



RESEARCH ARTICLE

STRUCTURE-BASED DESIGN OF NOVEL PYRIMIDINE CARBONITRILES ANALOGS TARGETING THE CYSTEINE PROTEASE FALCIPAIN 2 OF *PLASMODIUM FALCIPARUM* (pFP2) AT THE TROPHOZOÏTE STAGE WITH FAVORABLE ADME SPECIFICITIES

Yves Kily Hervé Fagnidi^{1,2} , Eric Ziki³ , Koffi N'Guessan Placide Gabin Allangba^{2,4,5,6,9} ,
 Beguems Toi^{7,8} , Eugene Megnassan^{2,8,9*}

¹Department of Science and Technology, University Alassane Ouattara, Ivory Coast. ²Fundamental and Applied Physics Laboratory, University Nangui Abrogoua, Ivory Coast. ³Laboratory of Material Sciences, the Environment and Solar Energy, University Felix Houphouët-Boigny, Ivory Coast. ⁴Laboratory of Environmental Science and Technology, University Jean Lorougnon Guédé, Ivory Coast. ⁵Laboratory of Biophysics and Nuclear Medicine (LBNM), University Félix Houphouët-Boigny, Ivory Coast. ⁶Department of Medical Physics, University of Trieste, Trieste, Italy. ⁷Ecole Normale Supérieure, Abidjan, Ivory Coast. ⁸Laboratory of Structural and Theoretical Organic Chemistry, University Felix Houphouët Boigny, Ivory Coast. ⁹ICTP-UNESCO, QLS, Strada Costiera 11, I 34151 Trieste, Italy.

Article Info:



Article History:

Received: 6 August 2023
 Reviewed: 9 September 2023
 Accepted: 26 October 2023
 Published: 15 November 2023

Cite this article:

Fagnidi YKH, Ziki E, Gabin Allangba KNP, Toi B, Megnassan E. Structure-based design of novel Pyrimidine carbonitriles analogs targeting the Cysteine protease Falcipain 2 of *Plasmodium falciparum* (pFP2) at the trophozoite stage with favorable ADME specificities. Universal Journal of Pharmaceutical Research 2023; 8(5):39-52. <https://doi.org/10.22270/ujpr.v8i5.1008>

*Address for Correspondence:

Eugene Megnassan, Fundamental and Applied Physics Laboratory, University Nangui Abrogoua, Ivory Coast. Tel: +225 0102363008; E-mail: megnase@gmail.com

Abstract

Aim and Objective: Structure-based drug design (SBDD) of new antimalarials at the moment of resistance of the most causative agent, *Plasmodium falciparum* to the more valuable artemisinin combination therapy (ACT) is even more urgent. Carbonitriles pyrimidine derivatives (CNP) has emerged as potential inhibitors of the cysteine protease falcipain 2 of *Plasmodium falciparum* (pFP2), so here we report virtual pharmacophore based screening of the CNP chemical subspace yielding novel CNP analogs with predicted high inhibitory potency against pFP2.

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: $\Delta\Delta G_{com}$) of pFP2-CNP complex formation and the experimental half-maximal enzymatic inhibition concentration (IC_{50}^{exp}). 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 (IC_{50}^{pre}) and pharmacokinetic profile.

Results: The QSAR model linear correlation equation: $pIC_{50}^{exp} = -0.1025 \times \Delta\Delta G_{com} + 7.2867$, $R^2=0.94$, the subsequent PH4 model linear correlation between experiment and PH4-estimated IC_{50} : $pIC_{50}^{exp} = 0.9366 \times pIC_{50}^{pre} + 0.2849$, $R^2=0.91$ 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 (IC_{50}^{pre}) 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 identifying novel CNP analogs against pFP2 and the handful of top predicted analogs are worth undergoing synthesis and biological evaluation.

Keywords: ADMET, CNP, FP2, Malaria, pharmacophore, QSAR, virtual screening.

INTRODUCTION

The malaria disease mainly occurs in tropical regions with a unicellular eukaryotic parasites of the genus

plasmodium as causative agent, the most virulent of which is *Plasmodium falciparum*¹. The WHO 2022 malaria reports deplore 247 million clinical cases for 619,000 deaths mostly under 5 years. Almost 80% of

malaria-related deaths in the WHO Africa region in 2021². The spread of resistance to the majority of existing drugs, even including chloroquine, which was one of the pillars of antimalarial treatment, constitute an obstacle³.

The same is also true for the case of artemisinin, our latest line of defense and artemisinin-based combination therapies (ACTs) which are also experiencing the emergence of new resistance hindering current efforts to counter malaria^{4,5,6,7,8}. 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 available⁹, there is an urgent need for a new therapy against malaria through the identification and development of new classes of drugs with high antimalarial potential¹⁰. 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 compound¹¹. The digestion of 3/4 of the hemoglobin present in the infected erythrocytes, provides the parasite with amino acids which are necessary for its growth and survival^{12,13,14}. The inhibition of parasitic proteases, in particular the cysteine proteases involved in the process of hemoglobin degradation, proves to be central to avoiding parasitic proliferation^{3,15}. The cysteine protease falcipain-2 (FP2) arouses great interest because of the central role it plays in the life cycle of *P. falciparum* through the degradation of hemoglobin¹⁶. FP2, the most expressed and best studied enzyme among falcipain, is a promising target for the development of new antimalarial drugs¹⁷. It is a logical target for antiparasitic chemotherapy and therefore we have been interested in the development of its inhibitors as antiparasitic¹⁸. FP2 has a fairly considerable active site comprising four pockets S1, S'1, S2 and S3 which can accommodate, for each of them, substituent's of substrates P1, P'1, P2 and P3 respectively. Previous studies have shown the selectivity of these pockets; the S1 pocket having a high affinity with compounds bearing a nitrile group at P1¹⁹ and the S2 pocket having a marked preference for compounds having a hydrophobic group at P2²⁰. Various non-peptide heteroaryl nitriles compounds have been studied as potential antiparasitic inhibitors of falcipain. This is the chemical class of 5-substituted-2-cyanopyrimidine known as carbonitriles pyrimidine (CNP) which constitutes a powerful and promising head of series¹⁸. Their P1 lateral chain containing the pyrimidine nucleus presents various substitution possibilities which is a very useful advantage in the fight against the drug resistance of certain pathogens and constitutes a reliable means in the design of antimalarial drugs with powerful bioavailable and favorable oral pharmacokinetic properties. 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 FP2²¹, by *in situ* modification to explore the active site of FP2 with CNPs. The study with the ADPNs provided valuable structural information that guided the simulations carried out with the CNPs and led to the design based on the structure of new and more powerful antimalarial agents. In this work, we start from a set of CNP molecules to build a QSAR model of FP2 inhibition with a single descriptor (Gibbs free energy, GFE during the formation of the FP2-Inhibitor complex) which has been correlated with the experimental biological activity IC_{50}^{exp} . Then a 3D-QSAR pharmacophore protocol was used to build a four-characteristic pharmacophore model (PH4) based on the conformations linked from the CNPs to the active site of FP2. In addition, the calculated enzyme-ligand interaction energy E_{int} correlates well with the experimental activities IC_{50}^{exp} making it possible to achieve its distribution for each residue of the active site. This latest structural information allowed us to select appropriate fragments P1 and P2 as a construction base for a virtual library (VL) of FP2 inhibitor. In order to access a good pharmacokinetic profile while avoiding toxicity problems, VL was targeted in priority on compounds with 0 property descriptors lying outside the range of values determined for 95% of the known drugs on 24 selected descriptors calculated by the QikProp. The aim of this study was to use the predictability of QSAR models obtained from binding the inhibitory enzyme to the pharmaceutical PH4 3D-QSAR model for VL screening. The best Hit Fits from the virtual PH4-based assay for VL were evaluated *in silico* MM-PB for predicted inhibitory activity up to the picomolar range for the most potent analogues.

