Original article

Design and *in silico* screening of combinatorial library of antimalarial analogs of triclosan inhibiting *Plasmodium falciparum* enoyl-acyl carrier protein reductase

Vladimir Frecer [a,b,*], Eugene Megnassan [a,c], Stanislav Miertus [a,*]

[a] *International Centre for Science and High Technology, UNIDO, AREA Science Park, Trieste I-34012, Italy*
[b] *Cancer Research Institute, Slovak Academy of Sciences, Bratislava SK-83391, Slovakia*
[c] *Laboratory of Fundamental and Applied Physics, University of Abobo-Adjame, Abidjan 02 BP 801, Cote d'Ivoire*

ABSTRACT

Enoyl-acyl carrier protein reductase of *Plasmodium falciparum* (PfENR) is an important target for anti-malarial agents that interfere with the FAS-II pathway of lipid synthesis, which is specific for the parasite. Recent studies showed that substituted analogs of triclosan (TCL) inhibit the purified PfENR enzyme with IC_{50} values below 200 nM when the suboptimal 5-chloro group was replaced by larger hydrophobic moieties. We have used computer-assisted combinatorial techniques to design, focus and *in silico* screen a virtual library of TCL analogs substituted at positions 5, 4′ and 2′. Our study can thus direct synthetic chemists working on the antimalarial FAS-II inhibitors towards the explored subset of the chemical space, which is predicted to contain compounds with PfENR inhibition potencies in the low nanomolar range and favorable ADME properties.

© 2009 Elsevier Masson SAS. All rights reserved.

1. Introduction

Malaria is endemic in inter-tropical areas, which include about 40% of the world's population and is responsible for over one million deaths annually, mainly in children under the age of 5 years [1]. *Plasmodium falciparum* infection is the most virulent form of malaria, accounting for the vast majority of mortalities and over 500 million infections each year [1]. The occurrence and spread of resistance to traditional antimalarial drugs underscore the need for new potent and orally bioavailable antimalarials, which are active also against the existing drug-resistant strains of *P. falciparum* [2].

Inhibition of type II fatty acid biosynthesis (FAS-II) pathway, which is different from the associative FAS-I pathway present in mammalian cells, appears to hold significant promise in devising

novel antimalarials [3]. The final reaction in the FAS-II pathway of *P. falciparum* is catalyzed by the enoyl-acyl carrier protein reductase (PfENR), which mediates the NADH-dependent reduction of *trans*-2-enoyl-acyl carrier protein (ACP) to acyl-ACP [4,5]. Fatty acids play a critical role in providing metabolic precursors of biological membranes and represent an important form of metabolic energy production in the parasite, making their biosynthetic pathway an excellent target for antimicrobial agents. Triclosan (TCL), a common antibacterial agent (Fig. 1), is an uncompetitive inhibitor of purified PfENR, which has demonstrated inhibitory potency against *P. falciparum* parasites cultured *in vitro* ($K_i = 50$ nM) [6–8]. TCL analogs that are non-toxic to mammalian cells [7,9], which do not have ENR homologues, have been recently explored as potential antimalarial therapeutics [8,10–13].

Publication of the X-ray crystal structure of TCL and its analogs plus NAD^+ co-factor bound to the PfENR has opened the possibilities of computer-assisted drug design of antimalarial agents, which target the plasmodium parasite lipid synthesis [8,13]. The phenol moiety of TCL has been demonstrated to be crucial to the binding to both the enzyme and the co-factor [13]. Replacements of the chlorine atoms at positions 5, 2′ and 4′ on the TCL's diaryl ether scaffold, where substitutions by larger groups could be introduced to gain energetically favorable interactions with both the enzyme and the co-factor, were studied by several laboratories [10–14]. Synthesis of 4′-derivatives of TCL resulted in minor gains in the
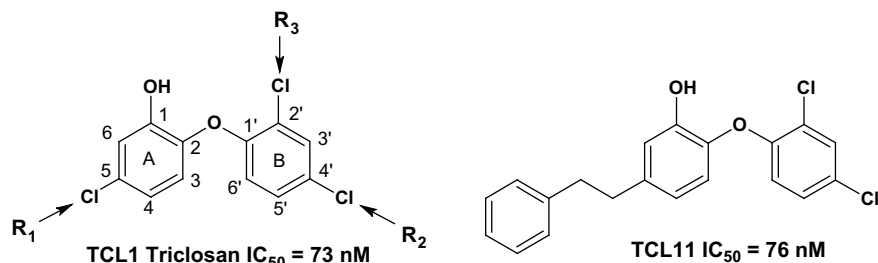
Fig. 1. Chemical structures of triclosan (TCL) and its analog TCL11 (Table 1). The PfENR inhibitory activities IC_{50}^{exp} were taken from Ref. [13]. Scaffold atom numbering and positions of the R-groups' attachment sites are indicated.

