pyrimidine family are in agreement with the structural information reported by the experimental study Structure Activity Relationship (SAR) of carbonitriles pyrimidine inhibitors of FP2 and FP3 and the RQSA complexation based on the in silico design of inhibitors of nitrile dipeptides of FP3; The major contribution of the substitutions on the different positions has led to the new CNPA analogs resulting in a significant increase in their activity inhibitory.

**ADME-related properties**

The properties related to ADME described in section 2-8 have been calculated and the profiles of the best CNPs analogs designed are presented in Table 10 in comparison with those of the drugs used for the treatment of malaria alone or in combination with Artemisinin (CTA) or in clinical trials (Table 10). With regard to our results, most of the CNP analogues designed are assumed to have a good level of drug likelihood with the descriptor #stars which lies in the validation interval between 0 and 5, just like the available antimalarial. It is also noted for all the analogues a strong human oral absorption in the gastrointestinal tract (HOA). None of them is outside the range of good oral bioavailability (< 25% - poor absorption;> 80% - good absorption) which is one of the main requirements for new antimalarial drugs, as declared by the WHO. This is also the case for the antimalarial drugs used.

**CONCLUSION**

The evaluation of key structural information on FP2 inhibition from FP2:CNPx complexes or from the in situ modification of an existing generic FP2 inhibitor constitutes a reliable means for designing powerful, bioavailable and favorable oral pharmacokinetic antimalarial. From the series of 15 CNPs inhibitors (12 for the training set and 3 for the validation set), we have, despite the low diversity of substitutions at the P1, P2 and P3 positions, established an RQSA model with a descriptor correlating the variation in the free enthalpy of complexation ΔΔGcom calculated during the formation of the complexes with the experimental activities for this family of series of inhibitors. This model sheds light on the main interactions having an impact on the activity. Subsequently, a 3D-RQSA pharmacophore protocol was used to prepare a four-function pharmacophore model (PH4); a model derived from the conformations of the inhibitors of this series linked to the active site of FP2. The robustness of the inhibitory activity PH4, due to the quality of the structural information on the inhibition of FP2 and having a good predictive level, made it possible to screen 83.300 CNPA analogues in order to design more powerful inhibitors of FP2 available orally. This was possible with the help of several substitutions on their molecular skeleton, in particular with the presence of the pyrimidine nucleus. The design strategy is based on the hydrophobic contact at the level of the S2 pocket in the best 3D model-RQSA PH4 generated while considering the size of the fragment in S1 comprising the 2-cyanopyrimidine nucleus offering a great possibility of substitution and respecting the hydrophobic spheres in S’1 and S3. The virtual combinatorial library was filtered by a set of pharmacokinetic descriptors and then subjected respectively to a virtual screening by pharmacophore. New analogues were selected, namely 52 CNPAs, which were evaluated again by complexation in order to determine theoretical inhibition constant molecules of the picomolar order. Presenting a good predictive pharmacokinetic profile, the best designed analogues with their predicted activity in brackets are: (Table 10): 1-15-38-48 (23 pM); 1-15-38-54 (34 pM) ; 1-15-38-56 (37 pM); 1-15-39-54 (34 pM); 1-15-40-50 (40pM); 1-15-40-54 (26 pM); 1-15-40-55 (14 pM); 1-15-40-56 (27 pM); 1-15-41-49 (23 pM); 1-15-41-53 (25 pM); 1-31-41-53 (19 pM). The activities of these analogs being located in the low picomolar range, are recommended for the synthesis and evaluation of the activity in order to verify the consistency of the structural information provided and to guide the future discovery of peptidomimetic inhibitors of FP2. In addition, although limited to the theoretical framework, the approach used in this work helps to understand the studied systems in order to better prepare them for validation by biological tests. This technique is to be encouraged for research and for the medical application of pathologies not yet controlled.

