Research Article

Computer-aided Design of New Hydroxamic Acid Derivatives Targeting the *Plasmodium falciparum* M17 Metallo-aminopeptidase with Favorable Pharmacokinetic Profile

Moussa Koné[1], Hermann N'Guessan[2], Aka J. N'gouan[1], Frederica M-Koblavi[1], Eugene Megnassan[1-4]

[1]Laboratoire de Cristallographie–Physique Moléculaire, University of Cocody (now Felix Houphouët-Boigny), Abidjan, Côte d'Ivoire.
[2]Laboratory of Fundamental and Applied Physics, University of Abobo Adjamé (now Nangui Abrogoua), Côte d'Ivoire.
[3]Laboratory of Structural and Theoretical Organic Chemistry, University of Cocody (now Felix Houphouët Boigny), Côte d'Ivoire.
[4]ICTP-UNESCO, QLS, Strada Costiera 11, I 34151 Trieste, Italy.

ARTICLE INFO

ABSTRACT

Through structure-based molecular design, we virtually design new subnanomolar range antimalarial, inhibitors of *Plasmodium falciparum* M17 aminopeptidase (PfA-M17). We developed the complexation QSAR models from hydroxamic acid derivatives (HDA). A linear correlation was established between the computed Gibbs free energies of binding (GFE: $\Delta\Delta G_{com}$) and observed enzyme inhibition constants (K_i^{exp}) for each training set pK_i^{exp} = , R^2 = 0.97. The predictive power of the QSAR model was validated with 3D-QSAR pharmacophore generation (PH4): pK_i^{exp} = 0.707×pK_i^{pred} – 2.5182, R^2 = 0.89. We then conducted a study on catalytic residues to exploit the different interactions (enzyme: inhibitor). Structural information from the models guided us in designing a virtual combinatorial library (VCL) of more than 56 thousand HDAs. The PH4 screening retained 48 new and potent HDAs with predicted inhibitory potencies pK_i^{pre} up to 73 times lower than that of HDA1 (pK_i^{exp} = 2.5 nM). Combining molecular modeling and PH4 *in-silico* screening of the VCL resulted in the proposed novel potent antimalarial agent candidates with favorable pharmacokinetic profiles.

INTRODUCTION

The recent WHO burden of malaria raises a crucial concern since 98% of 627 000 deaths (77%) are children aged under 5 years) are located in Africa where at the same time the emergence and spread of artemisinin-resistant *Plasmodium falciparum* in WHO African region, namely Rwanda, Uganda and Eritrea in East Africa and thereafter Burkina Faso and Mali in West Africa almost a decade after the Cambodia and Thailand Mekong Subregion.[1-4] For now almost 400 years after Quinine (*Cinchona* bark known as cardinal's bark or Jesuits'

bark) from South America in response to Vatican urgent needs of antimalarials,[5] 40 years after Artemisinin (*Artemisia annua*) from Asia as China response to USA malaria-based war strategy in Vietnam,[6,7] Africa is awaited to find this new antimalarial expected to face the critical future since today less than 1% of the needed antimalarial funding comes from the most affected continent.[4] Aminopeptidases cleaving dipeptides and oligopeptides to their constituent amino acids[8,9] during the erythrocyte stage of the disease are among the most studied parasite targets about antimalarial design [10-12]

*Corresponding Author: Prof. Eugene Megnassan

Address: *ICTP-UNESCO, QLS, Strada Costiera 11, I 34151 Trieste, Italy.*

Email ✉: megnase@yahoo.com

Tel.: +225 0102363008

especially the neutral M1 alanyl-aminopeptidase (*Pf*A-M1) and M17 Leucyl aminopeptidase (*Pf*A-M17) are the most studied and validated as essential *Pf* enzyme targets whose inhibition should be lethal[13,14] making them attractive for the design of a new class of antimalarials.[15] They are metal ion-containing exopeptidases where the metal ion acts as enzyme-catalyzed reaction promotor by favoring hydrolysis.[16,17]

Among the various scaffolds designed as inhibitors of *Pf*A-M17 starting from the natural product Bestatin (1) mimicking the dipeptide d-Phe-l-Leu with the hydroxyl containing zinc ion binding group in red and the Phe (benzyl) and Leu (isoleucine) moieties fitting in the S1 (phenyl ring in π---π contact with Tyr575) and S1' pockets respectively reaching a biological activity of 25 nM.[18] A para-nitro substitution on the P1 benzyl for a better filling of the S1 pocket resulted in a ninefold potency rise (K_i = 2.7 nM) against the parasite. The dual-targets inhibitors phosphinates dipeptide (2) (*Pf*A-M1 and *Pf*A-M17) binding to *Pf*A-M17 through a bivalent contact with the zinc ion supported by the P1 (S1 pocket) and P1' (S1' pocket) phenyls in π---π stacking with Tyr575 and the catalytic His496 respectively reaching (K_i = 13.2 nM).[19] Phosphonates inhibitors (3), (4), also are dual targets designed to fit specifically the large S1 pocket of *Pf*A-M17 the phosphate group interacting with the zinc ion with the most potent (4) at 11 nM, the hydrophobic contact of 4-(1H-pyrazol-1-yl) phenyl group filling the *Pf*AM17 S1 pocket increasing the binding affinity.[12] Hydroxamate is well known orally safe drug design zinc binding group[11] and malonic hydroxamate inhibitors exemplified by (5) in the rigidified phenylmeth-(Z/E)-ylidene geometry have high activity against *Pf*A-M17 (IC_{50} = 6 nM); this is explained by the Z-isomer instead of the E-isomer from docking experiment. The non-rigidified inhibitor is almost fivefold less potent (IC_{50} = 27 nM); the benzylmalonyl group fits into the S1' pocket and the phenylmeth-(Z)-ylidene moeity seats in the hydrophobic S1.[20] More recently amino-hydroxamic acid has emerged in place of phosphonates due to their better pharmacokinetic profile as common drug design zinc ion binding group combined with the promising 4-(1H-pyrazol-1-yl) phenyl in P1 and a pivaloyl group in P1' increasing the activity K_i from 11 (4) to 6 nM for N-(1-(4-(1H-Pyrazol-1-yl) phenyl)-2-(hydroxyamino)-2-oxoethyl)- pivalamide (5)[12,21] Replacement of 4-(1H-pyrazol-1-yl) phenyl in P1 by para bromophenyl (6) resulted in a drop of potency to K_i = 700 nM but the trifluophenyl in place of bromophenyl in P1 and keeping the t-butylcarbonyl moiety limited the drop to 60 nM (or 147 nM) (7).[22,23] Keeping in (7) the trifluorophenyl and enlarging the tertiary-butylcarbonyl to a tertiary-butyloxycarbonyl (adamantyl respectively) moiety is beneficial to activity (8) (90) respectively: K_i = 30 nM, (29 nM) [19,20]. All this suggests a lack of evidence about suitable pharmacophore of *Pf*A-M17 active
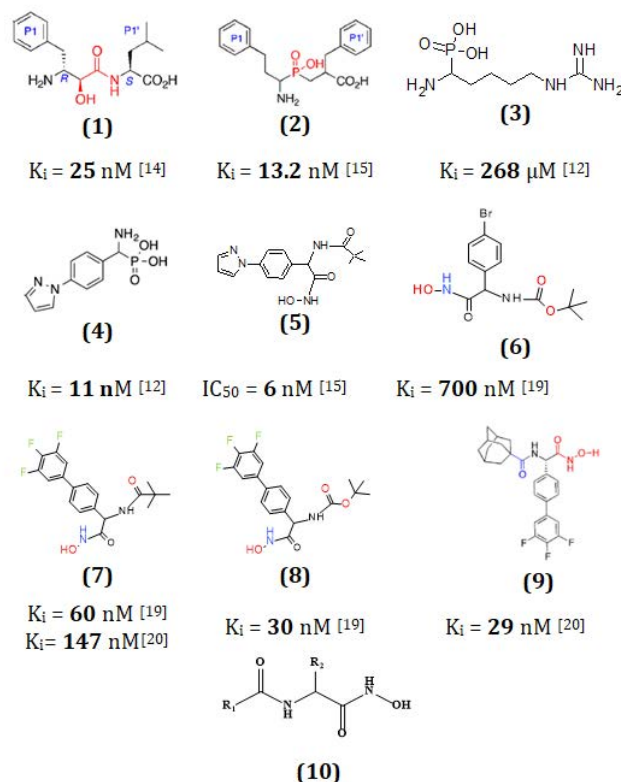