PfENR inhibition potencies [10] while investigation of analogs at the 2′-position yielded only micromolar activities [12]. Sullivan et al. [14] described synthesis and activity of 5-alkyl-substituted TCL derivatives towards purified PfENR and Chhibber et al. [11] studied TCL analogs with primarily hydrophilic 5-substituents, which were less potent than the TCL. Freundlich et al. [13] reported synthesis and activity testing of TCL analogs containing hydrophobic substituents in position 5. Their results demonstrated the possibility to improve the *in vitro* antimalarial potency of TCL by replacing the suboptimal 5-chloro group with larger hydrophobic moieties. The TCL scaffold has inspired numerous initiatives to design potent PfENR inhibitors in a search for new antimalarial drugs [15].

Combinatorial chemistry approaches have been developed to generate a large quantity of compounds prepared via parallel synthesis of libraries, which are screened in high-throughput bioassays to identify hits [16,17]. Computational methods are being increasingly used to assist library design, focusing and virtual screening by introducing selection criteria such as molecular diversity, drug likeness, Lipinski criteria for bioavailability and predicted receptor binding, to reduce the size of the combinatorial libraries to be actually synthesized and screened. Incorporation of structure-based methods into computer-assisted combinatorial chemistry has shifted the emphasize from large purely random diversity libraries to smaller focused combinatorial subsets. Virtual library design, focusing and *in silico* screening can thus significantly decrease the cost, time and labor required to generate new lead compounds. We have designed several classes of bioactive molecules using computer-assisted combinatorial approaches [18–20].

In the present study, we report the application of computer-assisted combinatorial techniques to design, focus and *in silico* screen a virtual library of TCL analogs. We have explored more extensively the chemical space around the TCL scaffold substituted at positions 2′, 4′ and especially at position 5 by smaller and larger groups with both hydrophobic and hydrogen-bonding characteristics, which could enhance the binding affinity to the PfENR, aiming at finding more potent selective FAS-II antimalarials. We have derived a small highly focused combinatorial subset of analogs of TCL, which contains virtual hits that are predicted to inhibit the PfENR in the low nanomolar concentration range and display favorable ADME properties.

2. Results and discussion

2.1. Library design and fragment-based focusing

The 2′-substituted TCL analogs can be prepared via coupling of phenolic A-ring and aromatic B-ring precursors through a nucleophilic aromatic substitution reaction, following the protocols published previously [10–14]. Aryl/alkyl substituents can be introduced at positions 5 and 4′ of the TCL scaffold (Fig. 1) as Grignard reagents

reacting with 5-aldehyde diaryl ether intermediate [13] or via acetylation, sufonation or reductive amination of 4′-aniline precursor [10], respectively. Relatively large diversity combinatorial library of TCL analogs ($\sim 10^6$ compounds) could be designed from the available reagents (building blocks, fragments, R-groups) listed in the databases of available chemicals [21]. A library of this size can exceed the capacity of parallel synthesis and high-throughput screening equipment. To design a focused library of a reduced size and increased content of drug-like orally bioavailable molecules, we have introduced a set of filters and penalties, which can help to select the most suitable reagents. The fragment-based library focusing relied on the predetermined optimum ranges of structural and physicochemical properties (descriptors) computed for fragments of known PfENR inhibitors, a subset of selected 10 most potent analogs of TCL (Table 1) taken from Refs. [10,12,13]. Twenty different descriptors that characterize a wide spectrum of various properties, such as molecular shape and size, polarity, ability to interact with the receptor, conformational flexibility, composition and structural complexity, were computed for the substituents on the TCL scaffold corresponding to R_1, R_2 and R_3-groups (Fig. 1) using