Table 10: Pharmacokinetic profile of some of the best CNPs analogues designed computed by QikProp

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CNPs a** | **#stars b** | **Mwc** | **Smold** | **Smol,hfoe** | **Vmolf** | **RotB g** | **HBdonh** | **HBacci** | **logPo/wj** | **logSwatk** | **logKHSAl** | **logB/B m** | **BIPcaco n** | ***#meta* o** | p | **HOAq** | **%HOAr** |
| 1-15-38-48 | 1 | 499 | 635.3 | 204.7 | 1655.3 | 11 | 3 | 9.5 | 3.4 | -6.6 | 0.4 | -2.4 | 61 | 4 | 23 | 1 | 100 |
| 1-31-38-48 | 0 | 485 | 601.4 | 209.6 | 1610.7 | 11 | 3 | 9.5 | 3.106 | -6.4 | 0.2 | -2.4 | 57 | 3 | 63 | 2 | 100 |
| 1-15-38-56 | 1 | 485 | 737.1 | 129.1 | 1680.0 | 12 | 2 | 8.5 | 4.727 | -7.2 | 0.9 | -1.2 | 147 | 2 | 37 | 1 | 100 |
| 1-15-39-52 | 1 | 500 | 664 | 137.7 | 1639.1 | 14 | 2 | 8.7 | 4.885 | -6.6 | 0.7 | -1.6 | 489 | 3 | 190 | 1 | 100 |
| 1-15-40-53 | 0 | 485 | 656.7 | 116.1 | 1581.9 | 13 | 2 | 8.5 | 4.241 | -5.2 | 0.7 | -0.9 | 195 | 2 | 150 | 1 | 86 |
| 1-15-40-54 | 0 | 499 | 708.6 | 90.3 | 1640.1 | 13 | 2 | 8.5 | 4.778 | -5.6 | 0.8 | -0.7 | 343 | 2 | 26 | 2 | 87 |
| 1-15-40-55 | 1 | 499 | 660.1 | 119.3 | 1626.4 | 13 | 2 | 8.5 | 4.504 | -5.3 | 0.8 | -1.0 | 182 | 2 | 14 | 2 | 96 |
| 1-15-40-56 | 0 | 499 | 688.3 | 121.9 | 1665.0 | 13 | 2 | 8.5 | 4.71 | -5.6 | 0.9 | -1.0 | 172 | 2 | 27 | 2 | 62 |
| 1-15-41-49 | 1 | 500 | 683.2 | 104.2 | 1633.2 | 13 | 3 | 8.5 | 4.72 | -5.5 | 0.4 | -1.2 | 935 | 2 | 23 | 2 | 98 |
| 1-15-41-53 | 0 | 500 | 681.2 | 113.5 | 1642.1 | 13 | 3 | 8.5 | 3.41 | -6.0 | 0.4 | -1.2 | 172 | 2 | 25 | 1 | 94 |
| 1-31-41-53 | 1 | 487 | 687.4 | 106.4 | 1611.1 | 13 | 2 | 8.5 | 4.17 | -5.3 | 0.6 | -1.3 | 195 | 2 | 19 | 1 | 87 |
| Amodiaquine | 1 | 334 | 603.2 | 131.7 | 1018.7 | 6 | 0 | 5 | 3.6 | -4.8 | 0.0 | -0.4 | 1689 | 0 |  | 3 | 100 |
| Artemether | 1 | 298 | 490.6 | 465.5 | 901.7 | 1 | 0 | 5.7 | 2.3 | -2.9 | -0.3 | 0.3 | 5729 | 0 |  | 3 | 100 |
| Artemisinin | 0 | 282 | 456.6 | 380.6 | 848.4 | 0 | 0 | 5.3 | 1.7 | -2.1 | -0.3 | 0.0 | 1886 | 1 |  | 3 | 96 |
| Artesunate | 0 | 384 | 644.1 | 465.1 | 1155.8 | 4 | 1 | 8 | 2.5 | -3.9 | -0.1 | -1.4 | 50.4 | 2 |  | 3 | 72 |
| Atovaquone | 0 | 367 | 620.6 | 136.9 | 1099.8 | 1 | 1 | 4.8 | 4.1 | -5.7 | 0.6 | -0.4 | 917.5 | 3 |  | 3 | 100 |
| Chloroquine | 1 | 294 | 594.1 | 188.9 | 982.9 | 6 | 0 | 3 | 4.6 | -4.5 | 0.4 | -0.1 | 3718.1 | 0 |  | 3 | 100 |
| Clindamycine | 0 | 425 | 721.5 | 534.2 | 1307.3 | 10 | 4 | 11.8 | 2 | -1.6 | -0.8 | -0.7 | 139.2 | 6 |  | 3 | 77 |
| Dapsone | 1 | 236 | 431.6 | 0.0 | 687.9 | 2 | 0 | 7 | -0.4 | -1.6 | -1.3 | -0.9 | 289.1 | 0 |  | 2 | 68.8 |
| Doxycycline | 4 | 422 | 602.2 | 174.0 | 1104.2 | 2 | 0 | 17.3 | -4.0\* | -0.9 | -2.9\* | -2.5 | 9.2\* | 4 |  | 1 | 20.8 |
| Halofantrine | 5 | 470 | 785.4 | 160.2 | 1351.8 | 5 | 0 | 3 | 7.6\* | -8.5\* | 1.5 | 0.2 | 2844.1 | 0 |  | 1 | 100 |
| Lumefantrine | 5 | 497 | 819.1 | 160.7 | 1437.5 | 7 | 0 | 3 | 8.3\* | -9.4\* | 1.7\* | 0.2 | 4337.2 | 0 |  | 1 | 100 |
| Mefloquine | 2 | 362 | 533.1 | 0.0 | 925.1 | 2 | 0 | 4 | 4.1 | -5.9 | 0.1 | 0.5 | 2903.1 | 0 |  | 3 | 100 |
| Pamaquine | 0 | 316 | 654.8 | 443.4 | 1148.1 | 9 | 1 | 4.8 | 4 | -3.3 | 0.4 | 0.2 | 1475.2 | 5 |  | 3 | 100 |
| Primaquine | 0 | 259 | 528.1 | 242.7 | 909.6 | 7 | 3 | 3.8 | 2 | -2.4 | -0.1 | -0.2 | 371.3 | 6 |  | 3 | 85 |
| Proguanil | 1 | 238 | 478.2 | 125.3 | 768.6 | 6 | 0 | 6 | 1.1 | -1.7 | -1.0 | -0.7 | 834.6 | 0 |  | 3 | 86 |
| Quinacrine | 0 | 370 | 680.5 | 268.8 | 1163.6 | 7 | 0 | 3.5 | 5.6 | -6.1 | 0.8 | -0.1 | 4435.7 | 1 |  | 1 | 100 |
| Quinine | 0 | 324 | 522 | 301.0 | 990.1 | 5 | 1 | 5.5 | 3.3 | -3.2 | 0.1 | 0.2 | 628.3 | 4 |  | 3 | 96 |
| Sulfadoxine | 1 | 296 | 510.6 | 152.3 | 849.5 | 5 | 0 | 9.5 | -0.8 | -1.7 | -1.7\* | -1.4 | 213.4 | 2 |  | 2 | 64 |
| Tetracycline | 5 | 422 | 604.5 | 173.1 | 1111.8 | 2 | 0 | 16 | -3.4\* | -1.4 | -2.5\* | -2.6 | 6.8\* | 5 |  | 1 | 22 |