Fig. 1: Inhibitors of PfA-M17.

site to guide straightforward selection of R-groups in order to fill S1 and S1' pockets advantageously since the S1 binding pocket can undergo conformational changes to accommodate bulky groups.[9] Accordingly, we retain in our study the better and safer zinc ion binding group (hydroxamate) along with R_2–phenyl in P1 and varying P1' substitution groups (R_1) as SAR[18] from which a complexation QSAR model (10) will be built as well as a subsequent 3D-QSAR pharmacophore model as navigator to virtually explore the hydroxamate analogs chemical subspace with the expectation of improving binding affinity.

The metrics describing interactions at the *Pf*A-M17 active site have been assessed from the X-rays crystal structure analysis of *Pf*A-M17 (PDB code 6EEE) in complex with one of the most active studied inhibitors in this work (9).[20] The zinc binding group is in various contacts with Asp449, Glu461, Asp379 and Asp399 (not displayed on Fig 2 2D diagram) and the hydroxamate in a network of HB contacts with Gly469, Asp459, Lys386, Glu461, Asp379, Lys374 and Leu487. The hydrophobic adamantyl contacts in S1' involve Ala460, Ile547 the deepest residues at the pocket base and Arg463, Asn457, Lys552 and Ser554. The hydrophobic S1 cavity (Ala577, Leu492, Trp581, Leu492, Phe583, Met392, Met396 and Gly390) accommodates the trifluorophenyl moiety in π---alkyl with Leu492 and Leu395, π---sulfur with Met396 and halogen bond with Ala577.
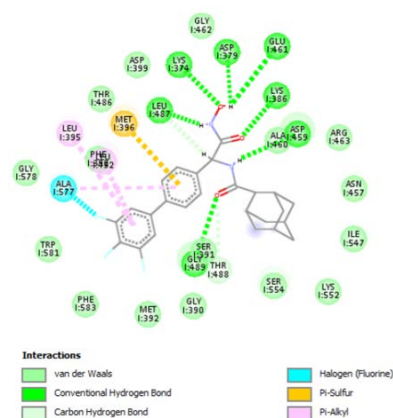
Fig. 2: *Pf*A-M17–(9) interactions at active site depicted in 2D.

METHODS

General Workflow

The general workflow start from SAR results along with the availability of X-rays structure of the target enzyme in complex with either one of the SAR compounds or at least any ligand. From this, *in-situ* modification will be performed to carefully build as many complexes as SAR compounds available in the training set. The computation of the Gibbs free energy ($\Delta\Delta G_{com}$) of each complex formation will lead, through a correlation with experimental activity (K_i), to the inhibitors' active conformations. The 3D-QSAR pharmacophore (PH4) generated from the active conformation (correlating HDA training set estimated K_i with experimental ones) will serve to screen a virtual library of HDA analogs to yield a handful of HDA best PH4 fit hits.

Training and Validation Sets

The chemical structures and biological activities (K_i^{exp}) of all the inhibitors used throughout this work are taken from the literature.[18-20] Their K_i^{exp} covers a very wide range ($2.5 \leq K_i^{exp} \leq 1700$ nM), more than four order of magnitude, suitable for a reliable QSAR model. Out of a total of 43 compounds, 36 were used for the training set (TS) and 7 for the validation (VS) set not used to build the QSAR model.