MATERIALS AND METHODS

Training set and validation set

The series of compounds used in this work is taken from the literature and belongs to the class of 5-substituted-2-cyanopyrimidine known as carbonitriles pyrimidine (CNP)¹⁸. Their activities cover a sufficiently wide range of activity to allow the construction of a reliable QSAR model. The training set containing 12 CNP ligands and the validation set 3 CNP ligands taken from the reference¹⁸.

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* modification 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 program²². No water molecules from the crystallographic structure were kept in the model. To identify the lowest-energy conformation bound of the modified inhibitor, an exhaustive conformation search of all freshly created residue bonds, coupled with minimization of the inhibitor energy of the protein's active site, was necessary. The structure of the

resulting low-energy complex is carefully optimized by minimizing the overall complex. In practice, the *in situ* 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. Calculation of the relative binding affinity of the ($\Delta\Delta G_{\text{com}}$) ligand has been fully described and reported previously.

$$\Delta\Delta G_{\text{com}} = \Delta G_{\text{com}}(\text{I}) - \Delta G_{\text{com}}(\text{I}_{\text{ref}}) = \Delta\Delta H_{\text{MM}} - \Delta\Delta TS_{\text{vib}} + \Delta\Delta G_{\text{sol}} \quad (1)$$

The enthalpic contribution relative to the GFE change related to intermolecular interactions in the E:I complex is described by $\Delta\Delta H_{\text{MM}}$ and derived from molecular mechanics (MM), $\Delta\Delta G_{\text{sol}}$ and $\Delta\Delta TS_{\text{vib}}$ represent, respectively, the relative solvation GFE and the simplified relative vibrational entropy.

Molecular mechanics

The modeling of CNP ligands and their FP2-CNP complexes was carried out by molecular mechanics using the CFF force field as widely described previously²⁴.

Conformational search

Conformational research is a method for calculating the relative energy associated with the conformation of a molecule²⁵. Its aim is therefore to find the minimum possible and to calculate Boltzmann's population, which gives us information on the population of occupied levels at a given temperature²⁶. This method has been described earlier²⁵.

Solvation Gibbs free energy

The biological medium is aqueous and proteins and ligands interaction mechanisms involved in the binding process take account of solvation phenomenon. Discovery Studio's DelPhi module calculates the electrostatic component of the solvation GFE, which includes the effect of the ionic strength by solving the non-linear Poisson-Boltzmann (PB) equation^{27,28,29}. This method has been described fully earlier²⁵.

Interaction energy

The non-bonded interactions (van der Waals and electrostatic interatomic potential terms) between two sets of atoms in all E:I complexes were calculated via the MM interaction energy (E_{int}) calculation protocol available in Discovery Studio 2.5³⁴ as described earlier²⁵.

Pharmacophore generation

The Discovery Studio's Catalyst Hypogen algorithm program³⁰ allowed us, based on the models of the various EI complexes used, to generate the hypotheses to construct a 3D-QSAR pharmacophore as described previously²⁵.

ADMET-related properties

The Qikprop program³¹ based on the Jorgensen method^{32,33,34} as fully described previously²⁵ was used to calculate the pharmacokinetic profile of CNPs.

Virtual library generation

The virtual library for CNPs analogues was generated according to the protocol described to the reference²⁵.

ADMET based library focusing

The criterion selection drug-likeness used to focus the initial virtual library of CNPs analogs was fully presented earlier²⁶.

Pharmacophore based library searching

PH4 based library screen process was described earlier²⁵.

In silico screening

The molecular structures corresponding to the best mapping of the pharmacophore model of 3D inhibition-QSAR PH4 Hypo1 have been selected and subjected to a screening of the QSAR model of complexation. The relative free enthalpy $\Delta\Delta G_{\text{com}}$ upon the formation of E:I complex was calculated for each new ligand selected from the focused virtual library, then used for the prediction of the predicted inhibitory potencies ($\text{IC}_{50}^{\text{pre}}$) against FP2. This parameter was inserted into the target-specific scoring function of the pFP2 receptor

$$\text{pIC}_{50}^{\text{pre}} = a \cdot \Delta\Delta G_{\text{com}} + b \quad (2)$$

RESULTS

Training and validation set

The series of CNP compounds comprising 12 ligands for the training set and 3 ligands for the validation set is a homogeneous series of pFP2 inhibitors with known and determined inhibitory activities from the same laboratory¹⁸. A variation of two positions R₁ and R₂ on the pyrimidine carbonitriles backbone was made to obtain the whole series (Table 1). Their experimental biological activities ($1 \leq \text{IC}_{50}^{\text{exp}} \leq 609 \text{ nM}$)¹⁸ from the literature extend sufficiently over a wide concentration range for the construction 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 literature³⁵.