a designed CNP analogs and known antimalarial agents, b drug likelihood, number of descriptors (24 out of a total of 49 calculated with QikProp, ver. 3.7, release 14) out of the interval for 95% of known drugs; c molecular weight in g.mol-1 (range for 95% of drugs: 130 - 725 g.mol-1) []; d total molecular surface accessible to the solvent, in Å2 (radius probe 1.4 Å) (range for 95% of drugs: 300 - 1000 Å2); e hydrophobic portion of the molecular surface accessible to the solvent, in Å2 (radius probe 1.4 Å) (range for 95% of drugs: 0 - 750 Å2); f total volume of molecule surrounded by the molecular surface accessible to the solvent, in Å3 (radius probe 1.4 Å) (range for 95% of drugs: 500 - 2000 Å3); g number of non-trivial single bonds (non-CX3), with non-prevented rotation (non-alkene, amide, ring with low number of links) (range for 95% of drugs: 0 - 15); h estimated number of potential hydrogen bonds as a donor with water molecules in aqueous solution. It is an average that can be non-integer (interval for 95% of drugs: 0.0 - 6.0); i estimated number of potential hydrogen bonds as acceptor with water molecules in aqueous solution. It is an average that can be non-integer (interval for 95% of drugs: 2.0 - 20.0); j logarithm of the partitioning coefficient between the n-octanol/water phases (range for 95% of drugs: -2 - 6.5); k logarithm of the aqueous solubility, log S. S in mol dm-3 it is the concentration of solute in a saturated solution in equilibrium with the crystal (range for 95% of drugs: -6.0 - 0.5); l logarithm of the constant of attachment to human serum albumin (interval for 95% of drugs: -1.5 - 1.5); m logarithm of the blood/brain partition coefficient. Note: QikProp forecasts for orally administered drugs for example, dopamine and serotonin are negative CNS because they are too polar to pass the blood-brain barrier (interval for 95% of drugs: -3.0 - 1.2); n predictive apparent permeability of Caco-2 cells on the Boehringer-Ingelheim scale, in [nm/s] (interval for 95% of drugs: < 25 low, > 500 important); o number of metabolic reactions (range for 95% of drugs: 1 - 8);ppredicted inhibition constants ,(pM). was predicted from computed ΔΔGcom using the regression Equation shown in Table 3;q oral absorption by humans (1 - low, 2 - medium, 3 - high); r percentage of oral absorption by humans appreciated at the gastrointestinal level (<25% - low, >80% high); (\*) the stars indicate that a descriptor is out of the range for 95% of known drugs.