Model Building

The receptor preparation is a very important step in the complexation process. Therefore, molecular models of free inhibitors and complexes have been prepared from high-resolution crystallographic structure using the Insight-II 2005 molecular modelling programme and Discovery Studio Suite.[24,25] The *Pf*A-M17 has 2 crystallographic structures, these structures are coded as: 6EE2 and 6EEE.[20] These structures are available in a database of crystallographic structures: The Protein Data Bank (PDB). For this work, we chose the structure coded 6EEE because the hexamer is complete. It includes an endogenous ligand at its active site: 6k (N-(2-(Hydroxyamino)-2-oxo-1-
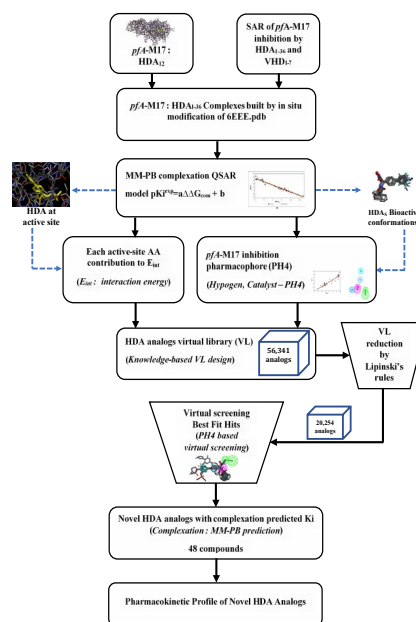
Scheme 1: General workflow of computer-aided design of the hydroxamates analogues.

(3′,4′,5′-trifluoro-[1,1′-biphenyl]-4-yl)ethyl)adamantane-1-carboxamide). The structure has been analyzed and prepared. No water molecule from the crystallographic structure was retained in the model. Based on the scaffold, substitutions in R1 and R2 have been made on the endogenous ligand to generate the whole series of parasite inhibitors. The washed structure undergo careful minimizations in order to preserve its tertiary structure. When simulating the crystal with a ligand other than the endogenous ligand, a systematic conformational research is carried out on the modified part of the ligand followed by its minimization extended to its vicinity in the active site over a radius of approximately 5Å.

Molecular Mechanics

The main advantage of molecular mechanics compared to other methods is the speed of the calculations suitable for very large molecular systems simulations. Since this method does not take into account the molecular electronic structure, it is not devoted to study systems in which electronic effects are predominant.[26] This study was carried out using the CFF force field.[21,22]

Conformational Search

The conformational search for the free inhibitors used in this study has been described fully earlier.[27]

Solvation Gibbs Free Energies

Ligand-receptor interactions take place in a solvent, which contributes through solvation phenomena to the binding process. However, the electrostatic component of Gibbs free energy (GFE) incorporating the effects of ionic strength through the resolution of the non-linear Poisson-

Boltzmann equation was calculated by the Delphi module of the Discovery Studio Suite.[28]

Calculation of Binding Affinity and QSAR Model

The calculation of binding affinity expressed as complexation GFE has been described fully earlier.[24]

Interaction Energy

The force field CFF was used to calculate the interaction energy (E_{int}) between the enzyme residues and the inhibitor as reported earlier.[24]

Pharmacophore Generation

The 3D-QSAR (PH4) pharmacophore generation protocol of DS[22] through its algorithm program Catalyst HypoGen[29] has been used to build the *Pf*A-M17 inhibition PHA as described earlier.[24]

ADME Properties

The pharmacokinetic profile of OHAs was calculated by the program QikProp[30] as described earlier.[24]

Virtual Library Generation

The virtual library generation has been described earlier.[24]

ADME-Based Library focusing

The orientation of the virtual library has been made thanks to numerous selection criteria as described earlier.[26]

Pharmacophore-based Library Searching

The pharmacophore model (PH4) derived from the bound conformations of HDAs at the active site of M17 served as a library searching tool as described earlier.[24]

Inhibitory Potency Prediction

The conformer with the best mapping on the PH4 pharmacophore in each cluster of the focused library subset was selected for in silico screening by the complexation QSAR model. The computed $\Delta\Delta G_{com}$ of each selected new analog was used for prediction of *Pf*A-M17 inhibitory potency (K_i^{pre}) of the focused virtual library of HDA analog by inserting this parameter into the target-specific scoring function in equation (1) parameterized using the complexation QSAR model of the training set of HDA inhibitors.[18-20]

$$pK_i^{pre} = -\log_{10}K_i^{pre} = a. \Delta\Delta G_{com} + b \qquad (1)$$

RESULTS

Training and Validation Sets

The 43 HDAs were selected from a homogeneous series of *Pf*A-M17 inhibitors with known experimentally determined inhibitory activities originating from a single laboratory.[18-20] The whole series was obtained by variations at two positions R1 and R2 on the backbone of

hydroxamic acid as show in Table 1. Their experimental inhibitory activities ($2.5 \leq K_i^{exp} \leq 1700$ nM) [18-20] cover a sufficiently wide concentration range as required for building 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.[31] In this study a training set of 36 HDAs and validation set of another 7 VDAs (Table 1) were created using the appropriate module of Discovery Studio Suite.[22]

The statistical data confirmed validity of the correlation Equations (A) and (B) plotted on Fig. 3. The ratio $K_i^{pre}/K_i^{exp} \approx 1$ (the K_i^{pre} values were estimated using correlation Equation B, Table 3) calculated for the validation set VDA1-7 documents the substantial predictive power of the complexation QSAR model from Table 2. Thus, the regression Equation B (Table 3) and computed $\Delta\Delta G_{com}$ GFEs can be used for prediction of inhibitory potencies K_i^{pre} against *Pf*A-M17 for novel HDA analogs, provided they share the same binding mode as the training set hydroxamic acid HDA1-36.
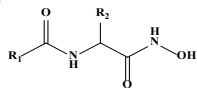
Binding Mode of HDAs

The inhibitors (HDAs) we have used throughout this work are a reported new series of hydroxamic acid obtained by synthesis.[18-20] Indeed, hydroxamic acids are used as metal ion chelators and the presence of the acid function in their molecular structure makes them particularly important for the inhibition of *Pf*A-M17. Active site have been assessed from the X-rays crystal structure analysis of *Pf*A-M17 (PDB code 6EEE) in complex with one of the most active studied inhibitors in this work (6).[20] The hydroxamate HB contacts active site with Gly469, Asp459, Lys386, Glu461, Asp379, Lys374 and Leu487. The hydrophobic adamantyl contacts in S1' involve Ala460, Ile547 the deepest residues at the pocket base and Arg463, Asn457, Lys552 and Ser554. The hydrophobic S1 cavity (Ala577, Leu492, Trp581, Leu492, Phe583, Met392, Met396 and Gly390) accommodates the trifluorophenyl moiety in π---alkyl with Leu492 and Leu395, π---sulfur with Met396 and halogen bond with Ala577.