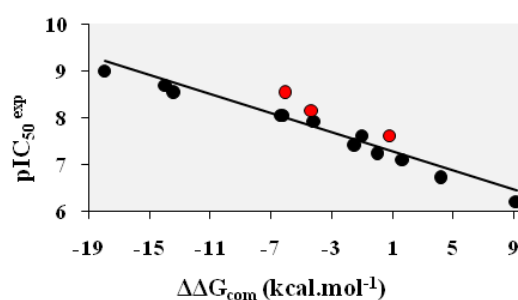


Figure 1: Plot of correlation equation between $\text{pIC}_{50}^{\text{exp}}$ and relative complexation GFE $\Delta\Delta G_{\text{com}}$ of the training set, all in kcal.mol⁻¹.

The validation set data points are shown in red color.

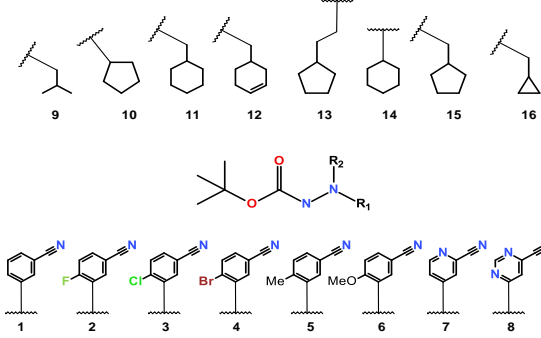
QSAR model

The relative GFE which describes the mutual affinity between the protein and the inhibitor is presented in Table 2 with its various components. The method of calculating its different quantities is explained extensively in Section 2-2. Calculated with an approximate approach, the relevance of the binding model was evaluated by a linear regression analysis, equation (2) which led to a linear correlation with the

experimental activity data IC_{50}^{exp} 18. Table 3 presents the correlation equation obtained for GFE $\Delta\Delta G_{com}$, equation (3) with relevant statistical data. The regression coefficient $R^2=97\%$ and the Fischer F

test=363.4 of the correlation with relatively high values show a strong relationship between the binding model and the experimental inhibition power IC_{50}^{exp} of the CNP series as indicated in Figure 2.

Table 1: Training set (CNP1-12) and validation set (CNPV 1-3) of P/FP2 inhibitors²⁰ used in the preparation of QSAR models of inhibitor binding.



Training set						
	CNP1	CNP2	CNP3	CNP4	CNP5	CNP6
#R ₁ -#R ₂	1-9	4-10	5-9	6-9	4-11	4-12
IC ₅₀ ^{exp} (nM)	60	9	40	79	24	2
Validation set						
	CNPV1	CNPV2	CNPV3			
#R ₁ -#R ₂	2-9	4-9	4-16			
IC ₅₀ ^{exp} (nM)	24	3	7			

In the first part of the table are numbered the different groups R₁ (bottom) and R₂ (top) in the form #R ≡ group index

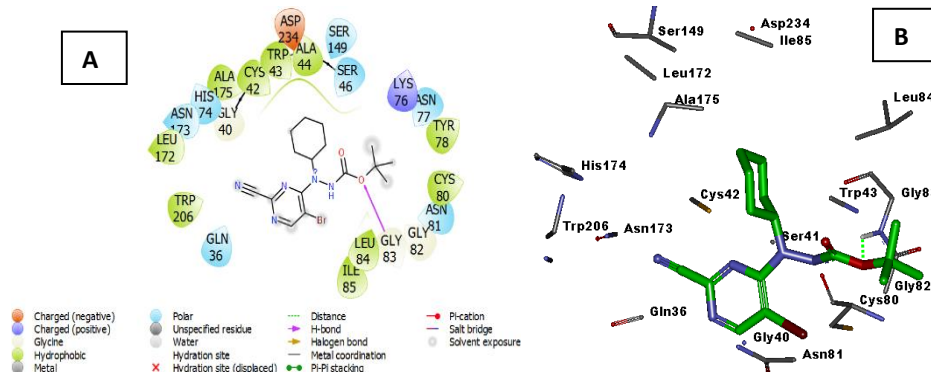


Figure 2: A: 2D schematic interaction diagram of the most potent inhibitor CNP9 (Table 1) at pFFP2 active-site; B: 3D structure of the active-site with bound inhibitor CNP9.

Binding mode and interaction energy

Binding mode

The active conformation of the most active of the carbonitriles pyrimidine 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

The distribution of the inhibitory enzymatic intermolecular E_{int} interaction energy of the key residues of the different pockets of the active site was subsequently calculated and listed in Table 4. Similarly, their correlation was determined as a function of the experimental activity IC_{50}^{exp} and represented in Figure 3. A comparative individual contribution between the most active CNP9 and the least active CNP11 compounds confirming the observed trend of experimental activities by pockets is presented in Table 5.

Table 2: Complexation GFE (binding affinity) and its different components respectively for the training and validation set of inhibitors (CNP 1-12 and CNPV 1-3).

Training set ^a	M_w ^b	$\Delta\Delta H_{MM}$ ^c	$\Delta\Delta G_{sol}$ ^d	$\Delta\Delta TS_{vib}$ ^e	$\Delta\Delta G_{com}$ ^f	pIC_{50}^{exp} ^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	$\Delta\Delta H_{MM}$	$\Delta\Delta G_{sol}$	$\Delta\Delta TS_{vib}$	$\Delta\Delta G_{com}$	$pIC_{50}^{pre} / pIC_{50}^{exp}$
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; ^b M_w is the molecular mass of the inhibitor; ^c $\Delta\Delta H_{MM}$ (kcal.mol⁻¹) is the relative enthalpic contribution to the GFE, CNP1 is the reference inhibitor; ^d $\Delta\Delta G_{sol}$ (kcal.mol⁻¹) is the relative solvation GFE contribution to the GFE; ^e $-\Delta\Delta TS_{vib}$ (kcal.mol⁻¹) is the relative inhibitor entropic contribution to the GFE; ^f $\Delta\Delta G_{com}$ (kcal.mol⁻¹) is the relative GFE related to the enzyme-inhibitor complex formation: $\Delta\Delta G_{com} \cong \Delta\Delta H_{MM} + \Delta\Delta G_{sol} - \Delta\Delta TS_{vib}$. ^g IC_{50}^{exp} is the experimental inhibition constant obtained from¹⁸.