AcknowledgEments

A brief acknowledgement section may be given after the conclusion section just before the references. The acknowledgments of people who provided assistance in manuscript preparation, funding for research, etc. should be listed in this section. All sources of funding should be declared as an acknowledgement. Authors should declare the role of funding agency, if any, in the study design, collection, analysis and interpretation of data; in the writing of the manuscript. If the study sponsors had no such involvement, the authors should so state.

Competing interests

The authors declare there is no conflict of interest in relation with the work presented herein.

Authors’ Contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Consent

It is not applicable.

Ethical approval

It is not applicable.

Definitions, Acronyms, Abbreviations

*2D : Two dimensional*

*3D : Three dimensional*

*ACTs : Artemisinin-based Combination Therapies*

*ADME : Absorption, distribution, metabolism, and excretion*

*ADPN : Azadipeptides Nitrile*

*CNP : Carbonitriles pyrimidine*

*ΔΔGcom : Relative complexation GFE*

*ΔΔHMM : Relative enthalpic contribution GFE*

*ΔΔGsolv : Relative solvation GFE*

*ΔΔTSvib : Relative vibrational entropy*

*E: Enzyme*

*Eint : Energy interaction*

*FP2 / FP3: falcipain 2 / falcipain 3*

*GFE: Gibbs free energy*

*HYD\_AL: hydrophobic aliphatic*

*HBA : hydrogen-bond acceptor*

*HYD : hydrophobic*

*HOA : human oral absorption*

*I : Inhibitor*

*MM : Mechanic molecular*

*MM-PB : Mechanic molecular – Poisson Boltzmann*

*PBD : Protein Data Bank*

*pfFP2 : falcipain 2 of Plasmodium falciparum*

*PH4 3D QSAR : Pharmacophore three dimensional quantitative structure activity relationship*