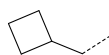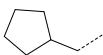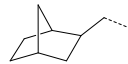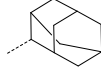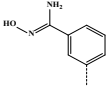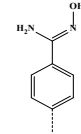### Interaction Energy

The energetic understanding of the different ligands (HDAx) used in the inhibition of *Pf*A-M17 is also provided by the energy interaction diagram (IE; ΔE_{int}). The distribution of the interaction energy in contribution of the residues to the active site of *Pf*A-M17 is done in three classes: a first class of ligands with maximum biological activity (Fig. 4), a second class of ligands with moderate biological activity and a third class of ligands with lowest biological activities.

For all classes, the interaction energy per residue is stronger with the residues Asp459, Leu487, Thr488, Ala460 and Gly489. No relevant difference is noticeable from class 1 to class 3 via class 2. Therefore in the lack of

Int. J. Pharm. Sci. Drug Res., May-June, 2023, Vol 15, Issue 3, 356-375

359

Table 1: Training set (HDA1-36) and validation set (VDA1-7) of *Pf*A-M17 inhibitors [18,19, 20] used in the elaboration of QSAR model of inhibitor binding. The R1 and R2 groups are numbered in the Table's column as #R ≡ group index.

#R	1	2	3	4	5	6	7
R-group	F	Br		F_3C			F_3C F F
#R	8	9	10	11	12	13	14
R-group		O					
#R	15	16	17	18	19	20	21
R-group	O						
#R	22	23	24	25	26	27	28
R-group	S	N–N		N	N	F	F
#R	29	30	31	32	33	34	35
R-group	F	CF_3	CF_3	F F	F F	F F	F F F
#R	36	37	38		39		
R-group	F F F	F F F F	NH_2 HO–N		H_2N OH N		

Training Set	HDA1	HDA2	HDA3	HDA4	HDA5	HDA6	HDA7	HDA8
#R₁-#R₂	6–36	6–33	6–28	6–34	6–24	6–23	6–37	6–39
	2.5	5.2	6.2	6.4	7	7.2	8.4	8.6
Training Set	HDA9	HDA10	HDA11	HDA12	HDA13	HDA14	HDA15	HDA16
#R₁-#R₂	6–38	6–27	9–39	21–35	6–29	9–27	9–23	9–29
	9.7	12	18	28.9	33	41	56	74
Training Set	HDA17	HDA18	HDA19	HDA20	HDA21	HDA22	HDA23	HDA24
#R₁-#R₂	8–35	19–35	9–28	17–35	9–31	6–35	12–35	14–35
	101	127	130	138	140	147	185	195
Training Set	HDA25	HDA26	HDA27	HDA28	HDA29	HDA30	HDA31	HDA32
#R₁-#R₂	9–35	7–35	9–25	11–35	15–35	13–35	16 - 35	5–35
	200	200	220	270	290	300	306	426
Training Set	HDA33	HDA34	HDA35	HDA36				
#R₁-#R₂	10–35	9–26	4–35	9–1				
	492	540	999	1700				
Validation Set	VDA1	VDA2	VDA3	VDA4	VDA5	VDA6	VDA7	
#R₁-#R₂	6–22	6–32	9–2	9–30	3–35	18–35	20–35	
	9,1	13	80	120	158	288	318	

Table 2: Gibbs free energy (binding affinity) and its components for the training set of *PfA*-M17 inhibitors HDA1-36 and validation set inhibitors VDA1-7.[18-20]

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]	K_i^{exp} [g]
HDA1	380	-10.43	2.25	-2.12	-6.06	2.5
HDA2	362	-7.43	1.33	-0.77	-5.33	5.2
HDA3	344	-5.67	0.11	0.45	-6.01	6.2
HDA4	362	-6.42	2.79	-0.38	-3.25	6.4
HDA5	326	-5.39	1.44	2.05	-6.00	7
HDA6	332	-5.54	1.05	1.27	-5.76	7.2
HDA7	416	-6.96	-1.05	-4.95	-3.06	8.4
HDA8	384	-6.83	2.67	-0.22	-3.94	8.6
HDA9	384	-6.50	2.87	0.00	-3.63	9.7
HDA10	344	-3.72	1.37	0.64	-2.99	12
HDA11	400	-7.67	3.02	-1.82	-2.83	18
HDA12	458	0.00	0.00	0.00	0.00	28.9
HDA13	344	-3.24	2.55	1.20	-1.89	33
HDA14	360	2.20	-2.01	2.10	-1.91	41
HDA15	346	1.34	3.16	2.23	2.27	56
HDA16	360	1.03	1.03	-0.35	2.41	74
HDA17	395	2.68	1.31	-0.31	4.30	101
HDA18	432	1.73	-1.18	-3.63	4.18	127
HDA19	360	4.95	0.45	0.18	5.22	130
HDA20	406	3.97	0.41	-0.55	4.93	138
HDA21	410	5.10	-3.20	-2.04	3.94	140
HDA22	380	3.64	-0.74	-2.62	5.52	147
HDA23	378	3.75	-0.90	-3.71	6.56	185
HDA24	406	4.20	-0.60	-2.91	6.51	195
HDA25	396	3.48	-0.86	-2.06	4.68	200
HDA26	442	2.31	-2.14	-7.11	7.28	200
HDA27	343	6.13	1.40	1.99	5.54	220
HDA28	378	5.57	-1.23	-3.97	8.31	270
HDA29	408	5.33	-1.50	-3.86	7.69	290
HDA30	392	5.86	0.18	-1.33	7.37	300
HDA31	392	5.02	-1.10	-3.59	7.51	306
HDA32	366	6.92	-1.47	-3.07	8.52	426
HDA33	364	7.71	-1.74	-4.18	10.15	492
HDA34	343	9.57	-0.08	-0.03	9.52	540
HDA35	406	10.27	-3.26	-5.56	12.57	999
HDA36	284	16.38	0.38	1.70	15.06	1700
VDA1	332	-3.48	1.07	2.08	-4.49	1.00
VDA2	362	-3.77	-0.51	-1.48	-2.80	0.99
VDA3	344	1.42	3.10	1.60	2.92	1.11
VDA4	410	3.63	-0.50	-3.87	7.00	0.96
VDA5	352	6.64	-1.24	-1.56	6.96	0.98
VDA6	420	6.73	0.86	0.13	7.46	1.00
VDA7	440	4.07	0.14	-2.18	6.39	1.03