From the analysis of Table 4, as it has been noted for the azadipeptides nitrile (ADPN) in previous study²⁰, here in the case of pyrimidine carbonitriles (CNP) a relevant stabilizing drop of S2 pocket residues contributions to Enzyme – Inhibitor interaction energy (E_{int}) from the least active TS CNP11 to the best active one CNP9 by almost 4 kcal.mol⁻¹.i.e. 50%, close to the 45% increase of their observed inhibitory potency pIC_{50}^{exp} from 6.21 to 9 respectively (see Table 5 and Figure 3 Top). Therefore the preeminent role of the hydrophobic S2 pocket of *PfFP2* active site in the design of potent molecules against *PfFP2* is confirmed

again as previously reported³⁶ devoting a central role of Leu172 in that S2 pocket (Figure 3, and Figure 4).

QSAR Pharmacophore Model

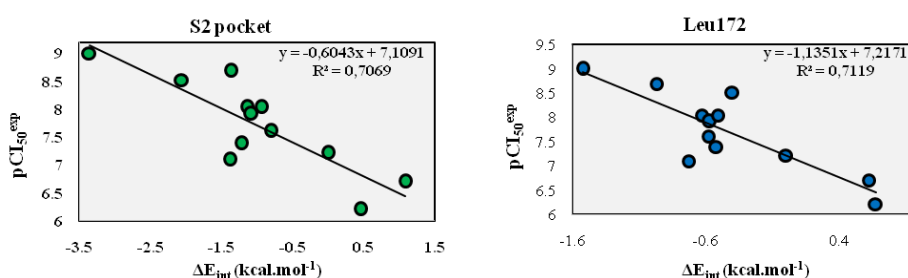
The QSAR 3D Pharmacophore model generation extensively was detailed previously²¹. The model was generated from the active conformation of CNPA compounds in the FP2 active site. These compounds cover a wide range of experimental activity (1 – 609 nM). The results of the 10 best hypotheses are presented in Table 6 showing the cost, the RMSD and the correlation coefficient between predicted and experimental activities.

Table 3: Regression analysis of computed binding affinities $\Delta\Delta G_{com}$ and their observed activities

$$pIC_{50}^{exp} = -\log_{10} IC_{50}^{exp} \text{ of CNPs}$$

Statistical Data of Linear Regression

$pIC_{50}^{exp} = -0.1025 \times \Delta\Delta G_{com} + 7.2867$ (3)	$\Delta\Delta G_{com}$
Number of compound n	12
Square correlation coefficient regression R^2	0.97
LOO cross-validated Square correlation R^2_{xv}	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 IC_{50}^{exp} [nM]	1 – 609

**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 *pfFP2*-CNPx.**

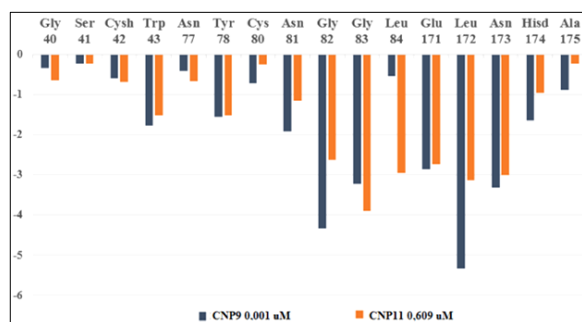


Figure 4: Molecular mechanics intermolecular interaction energy E_{int} breakdown (in kcal.mol⁻¹) to residue contributions for CNP9 and CNP11.

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 supreme indicator namely the gap $\Delta=280.644$ between the fixed (32.305) and the null costs (312.949) $\gg 70$ documents the quality and predictive character of the PH4 indicating the probability that the correlation between the IC_{50} values estimated by the PH4 and experimental ones is real at more than 90%³⁴. 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 (R^2) between 0.94 and 0.91 (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.

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 pIC_{50}^{exp} as a function of pIC_{50}^{pre} , estimated by Hypo1: $pIC_{50}^{exp}=0.9366 \times pIC_{50}^{pre} + 0.2849$ are listed in Table 7 ($n=12$; $R^2=0.91$; $R^2_{xv}=0.91$; $F\text{-test}=114.34$; $\sigma=0.551$; $\alpha > 95\%$) 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 QSAR 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.

Virtual screening

In silico screening of a virtual (combinatorial) library can lead to hit identification as it was shown in previous works on inhibitors design³⁷. From the different groups listed in Table 8, we have created a virtual combinatorial library by substitution in positions R_1 , R_2 , R_3 and R_4 on the pyrimidine nucleus and its side chain, the size of which is $R_1 \times R_2 \times R_3 \times R_4=14 \times 17 \times 14 \times 25=83300$ CNPA analogs (CNPs analogs). This virtual combinatorial library has been focused to 32856 (39%) orally administrable compounds with very good predictive pharmacokinetic profile.

Table 4: Active site residue contribution to E_{int} (in kcal/mol).

Pockets	CNP	1	2	3	4	5	6	7	8	9	10	11	12
	pIC_{50}^{exp}	7.2	8.1	7.4	7.1	7.6	8.7	7.9	8.1	9	8.5	6.2	6.7
S1	Gn136	-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	-1.4	-1.5	-1.8	-2.2	-1.5
	Cys80	-0.3	-0.3	-0.3	-0.3	-0.4	-0.4	-0.4	-0.3	-0.3	-0.1	-0.1	-0.2
	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
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	-0.3	-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