*QSAR : Quantitative structure activity relationship*

*RMSD:*

*SAR : Structure activity relationship*

*SBDD : Structure – based drug design*

*VL : Virtual Library*

*WHO : World Health Organization*

APPENDIX

ReferenceS

1. . Wongsrichanalai C, Meshnick SR. Declining artesunate–mefloquine efficacy against falciparum malaria on the Cambodia–Thailand border. *Emerg Infect Dis*. 2008;14(5):716–719. [*https://doi.org/10.3201/eid1405.071601*](https://doi.org/10.3201/eid1405.071601) [↑](#endnote-ref-2)
2. . [*World malaria report 2022. Geneva: World Health Organization; 2022. Licence: CC BY-NC-SA 3.0 IGO.*](https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2022) [↑](#endnote-ref-3)
3. . Hyde JE. Drug–resistant malaria–an insight. *FEBS J*. 2007;274(18):4688– 4698. [*https://doi.org/*](https://doi.org/)[*10.1111/j.1742-4658.2007.05999.x*](https://doi.org/10.1111/j.1742-4658.2007.05999.x) [↑](#endnote-ref-4)
4. . Sutherland, C. J.; Lansdell, P. L.; Sanders, M.; Muwanguzi, J.; van Schalkwyk, D. A.; Kaur, H. et al. pfk13-Independent treatment failure in four imported cases of Plasmodium falciparum malaria treated with artemether lumefantrine in the united kingdom. Antimicrob. Agents Chemother. 2017, 61, No. e02382. [*https://doi.org/10.1128/aac.02382-16*](https://doi.org/10.1128/aac.02382-16) [↑](#endnote-ref-5)
5. . Ashley, E. A.; Dhorda, M.; Fairhurst, R. M.; Amaratunga, C.; Lim, P.; Suon, S et al.. Spread of artemisinin resistance in Plasmodium falciparum malaria. N. Engl. J. Med. 2014, 371, 411-423. [*https://doi.org/10.1056/NEJMoa1314981*](https://doi.org/10.1056/NEJMoa1314981)*.* [↑](#endnote-ref-6)
6. . Wongsrichanalai C, Meshnick SR. Declining artesunate–mefloquine efficacy against falciparum malaria on the Cambodia–Thailand border. *Emerg Infect Dis*. 2008;14(5):716–719. [*https://doi.org/10.3201/eid1405.071601*](https://doi.org/10.3201/eid1405.071601) [↑](#endnote-ref-7)
7. . Dondorp AM, Nosten F, Yi P, *et al*. Artemisinin resistance in *Plasmodium falciparum* malaria. *Eng J Med*. 2009;361(15):455–467. [*https://doi.org/10.1056/NEJMoa0808859*](https://doi.org/10.1056/NEJMoa0808859) [↑](#endnote-ref-8)
8. . Garcia Linares GE, Rodriguez JB. Current Status and Progresses Made in Malaria Chemotherapy. *Curr Med Chem*. 2007;14(3): 289–314. [*https://doi.org/10.2174/092986707779941096*](https://doi.org/10.2174/092986707779941096) [↑](#endnote-ref-9)
9. . Juliane Wunderlich, Petra Rohrbach, John Pius Dalton. The malaria digestive vacuole. *Front Biosci.* 2012;4:1424–1448. [*https://doi.org/10.2741/s344*](https://doi.org/10.2741/s344) [↑](#endnote-ref-10)
10. . Rosenthal PJ. *Protease Inhibitors.* In: Antiparasitic Chemotherapy: Mechanisms of Action, Resistance, and New Directions in Drug Discovery. New York: Springer; 2001:325–345. [↑](#endnote-ref-11)
11. . Esmel A, Keita M, Megnassan E, et al. Insight into binding mode of nitrile inhibitors of *Plasmodium falciparum* Falcipain–3, QSAR and Pharmacophore models, virtual design of new analogues with favorable pharmacokinetic profiles. *J Comput Chem Molec Model*. 2017;2–1:1–21. [*http://dx.doi.org/10.25177/JCMP.2.1.5*](http://dx.doi.org/10.25177/JCMP.2.1.5) [↑](#endnote-ref-12)