[a] for the chemical structures of the training set of inhibitors see Table 1; [b] Mw is the molar mass of inhibitors; [c] $\Delta\Delta H_{MM}$ is the relative enthalpic contribution to the GFE change related to E-I complex formation derived by MM; $\Delta\Delta H_{MM} \approx [E_{MM}\{E\text{-}I_x\} - E_{MM}\{I_x\}] - [E_{MM}\{E\text{-}I_{ref}\} - E_{MM}\{I_{ref}\}]$, Iref is the reference inhibitor HDA1; [d] $\Delta\Delta G_{sol}$ is the relative solvent effect contribution to the GFE change of E-I complex formation: $\Delta\Delta G_{sol} = [G_{sol}\{E\text{-}I_x\} - G_{sol}\{I_x\}] - [G_{sol}\{E\text{-}I_{ref}\} - G_{sol}\{I_{ref}\}]$; [e] $-\Delta\Delta TS_{vib}$ is the relative entropic contribution of inhibitor to the GFE of E-I_x complex formation: $\Delta\Delta TS_{vib} = [TS_{vib}\{I_x\}E - TS_{vib}\{I_x\}] - [TS_{vib}\{I_{ref}\}E - TS_{vib}\{I_{ref}\}]$; [f] $\Delta\Delta G_{com}$ is the overall relative GFE change of E-I_x complex formation: $\Delta\Delta G_{com} \approx \Delta\Delta H_{MM} + \Delta\Delta G_{sol} - \Delta\Delta TS_{vib}$; [g] K_i^{exp} is the experimental inhibitory concentration of *PfA*.M17 obtained from ref.;[18-20] [h] ratio of predicted and experimental inhibitory potencies pK_i^{pre}/pK_i^{exp} ($pK_i^{pre} = -\log_{10}K_i^{pre}$) was predicted from computed $\Delta\Delta G_{com}$ using the regression equation for *PfA*-M17 shown in Table 3.

structural substitution directed to any active site residue, only a virtual screening of a diverse combinatorial library is expected to bring to novel more potent HDA analogs.

3D-QSAR Pharmacophore Model

PfA-M17 Active Site Pharmacophore

The interaction generation protocol in Discovery Studio molecular modeling program[22] provides the pharmacophore features of the active site of a protein. *PfA*-M17 predominantly displays aliphatic and hydrophobic aromatic features at the active site.

Generation and Validation of Pharmacophore

PfA-M17 inhibition 3D-QSAR pharmacophore was generated from the active conformation of 36 TS HDA$_{1-36}$ and evaluated by 7 VS VDA$_{1-7}$ covering a large range of experimental activity (2.5–1700 nM) spanning almost three orders of magnitude, suitable for a reliable pharmacophore hypothesis. The generation process is divided into three main steps: (i) the constructive step, (ii) the subtractive step, and (iii) the optimization step[24] as described earlier.

Accordingly, none of the training set HDAx was inactive and no starting PH4 features were removed. Finally, during the optimization phase, the score of the pharmacophoric hypotheses was improved. Hypotheses were scored according to errors in activity estimates from regression and complexity via a simulated annealing approach. At the end of the optimization, the top scoring 10 unique pharmacophore hypotheses were kept, all displaying five-point features. The cost values, correlation coefficients, root-mean square deviation (RMSD) values,

Table 3: Analysis of computed binding affinities $\Delta\Delta G_{com}$, its enthalpic component $\Delta\Delta H_{MM}$, and experimental inhibitory concentrations pK_i^{exp} = $-\log K_i^{exp}$ of HDAs towards *Pf*A-M17 [18,19, 20]

Statistical Data of Linear Regression	(A)	(B)
$pK_i^{exp} = -0.1171 \times \Delta\Delta H_{MM} + 7.3043$ (A)		
$pK_i^{exp} = -0.1253 \times \Delta\Delta G_{com} + 7.5022$ (B)		
Number of compound n	36	36
Squared correlation coefficient of regression R^2	0.94	0,97
LOO cross-validated Squared Correlation coef. R^2_{XV}	0.94	0,97
Standard error of regression σ	0.19	0,14
Statistical significance of regression. Fisher F-test	523.6	969,7
Level of statistical significance α	95%	95%
Range of activities K_i^{exp} [nM]	2.5–1700	

Fig. 3: (Top) plot of correlation equation between pK_i^{exp} and relative enthalpic contribution to the GFE ($\Delta\Delta H_{MM}$ [kcal.mol^{-1}]). (Bottom) similar plot for relative complexation Gibbs free energies of the *Pf*A.M17-HDA complex formation $\Delta\Delta G_{com}$ [kcal.mol^{-1}] of the training set.[18-20] The validation set data points are shown in red color.

the pharmacophore features, and the max-fit value of the top 10 ranked hypotheses (Hypo1-Hypo10) are listed in Table 4. They were selected based on significant statistical parameters, such as high correlation coefficient, low total cost, and low RMSD.

The generated pharmacophore models were then assessed for their reliability based on the calculated cost parameters ranging from 335.24 (Hypo1) to 434.97 (Hypo10). The relatively small gap between the highest and lowest cost parameter corresponds well with the homogeneity of the generated hypotheses and consistency of the TS of HDAx. For this PH4 model, the fixed cost (73.53) is lower than the null cost (2477.37) by a difference **Δ** = 2403.84. As reported earlier,[24] this difference is a major quality indicator of the PH4 predictability (**Δ** > 70 corresponds to an excellent chance or a probability higher than 90% that the model represents a true correlation.[22] A hypothesis has to be as close as possible to the fixed cost and as far as possible from the null cost to be statistically significant. For the set of 10 hypotheses, the difference **Δ** ≥ 597.58 attests to the pharmacophore model's high quality. The standard indicators such as the RMSD between the hypotheses, ranged from 3.78 to 4.46, and the squared correlation coefficient (R^2) falls to an interval from 0.95 to 0.92. The first PH4 hypothesis with the total costs (335.24) and best RMSD and R^2 was retained for further analysis. The statistical data for the set of hypotheses (costs, RMSD, R^2) are listed in Table 4. The configuration cost (14.78 for all hypotheses) far below 49 confirms this pharmacophore as reasonable. The evaluation of Hypo 1 is the mapping of the best active training set HDA1 (Fig. 5 (D)) displaying the geometry of the Hypo1 pharmacophore of M17 activation. The regression equation for pK_i^{exp} vs. pK_i^{pre} estimated from Hypo1: $pK_i^{exp} = 0.707 \times pK_i^{pre} - 2.5182$ (n = 36. R^2 = 0.89. R^2_{xv} = 0.89. F-test = 286.23. σ = 0.25, α > 95%) is also

plotted on Fig. 5 (E). Therefore the PH4 is good potential to choice the new HDA analogs.