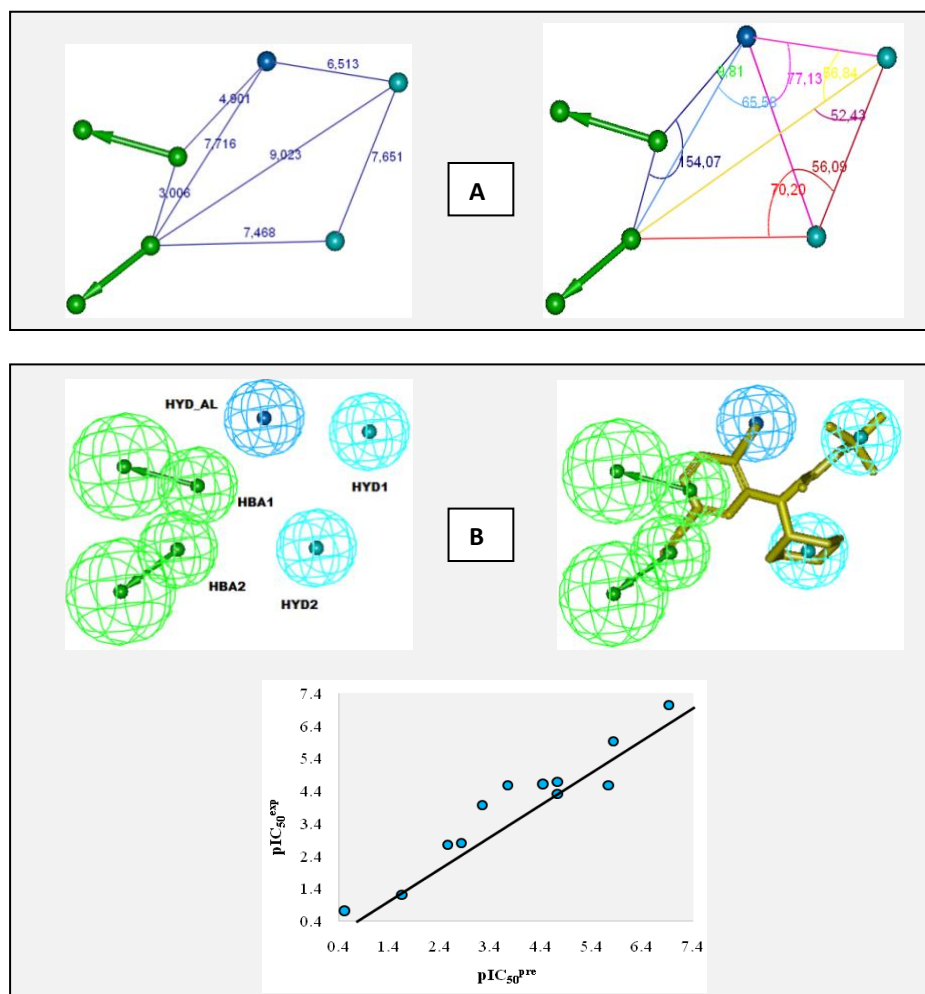


Figure 5: A: Coordinates, B: features and mapping of the FP2 inhibitor pharmacophore with the best training set CNP9 (yellow). The correlation plot of experimental vs. predicted inhibitory activity is displayed at the bottom.

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.

It was then screened with the generated PH4 3D-QSAR model of Hypo1 to yield 127 analogues mapping to the PH4 and all were subjected to a last evaluation of their predicted IC_{50}^{pre} with the complexation method (IC_{50}^{pre} is calculated from the correlation equation (3) Table 3). Finally, 52 analogues with better scores are selected and listed in Table 8. The histograms of frequency of appearance of groups R_1 , R_2 , R_3 and R_4 for the 52

CNPA (Figure 6) reveal the fragments of greater occurrences (value in parentheses): 1(49), 2(1), 3(1), 4(1) in R_1 for filling the S3 pocket; 15(32), 31(20) in R_2 and 38(14), 39(10), 40(12), 41(9) at R_3 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 R_4 for filling the pocket S2.

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

Hypothesis	RMSD ^a	R ² ^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

Table 7: Regression analysis of pIC_{50}^{exp} and computed pIC_{50}^{pre} of CNPs toward PfFP2.

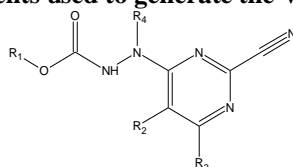
Statistical Data of Linear Regression for Hypo 1	
$pIC_{50}^{exp} = 0.9366 \times pIC_{50}^{pre} + 0.2849$	
Number of compound, n	12
Square correlation coefficient regression, R^2	0.91
LOO cross-validated Square correlation, R^2_{xv}	0.91
Standard error of regression, σ	0.551
Statistical significance of regression, Fisher F-test	114.34
Level of statistical significance, α	> 95 %
Range of activities IC_{50}^{exp} [nM]	1- 609

Novel CNPs analogs

The predictive activities IC_{50}^{pre} 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 ($IC_{50}^{exp} = 1$ nM) proposed by Cotereon *et al.*,²⁰ are presented in Table 9.

The best CNP analogs (CNPA) with their predicted activity value (IC_{50}^{pre}) 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 (IC_{50}^{pre}) of approximately 70 times greater than that of the best CNP9 inhibitor of the training set ($IC_{50}^{exp} = 1$ nM).

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

*R ₁	R ₂	R ₃	R ₄
1 -C ₆ H ₁₁	15 -CH ₃	32 -C ₃ H ₇ O	
2 -C ₆ H ₁₀ OH	16 -OH	33 -CONH ₂	
3 -Me(C ₆ H ₁₁)	17 -CH ₂ F	34 -CH ₂ CO ₂	
4 1,2-diMe(C ₆ H ₁₁)	18 -CH ₂ Cl	35 -C ₂ H ₅ CONH ₂	
5 -Phe	19 -NO	36 -OC ₃ H ₇	
6 -Me(phe)	20 -NO ₂	37 -OC ₂ H ₅	
7 Piperidine	21 -CH ₂ I	38 -C ₄ H ₉	
8 2-Me(piperidine)	22 -CH ₂ Br	39 -C ₄ H ₉ OH	
9 3-Me(piperidine)	23 -F	40 -C ₃ H ₁₁	
10 4-Me(piperidine)	24 -Cl	41 -C ₃ H ₁₁ OH	
11 Pyridine	25 -Br	42 -C ₃ H ₁₁ F	
12 2-Me(pyridine)	26 -I	43 -C ₃ H ₁₁ Cl	
13 3-Me(pyridine)	27 -CCl ₃	44 -C ₃ H ₁₁ Br	
14 4-Me(pyridine)	28 -CF ₃	45 -C ₃ H ₁₁ I	
	29 -CBr ₃		
	30 -Cl ₃		
	31 -H		
R ₄			
46	2,4-diMe(C ₆ H ₁₂)carboxamide		
47	2,3-diMe(C ₆ H ₁₂)carboxamide		
48	4-Et-2-Me(C ₆ H ₁₂)carboxamide		
49	But(C ₆ H ₁₂)		
50	1-But-2-Me(C ₆ H ₁₂)		
51	1-But-3-Me(C ₆ H ₁₂)		
52	1-But-4-Me(C ₆ H ₁₂)		
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

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): (c) molar mass: $300 \leq M_w \leq 500$

$g \cdot mol^{-1}$; (d) total solvent-accessible molecular surface, (probe radius 1.4 Å): $300 \leq S_{mol} \leq 1000 \text{ Å}^2$; (e) hydrophobic portion of the solvent-accessible molecular surface, (probe radius 1.4 Å): $0 \leq S_{molhfo} \leq 750 \text{ Å}^2$; (f) total volume of molecule enclosed by solvent-accessible molecular surface (probe radius 1.4