12. . Francis SE, Sullivan DJ Jr, Goldberg DE (1997) Hemoglobin metabolism in the malaria parasite Plasmodium falciparum. Annu Rev Microbiol 51: 97-123. [*https://doi.org/10.1146/annurev.micro.51.1.97*](https://doi.org/10.1146/annurev.micro.51.1.97) [↑](#endnote-ref-13)
13. . Rosenthal PJ. Hydrolysis of erythrocyte proteins by proteases of malaria parasites. Curr Opin Hematol. 2002 Mar;9(2):140-5. [*https://doi.org/10.1097/00062752-200203000-00010*](https://doi.org/10.1097/00062752-200203000-00010) [↑](#endnote-ref-14)
14. . Goldberg DE, Slater AF, Cerami A, Henderson GB. Hemoglobin degradation in the malaria parasite Plasmodium falciparum: an ordered process in a unique organelle. Proc Natl Acad Sci U S A. 1990 Apr;87(8):2931-5. [*https://doi.org/10.1073/pnas.87.8.2931*](https://doi.org/10.1073/pnas.87.8.2931) [↑](#endnote-ref-15)
15. . Kerr ID, Lee JH, Pandey KC, Harrison A, Sajid M, Rosenthal PJ, Brinen LS. Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. J Med Chem. 2009 Feb 12;52(3):852-7. [*https://doi.org/10.1021%2Fjm8013663*](https://doi.org/10.1021%2Fjm8013663) [↑](#endnote-ref-16)
16. . Rosenthal PJ, McKerrow JH, Aikawa M, et al. A malarial cysteine proteinase is necessary for hemoglobin degradation by *Plasmodium falciparum*. J Clin Invest. 1988;82(5):1560–1566. [*https://doi.org/10.1172/jci113766*](https://doi.org/10.1172/jci113766) [↑](#endnote-ref-17)
17. . Gamboa de Dominguez N D, Rosenthal PJ. Cysteine proteinase inhibitors block early steps in hemoglobin degradation by cultured malaria parasites. *Blood*. 1996;87:4448–4454. [↑](#endnote-ref-18)
18. . Boris, D. B., Fidele, N.-K., Luc, C. O. O., & Eugene, M. (2016). Targeting Cysteine Proteases from Plasmodium falciparum: A General Overview, Rational Drug Design and Computational Approaches for Drug Discovery. *Current Drug Targets, 17*, 1-26. [*http://dx.doi.org/10.2174/1389450117666161221122432*](http://dx.doi.org/10.2174/1389450117666161221122432)*.* [↑](#endnote-ref-19)
19. . Bekono BD, Ntie-Kang F, Owono Owono LC, Megnassan E. Targeting Cysteine Proteases from Plasmodium falciparum: A General Overview, Rational Drug Design and Computational Approaches for Drug Discovery. Curr Drug Targets. 2018;19(5):501-526. [*https://doi.org/10.2174/1389450117666161221122432*](https://doi.org/10.2174/1389450117666161221122432) [↑](#endnote-ref-20)
20. . Cotereon JM, Catterick D, Castro J, et al. Falcipain Inhibitors: Optimization Studies of the 2–Pyrimidine–carbonitrile Lead Series. *J Med Chem*. 2010;53(16):6129–6152. [*http://dx.doi.org/10.1021/jm101228f*](http://dx.doi.org/10.1021/jm101228f) [↑](#endnote-ref-21)
21. . R. Löser et al. Interaction of papain-like cysteine proteases with dipeptide-derived nitriles. *Journal of Medicinal Chemistry*, 48(24) :7688–7707, 2005. [*https://doi.org/10.1021/jm050686b*](https://doi.org/10.1021/jm050686b) [↑](#endnote-ref-22)