We can carry out computational design and selection of new HDA analogs with elevated inhibitory potencies against *Pf*A-M17, based on a strategy using the noticeable presence of the hydrophobic features included in the best pharmacophore model at the position of R1 and R2 coupled with mapping to the hydrophobic aromatic ring feature and the appropriate ring substitution to the hydrophobic aliphatic feature in Hypo1 (Fig. 5).

Virtual Screening

In silico screening of a virtual (combinatorial) library can lead to hit identification as it was shown in our previous works on inhibitors design.[24,32]

Virtual Library

An initial virtual library (VL) was generated by substitutions at positions for R1 and R2 (see Table 5) on the scaffold: R1× R2 = 547 × 103 = 56,341 analogs. In order to match the substitution pattern of the best training set ligand HDA1 and consider the reported structural information about S pockets filling suitable for substitution not excluded through the Lipinski's rule violation (M_w > 500 g/mol[33] the VL underwent a focusing. To increase the content of drug-like and orally bioavailable analogs, the initial VL was filtered in an ADME-based focusing step. Only those molecules that satisfied the Lipinski's rule of five[30] computed using QikProp,[27] were kept. From the initial set of 56,341 analogs, 20,254 fulfilled the Lipinski test. 48 best fitting analogs (PH4 hits) were retained and submitted to structure-based screening using the QSAR model and computed GFE of the M17: HDA
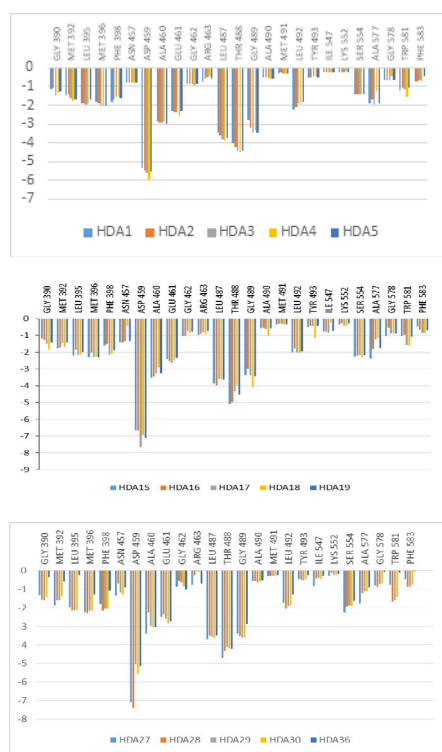
Fig. 4: Molecular Mechanics intermolecular interaction energy E_{int} breakdown to residue contributions in [kcal.mol^{-1}]: (A: Top) the most active inhibitors HDA1-5, (B: Middle) moderately active inhibitors HDA15-19, (C: Bottom) less active inhibitors HDA27-36,[18-20]

complex formation. The calculated $\Delta\Delta G_{com}$ of the M17: HDA complexes of the hits, their components, and predicted activities K_i^{pre} estimated from the correlation equation (B) (Table 3) are listed in Table 6.

In-silico Screening of Library of HDAs

The focused library of 20,254 analogs was further screened for molecular structures matching the 3D-QSAR PH4 pharmacophore model Hypo1 of *Pf*A-M17 inhibition. 48 HDAs mapped to at least 5 features of the pharmacophore. These best-fitting analogs (PH4 hits) then underwent complexation QSAR model screening. The computed GFE of *Pf*AM17-HDAx complex formation, their components, and predicted inhibitory activities K_i^{pre} calculated from the correlation equation B (Table 3) are listed in Table 6.

Novel HDA Analogs

The design of the virtual library of novel analogs was guided by structural information retrieved from the HDAx active conformation and was used to select appropriate substituents (R1- and R2-groups). In order to identify which substituents lead to new inhibitor candidates with the highest predicted potencies towards the *Pf*A-M17, we have prepared histograms of the frequency of occurrence of R1- and R2-groups among the 48 best fit PH4 hits (Fig. 6). The histograms show that the R1 group 530, 529 and 1 were represented with the highest

Table 4: Parameters of 10 generated PH4 pharmacophoric hypotheses for PfA-M17 inhibitor after Cat-Scramble validation procedure (49 scrambled runs for each hypothesis at the selected level of confidence of 98%).

Hypothesis	RMSD	R^{2b}	Total Cost ^{c}	Costs Difference ^{d}	Closest Random ^{e}
Hypo 1	3.78	0.95	335.2	2142.2	1033.4
Hypo 2	3.79	0.95	336.9	2140.5	1034.6
Hypo 3	3.97	0.94	362.5	2114.9	1034.7
Hypo 4	3.99	0.94	365.9	2111.5	1034.7
Hypo 5	4.24	0.93	402.5	2074.9	1034.7
Hypo 6	4.26	0.93	404.5	2072.9	1043.9
Hypo 7	4.26	0.93	405.7	2071.7	1099.8
Hypo 8	4.33	0.93	415.6	2061.8	1107.1
Hypo 9	4.45	0.92	433.6	2043.8	1116.6
Hypo 10	4.46	0.92	435.0	2042.4	1119.9
Fixed Cost	0	1	73.5		
Null Cost	11.6	0	2477.4		