Å): $500 \leq V_{\text{mol}} \leq 2000 \text{ Å}^3$; (g) number of non-trivial (not CX3), non-hindered (not alkene, amide, small ring) rotatable bonds: $0 \leq \# \text{rotB} \leq 15$; (h) estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution, values are averages taken over a number of configurations, so they can assume non-integer values: $0.0 \leq \text{HB}_{\text{don}} \leq 6.0$; (i) estimated number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution, values are averages taken over a number of configurations, so they can assume non-integer values: $2.0 \leq \text{HB}_{\text{acc}} \leq 20.0$; (j) logarithm of partitioning coefficient between n-octanol and water phases: $-2 \leq \text{LogP}_{\text{o/w}} \leq 6.5$; (k) logarithm of predicted aqueous solubility: $\log S, S$ in $[\text{mol} \cdot \text{dm}^{-3}]$ is the concentration of the solute in a saturated solution that is in equilibrium with the crystalline solid: $-6.0 \leq \text{LogS}_{\text{wat}} \leq 0.5$; (l) logarithm of predicted binding constant to human serum albumin: $-1.5 \leq \text{LogK}_{\text{hsa}} \leq 1.5$; (m) logarithm of predicted

brain/blood partition coefficient: $-3.0 \leq \text{Log B/B} \leq 1.2$; (n) predicted apparent Caco-2 cell membrane permeability in Boehringer-Ingelheim scale in $[\text{nm s}^{-1}]$: $\text{BIP}_{\text{caco}} < 25$ poor, $\text{BIP}_{\text{caco}} > 500$ $\text{nm} \cdot \text{s}^{-1}$ great; (o) number of likely metabolic reactions: $1 \leq \# \text{metab} \leq 8$; (p) predicted inhibition constants $\text{IC}_{50}^{\text{pre}}$. $\text{IC}_{50}^{\text{pre}}$ in pM was predicted from computed $\Delta\Delta G_{\text{com}}$ using the regression Equation B shown in Table 3; (q) HOA: human oral absorption: 1=low, 2=medium, 3=high; (r)%HOA: percentage of human oral absorption in gastrointestinal tract: $\geq 80\%$ =high. The value of the drug likelihood descriptor #stars (Table 10, column 2) indicates how through 24 descriptors that comply with those of 95% of drugs the CNPs analogs perform to have a better profile than the majority of ACTs. These results also present the percentage of absorption by oral route (%HOA) in response to WHO recommendation for new antimalarials: %HOA $\geq 80\%$.

Table 9: Complexation Gibbs free energy and its components and predictive activities $\text{IC}_{50}^{\text{pre}}$ of the best new analogs.

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

^a M_w (g/mol) is molecular mass of the inhibitor; ^b $\Delta\Delta H_{\text{MM}}$ (kcal/mol) is the relative enthalpic contribution to the Gibbs free energy change related to the FP2-CNP complex formation $\Delta\Delta G_{\text{com}}$ (for details see footnote of Table 2); ^c $\Delta\Delta G_{\text{sol}}$ (kcal/mol) is the relative solvation Gibbs free energy contribution to $\Delta\Delta G_{\text{com}}$; ^d $\Delta\Delta T S_{\text{vib}}$ (kcal/mol) is the relative entropic (vibrational) contribution to $\Delta\Delta G_{\text{com}}$; ^e $\Delta\Delta G_{\text{com}}$ (kcal/mol) is the relative Gibbs free energy change related to the enzyme-inhibitor FP2-CNP complex formation $\Delta\Delta G_{\text{com}} - \Delta\Delta H_{\text{MM}} + \Delta\Delta G_{\text{sol}} + \Delta\Delta T S_{\text{vib}}$; ^f $\text{IC}_{50}^{\text{pre}}$ (pM) is the predicted half-maximal inhibitory concentration of CNP_x towards PjFP2 calculated from $\Delta\Delta G_{\text{com}}$ using correlation Equation (B), Table 3; ^g $\text{IC}_{50}^{\text{exp}}$ is given for the reference inhibitor CNP9 instead of $\text{IC}_{50}^{\text{pre}}$ (pM)²⁰.

DISCUSSION

The study of the binding mode of PjFP2:CNP from its QSAR 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 PjFP2 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.

QSAR model

The robustness of this QSAR model with a descriptor is evaluated through the components of the GFE ($\Delta\Delta G_{\text{com}}$), namely the contribution of the $\Delta\Delta H_{\text{MM}}$ enthalpy, the $\Delta\Delta G_{\text{solv}}$ solvation and the loss of

vibrational entropy $\Delta\Delta TS_{\text{vib}}$ 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 TS_{vib} to explain approximately 97% of the variations of the IC_{50}^{exp} by that of the GFE.

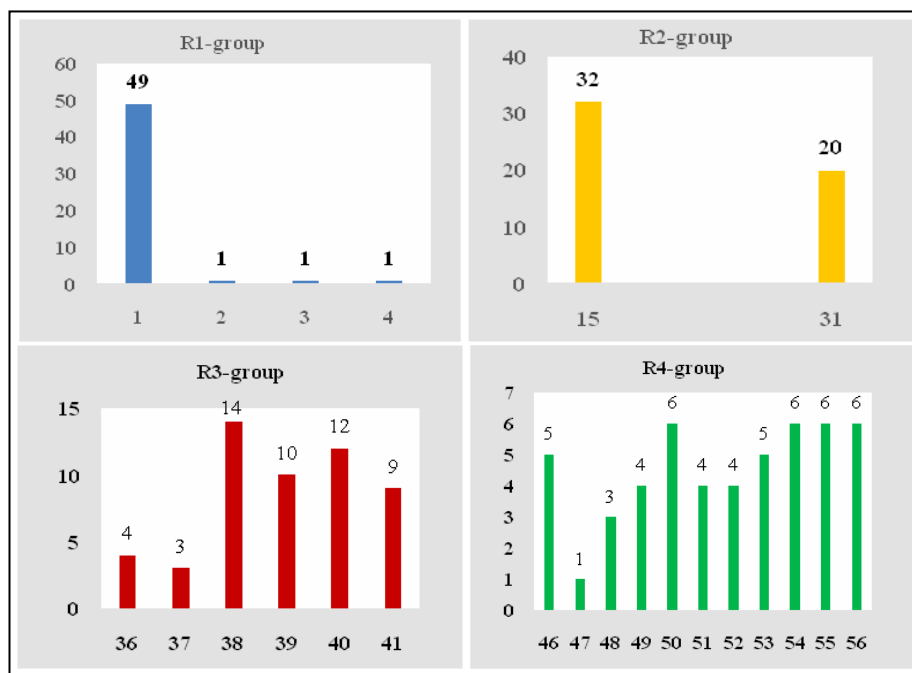


Figure 6: Histograms of frequency of occurrence of the individual R-groups in the 52 best analogues selected and mapped to the four characteristics of the ph4 hypothesis of the hypo1 pharmacophore.