22. . Reik Löser , Jiri Gut, Philip J. et al. Antimalarial activity of azadipeptides nitriles. *Bioorganic & Medicinal Chemistry Letters*. 2010; 20 (1):252–255.[*https://doi.org/10.1016/j.bmcl.2009.10.122*](https://doi.org/10.1016/j.bmcl.2009.10.122) [↑](#endnote-ref-23)
23. . Ettari R, Bova F, Zappala M, et al. Falcipain–2 inhibitors. *Medicinal Research Reviews*. 2010;30(1):136–167. [*https://doi.org/10.1002/med.20163*](https://doi.org/10.1002/med.20163) [↑](#endnote-ref-24)
24. . K. Singh et al. Pyrimidine-based antimalarials : design strategies and antiplasmodial effects. *Med. Chem. Commun.*, 7 :749–768, 2016. [*https://doi.org/10.1039/C6MD00084C*](https://doi.org/10.1039/C6MD00084C) [↑](#endnote-ref-25)
25. . Fagnidi YKH, Toi B, Megnassan E, et al. *In silico* design of *Plasmodium falciparum* cysteine protease falcipain 2 inhibitors with favorable pharmacokinetic profile. *J Anal Pharm Res.* 2018;7(3):298‒309. [*https://doi.org/10.15406/japlr.2018.07.00244*](https://doi.org/10.15406/japlr.2018.07.00244) [↑](#endnote-ref-26)
26. . [*QikProp, version 3.7, release 14, X Schrodinger, LLC, New York, NY, 2014.*](https://www.schrodinger.com/products/qikprop) [↑](#endnote-ref-27)
27. . [*Discovery Studio molecular modeling and simulation pro­gram, version 2.5, Accelrys, Inc., San Diego, CA, 2009, California*](https://drug-discovery.creative-biostructure.com/molecular-dynamics-simulation-p18?gclid=EAIaIQobChMI8ZeIhvP3gQMVhzUGAB2mpQ6PEAAYASAAEgKJhPD_BwE) [↑](#endnote-ref-28)
28. . Kouassi AF, Kone M, Keita M, et al. Computer–Aided Design of Orally Bioavailable Pyrrolidine Carboxamide Inhibitors of Enoyl–Acyl Carrier Protein Reductase of *Mycobacterium tuberculosis* with Favorable Pharmacokinetic Profiles. *Int J Mol Sci*. 2015;16(12):29744–29771. [*https://doi.org/10.3390/ijms161226196*](https://doi.org/10.3390/ijms161226196) [↑](#endnote-ref-29)
29. . Keita M, Kumar A, Dali B, et al. Quantitative structure–activity relationships and design of thymine–like inhibitors of thymidine monophosphate kinase of *Mycobacterium tuberculosis* with favourable pharmacokinetic profiles. *RSC Advances*. 2014;4(99):55853–55866. [*https://doi.org/10.1039/C4RA06917J*](https://doi.org/10.1039/C4RA06917J) [↑](#endnote-ref-30)
30. . Dugas H. Basic principles in molecular modeling, theoretical and practical aspects. 4th ed, University of Montreal Bookstore; 1996. [↑](#endnote-ref-31)
31. . Bartol J, Comba P, Melter M, Zimmer M. Conformational searching of transition metal compounds. J Comput Chem. 1999; 20(14):1549-58. [*https://doi.org/10.1002/(sici)1096-987x(19991115)20:14%3C1549::aid-jcc8%3E3.0.co;2-f*](https://doi.org/10.1002/(sici)1096-987x(19991115)20:14%3C1549::aid-jcc8%3E3.0.co;2-f) [↑](#endnote-ref-32)
32. . Gilson MK, Honig B. The inclusion of electrostatic hydration energies in molecular mechanics calculations. *J Comput Aided Mol Des*. 1991;5(1):5–20. [*https://doi.org/10.1007/bf00173467*](https://doi.org/10.1007/bf00173467) [↑](#endnote-ref-33)