Configuration cost = 11.40.
^{a} root mean squared deviation; ^{b} squared correlation coefficient; ^{c} overall cost parameter of PH4 pharmacophore; ^{d} cost difference between Null cost and hypothesis total cost; ^{e} lowest cost from 49 scrambled runs at a selected level of confidence of 98%
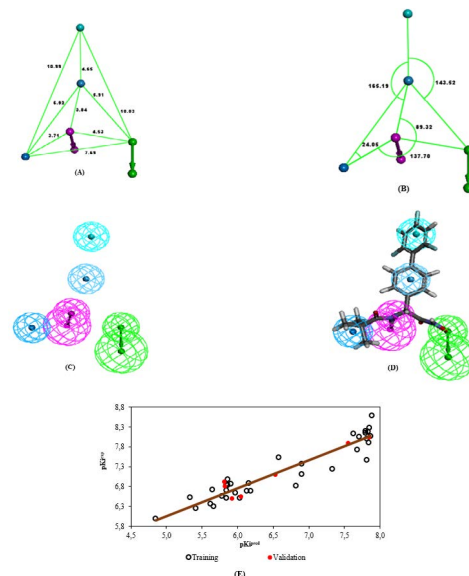
Fig. 5: (A) Distances between centers, (B) angles between centers of pharmacophoric features, (C) Features coordinates of centers, (D) mapping of pharmacophore of *Pf*A-M17 inhibitor with the most potent molecule HDA1. Feature legend: HYDA = Hydrophobic Aliphatic (blue), HYDAr = Hydrophobic Aromatic (sky blue), HBA = Hydrogen bond Acceptor (green), HBD = Hydrogen bond Donor (pink). (E) Correlation plot of experimental vs. predicted inhibitory activity.

frequency of occurrence (8, 4 and 4 respectively) among the 48 HDA hits. The R2-groups most frequently represented in this subset are 5, 1 and 7 with occurrences of 9, 5 and 5 respectively. The top ten scoring virtual

hits, namely, analogs are 530-29 ($K_i^{pre.}$ = 34 pM), 529–99 (36 pM), 529–100 (43 pM), 19–1 (46 pM), 530–98 (48 pM), 530–103 (50 pM), 530–95 (52 pM), 19–99 (64 pM), 530–94 (84 pM) and 530–102 (102 pM). They include the following substituents at R1 position : 530 : 2-adamantyl (6), 529 : adamantyl (2) and 19 : tert-butoxy (2), or at R2 position : 29 : 4'-((dimethylphosphoryl)oxy)-[1,1'-biphenyl]-4-yl (1), 99 : 7-phosphonopyren-2-yl (2), 100 : 4-(4-((dimethylphosphoryl)- oxy)cyclopenta-1,4-dien-1-yl) phenyl (1), 1 : (1), 98 : 4'-phospho-no-[1,1'-biphenyl]-4-yl (1), 103 : 4'-(dimethoxyphosphoryl)-3',5'-difluoro-[1,1'-biphenyl] -4-yl (1), 95 : 6,8-difluoro-7-phos-phonopyren-2-yl (1), 94 : 5,6,10-trifluoro-*9H*-cyclopenta[*a*]py-ren-2-yl (1) and 102 : 3',5'-difluoro-4'-(hydroxy(metoxy)phos-phoryl)-[1,1'-biphenyl]-4-yl (1). These R2-groups, all of which are hydrophobic rings, have a suitable substituent in the p-position that explores the depth of the S1 pocket. Due to the amino acid composition of the larger hydrophobic S1' pocket, R1-groups show preferences for larger aliphatic building blocks from butyl to decyl.[20]

Substitutions in the R1 position sufficiently occupying the S1' pocket and in the R2 position deeply embedded in the narrow S1 hydrophobic pocket of the HDAs led to an overall increase in the binding affinity of *Pf*A-M17, as illustrated by the inhibitory potencies of the majority of the newly designed analogues. The best-designed HDA, 530–29, has a predicted inhibitory potency of K_i^{pre} = 34 pM which is approximately 70 times lower than the most active compound in the TS series from the work of Drinkwater *et al.*,[19] and Vinh *et al.*[20]

Pharmacokinetic Profile of Novel HDA Analogs

The question of the pharmacokinetic profile has been a very important issue in the search for new inhibitors, particularly *Pf*A-M17 inhibitors.[18-20] For example, the best-designed triclosan derivative with very low oral bioavailability due to its low water solubility and rapid phase II metabolism, which must be optimized for possible use as an antituberculosis drug, possibly as an antimalarial drug in cases of high affinity with *Pf*EACP.[34,35]

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 18 descriptors related to adsorption, distribution. Metabolism and excretion (ADME properties) of the ligands were computed by the QikProp program[31] based on the methods of Jorgensen. According to these methods, experimental results of more than 710 compounds, including about 500 drugs and related heterocycles, were correlated with computed physicochemical descriptors, accurately predicting the molecule's pharmacokinetic profile. Drug likeness (#stars) is represented by the number of descriptors that exceed the range of values determined for 95% of known drugs out of 24 selected descriptors computed by QikProp.[31]
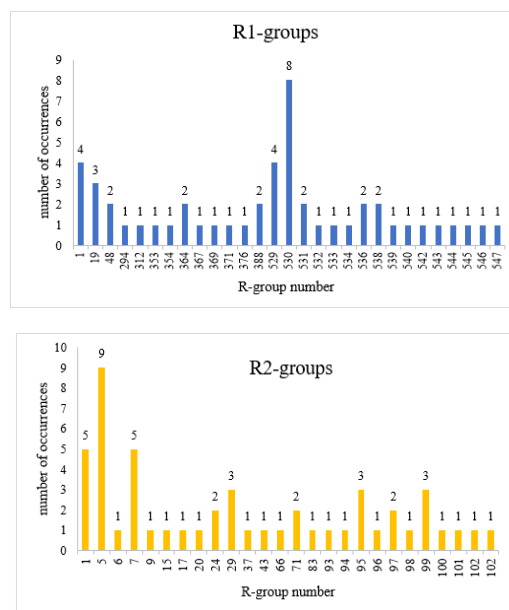

Fig. 6: Histograms of frequency of occurrence of individual R-groups in the 48 best selected analogs mapping to four features of the PH4 pharmacophore hypothesis Hypo1

Drug-likeness was used as the global compound selection criterion related to ADME properties. The selected ADME descriptors were calculated from 3D structures of the compounds considered. They were used to assess the pharmacokinetics profile of designed compounds (Table 7) The values for the best-designed HDAs are compared with those calculated for current drugs used for the treatment of malaria or in clinical trials (Table 7).