This last contribution is one of the most reliable indicators of the predictive power of the QSAR model as reported by Freire *et al.*,³⁸. Consequently, the correlation equation (3) and the calculated quantities $\Delta\Delta G_{\text{com}}$ can be used respectively to predict the IC_{50}^{pre} 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 pIC_{50}^{pre} and the values of the experimental activities pIC_{50}^{exp} and documents the considerable predictive power of the QSAR model of complexation ($pIC_{50}^{\text{pre}}/pIC_{50}^{\text{exp}}$, pIC_{50}^{pre} 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: CNP_x 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 P1⁴⁸ offering a multiple substitution

positions. The halogenation of the C5 favors strong interactions with the Trp43, Asn173 and His174 residues. The hydrophobic fragments P2 fills in the S2 pocket better in contact with Ser149, Leu172, Ala175 and Asp234 increasing affinity as reported by Löser *et al.*,²⁰. The, pyrimidine carbonitriles exhibit the same similarity of interaction with the active site of FP2 as ADPNs²⁰. For the design of new analogs, the interaction energy (E_{int}) between each residue of the active site (Table 4) were calculated in order to guide the search of substituent at R1 to R4 positions as structural characteristics of the binding affinity improvement by filling suitably the S1, S'1, S2 and S3 pockets. In this way, the key residues of these pockets contribute to the overall FP2–CNP_x interaction energy. Specifically, the S2 pocket filling and Leu172 interactions, impact FP2 inhibition by more than 70% (Figure 3) as reported previously^{19,20}. In contrast, Figure 4 comparing each active site residue contribution to the E_{int} for the best active TS CNP9 and the least active TS CNP11 confirms the observed trend but cannot justify the large gap in their IC_{50}^{exp} as in our recent work justifying successfully the observed 37.5% jump in the $pFA-M17K_i^{\text{exp}}$ between methylphosphonic arginine and the hydroxamic acid from the enzyme active site residues contribution to E_{int} at a level of

35%³⁹. Therefore, the essential structural information in the design of new powerful analogues of CNP will not only be derived from the E_{int} but also from a more predictive approach. 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)³⁹.

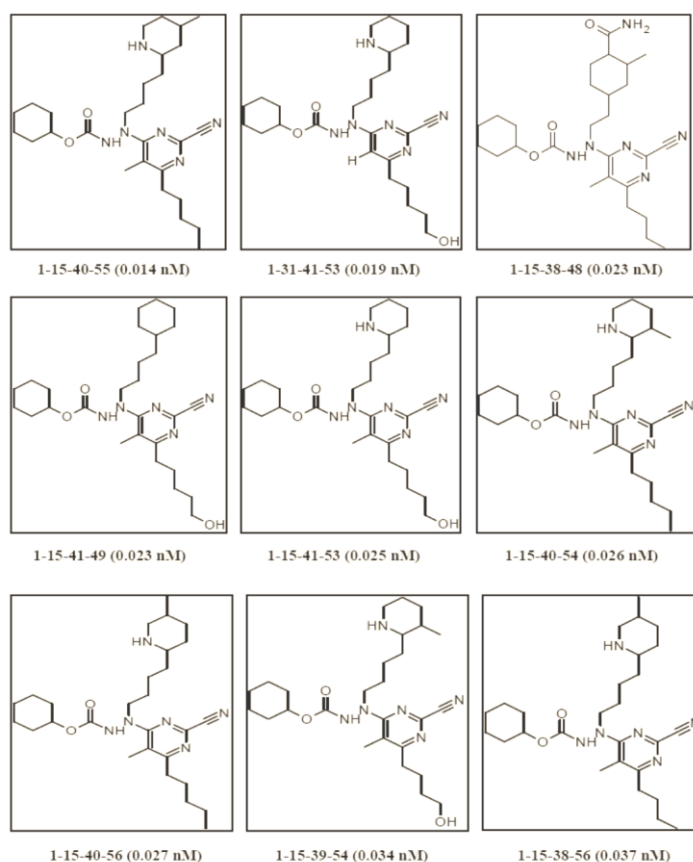


Figure 7: Presentation of some designed analogs of CNPA.

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: $IC_{50}^{pre} = 14$ pM for 1-15-40-55, 70-fold the best TS CNP9 ($IC_{50}^{exp} = 1$ nM). The representation of 1-15-40-55 in Figure 8 shows the cyclohexane ($R_1=1$) replacing in P3 the methylpropane of TS for better S3 pocket filling. In P1, two substitutions have been made at positions 4 and 5: the bromine (Br) and hydrogen (H) atoms were replaced by methyl ($R_2 = 15$) and pentyl

($R_3=40$) respectively to substantially increase the hydrophobic contact at S1. The lipophilic S2 pocket contains a larger 2-But-4-Me(piperidine) ($R_4=55$) in place of the cyclohexane fragment, for better hydrophobic contact beyond Structure Activity Relationship (SAR) results of carbonitriles pyrimidine inhibitors of FP2 and FP3²⁰ and similar to the strong potency increase in our previous study on nitrile dipeptides inhibitors of FP3¹¹.

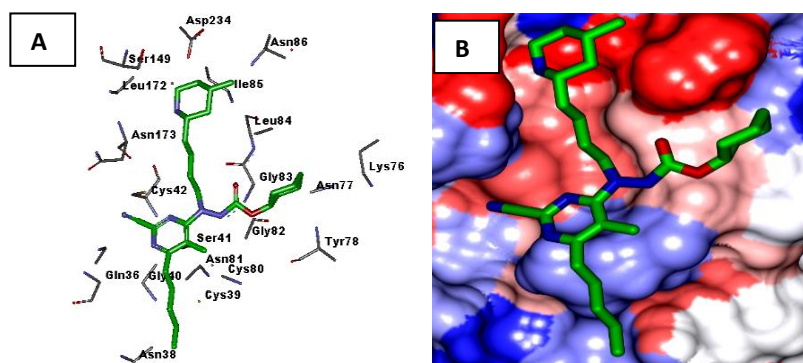


Figure 8: A: Close up of virtual hit 1-15-40-55 at the active-site of FP2.