33. . Rocchia W, Sridharan S, Nicholls A, et al. Rapid grid–based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *J Comput Chem*.2002;23(1):128–137. [*https://doi.org/10.1002/jcc.1161*](https://doi.org/10.1002/jcc.1161) [↑](#endnote-ref-34)
34. . [*Discovery Studio Molecular Modeling and Simulation Software. USA: San Diego; 2009.*](https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-discovery-studio/) [↑](#endnote-ref-35)
35. . [*Böttcher CJF. Theory of electric polarization. Amsterdam, The Netherlands: Elsevier; 1973.*](https://shop.elsevier.com/books/theory-of-electric-polarization/luisa/978-0-444-41019-1) [↑](#endnote-ref-36)
36. . Duffy EM, Jorgensen WL. Prediction of properties from simulations: Free energies of solvation in hexadecane, octanol, and water. *J Am Chem Soc*. 2000;122:2878–2888. [*https://doi.org/10.1021/ja993663t*](https://doi.org/10.1021/ja993663t) [↑](#endnote-ref-37)
37. . Jorgensen WL, Duffy EM. Prediction of drug solubility from Monte Carlo simulations. *Bioorg Med Chem Lett*. 2000;10(11):1155–1158. [*https://doi.org/10.1016/s0960-894x(00)00172-4*](https://doi.org/10.1016/s0960-894x(00)00172-4) [↑](#endnote-ref-38)
38. . Jorgensen WL, Duffy EM. Prediction of drug solubility from structure. *Adv Drug Deliv Rev*. 2002;54(3):355–366*.* [*https://doi.org/10.1016/s0169-409x(02)00008-x*](https://doi.org/10.1016/s0169-409x(02)00008-x) [↑](#endnote-ref-39)
39. . Frecer V, Miertuš S. Polarizable continuum model of solvation for biopolymers. Int J Quantum Chem. 1992;42(5):1449-68. [*https://doi.org/10.1002/qua.560420520*](https://doi.org/10.1002/qua.560420520) [↑](#endnote-ref-40)
40. Allangba K. N. P. G., Keita M.,Kre N’Guessan R., Megnassan E., Vladimir F., Miertus S. Virtual design of novel Plasmodium falciparum cysteine protease falcipain-2 hybrid lactone–chalcone and isatin–chalcone inhibitors probing the S2 active site pocket, Journal of Enzyme Inhibition and Medicinal Chemistry, 2019, 34:1, 547-561. [*https://doi.org/10.1080/14756366.2018.1564288*](https://doi.org/10.1080/14756366.2018.1564288) [↑](#endnote-ref-41)
41. . Krovat EM, Frühwirth KH, Langer T. Pharmacophore identification, *in silico* screening, and virtual library design for inhibitors of the human factor Xa. *J Chem Inf Model.* 2005;45(1):146-59.[*https://doi.org/10.1021/ci049778k*](https://doi.org/10.1021/ci049778k) [↑](#endnote-ref-42)
42. . Freire E. Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov Today.* 2008;13(19-20):869–874. [*https://doi.org/10.1016/j.drudis.2008.07.005*](https://doi.org/10.1016/j.drudis.2008.07.005) [↑](#endnote-ref-43)
43. . N’Guessan H, Megnassan E. *In silico* Design of Phosphonic Arginine and Hydroxamic Acid Inhibitors of *Plasmodium falciparum* M17 Leucyl Aminopeptidase with Favorable Pharmacokinetic Profile. *J Drug Design Med Chem*. 2017;3(6):98–125. [*http://dx.doi.org/10.11648/j.jddmc.20170306.13*](http://dx.doi.org/10.11648/j.jddmc.20170306.13) [↑](#endnote-ref-44)