DISCUSSIONS

The structural requirements for tight binding of HDAs to *Pf*A-M17 include hydrophobic moiety in S1 pocket which can, through conformational changes, accommodate bulkier group, fitting S1' hydrophobic pocket and a zinc binding group. The best-designed HDA analogs are displayed on Fig. 7 through their mapping to the *Pf*A-M17 inhibition PH4, 3D interactions and Connoly surface of the active site. The studied nanomolar range inhibitory concentration HDAx series bind to the target enzyme through hydrogen bonds, π---π, π---alkyl, π---sulfur and hydrophobic contacts which are assessed by *Pf*A-M17:HDAx interaction energy breakdown to each active site residue, a quantitative descriptor useful to evaluate residue – inhibitor interactions unavoidable in case of hydrophobic contacts. Comparing the best active training set HDA1 and one of the top designed HAD analog 530-29 through the residues' contribution to interaction energies (Fig. 8) one can notice: for the residues involved in S1 – Inhibitor interactions are Ala577 (HDA1: ΔE_{int} = -2, and 530-29: ΔE_{int} =-3 kcal.mol^{-1}), Leu492 (-2.2, and 530-29: -2.7 kcal.mol^{-1}), Trp581 (-1.2, and 530-29: -2.3 kcal.mol^{-1}), Phe583 (-0.8, and 530-29: -1.6 kcal.mol^{-1}), Met392 (-1.5,

Table 5: R-groups (fragments. Building blocks. Substituents) used in the design of the diversity VL of HDA analogs.
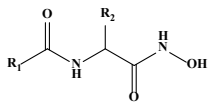
R1-groups					
1	tert-butyl	2	isopropyl	3	ethyl
4	2,2,2-trifluoroethyl	5	perfluoroethyl	6	Buta-1,3-dien-2-yl
7	penta-1,3-dien-3-yl	8	penta-1,4-dien-3-yl	9	3-vinylpenta-1,4-dien-3-yl
10	3-neopentyl	11	3-isobutyl	12	3-propyl
13	3,3,3-trifluoropropyl	14	2,2,3,3,3-pentafluoropropyl	15	2-methylenebut-3-en-1-yl
16	2-vinylbut-2-en-1-yl	17	2-vinylbut-3-en-1-yl	18	2,2-divinylbut-3-en-1yl
19	tert-butoxy	20	isoprotoxy	21	ethoxy
22	2,2,2-trifluoroethoxy	23	perfluoroethoxy	24	buta-1,3-dien-2-yloxy
25	penta-1,3-dien-3-yloxy	26	penta-1,4-dien-3-yloxy	27	(3-vinylpenta-1,4-dien-3-yl)oxy
28	tert-butylthio	29	isopropylthio	30	ethylthio
31	(2,2,2-trifluoroethyl)thio	32	(perfluoroethyl)thio	33	buta-1,3-dien-2-ylthio
34	penta-1,3-dien-3-ylthio	35	penta-1,4-dien-3-ylthio	36	(3-vinylpenta-1,4-dien-3-yl)thio
37	cyclopropyl	38	oxiran-2-yl	39	aziridin-2-yl
40	1-formylaziridin-2-yl	41	1-(hydrazineylidenemethyl)aziridin-2-yl	42	cyclobuta-1,3-dien-1-yl
43	cyclobutyl	44	3-methylcyclobuta-1,3-dien-1-yl	45	2-methylcyclobuta-1,3-dien-1-yl
46	azet-3-yl	47	4-(trifluoromethyl)cyclobuta-1,3-dien-1-yl	48	3,4-dimethylcyclobuta-1,3-dien-1-yl
49	4-methyl-3-(trifluoromethyl)cyclobuta-1,3-dien-1-yl	50	2-aminoazet-3-yl	51	3-aminocyclobuta-1,3-dien-1-yl
52	3-(difluoroamino)cyclobuta-1,3-dien-1-yl	53	3-(1-aminovinyl)cyclobuta-1,3-dien-1-yl	54	3-aminocyclobut-2-en-1-yl
55	3-methylcyclobut-2-en-1-yl	56	2-methylcyclobut-2-en-1-yl	57	2-aminocyclobut-2-en-1-yl
58	4-aminocyclobut-1-en-1-yl	59	cyclobut-1-en-1-yl	60	3-methylcyclobut-1-en-1-yl
61	4-methylcyclobut-1-en-1-yl	62	3-aminocyclobut-1-en-1-yl	63	3-carbamoylcyclobut-1-en-1-yl
64	3-fluorocyclobut-1,3-dien-1-yl	65	2,3-difluorocyclobut-1,3-dien-1-yl	66	3,4-difluorocyclobut-1,3-dien-1-yl
67	4-fluorocyclobut-1,3-dien-1-yl	68	2-fluorocyclobut-1,3-dien-1-yl	69	2,3,4-trifluorocyclobut-1,3-dien-1-yl
70	2,4-difluoro-3-(trifluoromethyl) cyclobut-1,3-dien-1-yl	71	3-amino-2,4-difluorocyclobut-1,3-dien-1-yl	72	1*H*-pyrrol-1-yl
73	3-fluoro-1*H*-pyrrol-1-yl	74	3,4-diamino-1*H*-pyrrol-1-yl	75	3,4-difluoro-1*H*-pyrrol-1-yl
76	3-chloro-4-fluoro-1*H*-pyrrol-1-yl	77	3-fluoro-4-formyl-1*H*-pyrrol-1-yl	78	3-fluoro-4-(2,2,2-trifluoroacetyl)-1*H*-pyrrol-1-yl
79	3-(cyanocarbonyl)-4-fluoro-1*H*-pyrrol-1-yl	80	3-cyano-1*H*-pyrrol-1-yl	81	6-cyanopyridazin-3-yl
82	5-cyanopyridin-2-yl	83	5-cyanopyrimidin-2-yl	84	5-cyanopyrazin-2-yl
85	Perfluoro-1*H*-pyrrol-1-yl	86	2,4-dibromo-3,5-difluoro-1*H*-pyrrol-1-yl	87	3-methyl-1*H*-pyrrol-1-yl
88	3-(trifluoromethyl)-1*H*-pyrrol-1-yl	89	3,5-dihydro-4*H*-1,2,4-triazol-4-yl	90	2,5-dihydro-1*H*-pyrrol-1-yl
91	3,4-diamino-2,5-dihydro-1H-pyrrol-1-yl	92	2*H*-pyrrol-3-yl	93	3*H*-pyrazol-4-yl