B: Connolly surface of the active-site of PjFP2 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.

Table 10: Pharmacokinetic profile of some of the best CNPs analogues designed computed by QikProp³¹.

CNPs ^a	#stars ^b	M _w ^c	S _{mol} ^d	S _{mol,hfo} ^e	V _{mol} ^f	RotB ^g	HB _{don} ^h	HB _{acc} ⁱ	logP _{ow} ^j	logS _{wat} ^k	logK _{HSA} ^l	logB/B ^m	BIPcaco ⁿ	#meta ^o	IC ₅₀ ^{pre} ^p	HOA ^q	%HOA ^r
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	131.9	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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, release14³¹) out of the interval for 95% of known drugs; (*) the stars indicate that a descriptor is out of the range for 95% of known drugs.

ADMET-related properties

The properties related to ADMET described in section 2-8 and the results in Table 10 indicate that the CNP analogues designed possess a good level of drug likelihood since the descriptor #stars (column 2) which lies in the validation interval between 0 and 5, just like the available antimalarial drug. Moreover all the analogues' human oral absorption in the gastrointestinal tract (HOA) is strong because none of them falls below the range of good oral bioavailability admitted as the WHO main requirements for new antimalarial drugs.

Further improvement to this work

The predictive results of this study are expected to undergo synthesis and biological evaluation. Since this last step is expensive an intermediary step would be a crosscheck of the stability of the FP2 – CNPAx conformation and the related MM-PB interaction patterns and metrics by time-consuming but relevant Molecular Dynamics (MD) runs. We're operating to fulfill this achievement in shirt future.

CONCLUSIONS

The evaluation of key structural information on FP2 inhibition from FP2 – CNPAx 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, through substitutions at the P1, P2 and P3 positions, established a QSAR model explaining the variation experimental activities IC₅₀^{exp} by that of

complexation GFE($\Delta\Delta G_{com}$) of calculated during the FP2 – CNPAx complex formation, shedding light on the main determinant of the activity. A subsequent 3D-QSAR pharmacophore (PH4) helped in screening a virtual combinatorial library 83300 CNPA analogues with the purpose to substantially increase the hydrophobic contact at in the S2 pocket and, at a lower level, S'1 and S3. The new 52 CNP analogues (CNPA), which were crosscheck evaluated by QSAR complexation exhibit picomolar range predicted potency IC₅₀^{pre} and favorable predictive pharmacokinetic profile, (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). They are recommended for the synthesis and biological evaluation of FP2 inhibitory activity.

ACKNOWLEDGEMENTS

Authors are thankful for Fundamental and Applied Physics Laboratory, University Nangui Abrogoua, Ivory Coast to provide necessary facilities for this work.

AUTHOR'S CONTRIBUTIONS

Fagnidi YKH: writing original draft, methodology, investigation. **Ziki E:** writing, review and editing, methodology. **Gabin Allangba KNP:** formal analysis, supervision. **Toï B:** writing, review, methodology, data curation. **Megnassan E:** writing, review, and editing,

data curation. The final manuscript was read and approved by all authors.

DATA AVAILABILITY

The data and material are available from the corresponding author on reasonable request.

CONFLICT OF INTEREST

None to declare.

REFERENCES

- 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>
- World malaria report 2022. Geneva: World Health Organization; 2022. Licence: CC BY-NC-SA 3.0 IGO.
- Hyde JE. Drug-resistant malaria—an insight. *FEBS J*. 2007; 274(18):4688-4698. <https://doi.org/10.1111/j.1742-4658.2007.05999.x>
- Sutherland CJ, Lansdell PL, Sanders M, 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>
- 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>
- 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>
- 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>
- 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>
- Juliane Wunderlich, Petra Rohrbach, John Pius Dalton. The malaria digestive vacuole. *Front Biosci* 2012;4:1424-1448. <https://doi.org/10.2741/s344>
- Rosenthal PJ. Protease Inhibitors. In: *Antiparasitic Chemotherapy: Mechanisms of Action, Resistance, and New Directions in Drug Discovery*. New York: Springer; 2001:325-345.
- 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>
- 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>
- 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>
- 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 USA* 1990 Apr; 87(8):2931-5. <https://doi.org/10.1073/pnas.87.8.2931>
- 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%2Fjfm8013663>
- Boris DB, Fidele NK, Luc COO, Eugene M. (2016). Targeting cysteine proteases from *Plasmodium falciparum*: A general overview, rational drug design and computational approaches for drug discovery. *Current Drug Targets* 2016; 17, 1-26. <http://dx.doi.org/10.2174/1389450117666161221122432>
- 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>
- 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>
- Löser R, Schilling K, Dimmig E, Gütschow M. Interaction of papain-like cysteine proteases with dipeptide-derived nitriles. *J Med Chem* 2005; 48(24):7688-7707. <https://doi.org/10.1021/jm050686b>
- Löser R, Gut J, Philip J, et al. Antimalarial activity of azadipeptides nitriles. *Bioorg Med Chem Lett* 2010; 20 (1):252-255. <https://doi.org/10.1016/j.bmcl.2009.10.122>
- 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>
- Discovery Studio molecular modeling and simulation program, version 2.5, Accelrys, Inc., San Diego, CA 2009; California
- 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>
- 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>
- Dugas H. Basic principles in molecular modeling, theoretical and practical aspects. 4th ed, University of Montreal Bookstore; 1996.
- 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)
- 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>
- 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>
- Discovery Studio Molecular Modeling and Simulation Software. USA: San Diego; 2009.
- Böttcher CJF. Theory of electric polarization. Amsterdam, The Netherlands: Elsevier; 1973.
- Schrodinger. *QikProp, 6.5 (Release 139)*; Schrodinger LLC: New York, NY, USA, 2019.
- 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>
33. 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)
34. 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)
35. 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>
36. Allangba KNPG, 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, *J Enzyme Inhib Med Chem* 2019, 34:1, 547-561. <https://doi.org/10.1080/14756366.2018.1564288>
37. 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>
38. 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>
39. 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>