Table 1
Training and validation sets of TCL derivatives used in the QSAR model of PfENR reductase inhibition [10,12,13] and in parameterization of the target-specific scoring function.

Training set	R_1[a]	R_2	R_3	IC_{50}^{exp} [b] [nM]
TCL1 triclosan	…Cl	…Cl	…Cl	73
TCL2	…C(=O)NH$_2$	…Cl	…Cl	21,000
TCL3	…(CH$_2$)$_3$CH$_3$	…Cl	…Cl	480
TCL4	…CH$_2$CH(CH$_3$)$_2$	…Cl	…Cl	180
TCL5	…CH$_2$CH(CH$_3$)CH$_2$CH$_3$	…Cl	…Cl	290
TCL6	…(CH$_2$)$_2$CH(CH$_3$)$_2$	…Cl	…Cl	120
TCL7	…Ph–(o)CH$_3$	…CN	…Cl	410
TCL8	…Ph–(m)CH$_3$	…Cl	…Cl	230
TCL9	…Ph–(p)CH$_3$	…Cl	…Cl	190
TCL10	…CH$_2$–Ph	…Cl	…Cl	71
TCL11[c]	…(CH$_2$)$_2$–Ph	…Cl	…Cl	76
TCL12	…2-Pyr	…CN	…Cl	700
TCL13	…CH$_2$–2-Pyr	…Cl	…Cl	640
TCL14	…CH$_2$–3-Pyr	…Cl	…Cl	840
TCL15	…Cl	…H	…NH$_3^+$	7000
TCL16	…Cl	…H	…CH$_2$NH(CH$_2$)$_2$–Ph	2500

Validation set	R_1[a]	R_2	R_3	$IC_{50}^{pre}/IC_{50}^{exp}$ [d]
VCL1	…CH$_2$–CycHex	…Cl	…Cl	1.12
VCL2	…Ph–(p)F	…Cl	…Cl	1.16
VCL3	…CH$_2$–4-Pyr	…CN	…Cl	1.19
VCL4	…Cl	…Cl	…CH$_2$NHCH$_2$–Ph	1.28

[a] For the position of R_1–R_3-groups see Fig. 1, dashed bonds indicate the attachment points of fragments.
[b] Experimental IC_{50}^{exp} of the training and validation sets of the TCL derivatives towards PfENR were taken from Refs. [10,12,13].
[c] Enzyme–inhibitor complex used in the virtual screening (PDB code 2OOS) [13].
[d] Ratio of predicted IC_{50}^{pre} and experimental IC_{50}^{exp} activities of the validation set, IC_{50}^{pre} was computed from the QSAR equation (1).
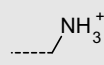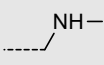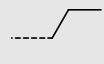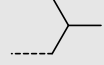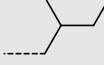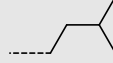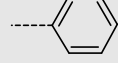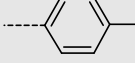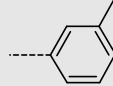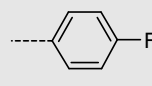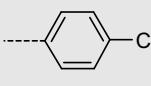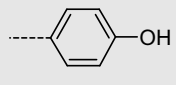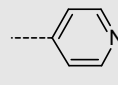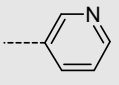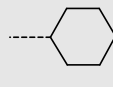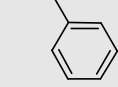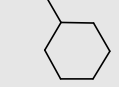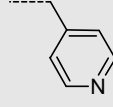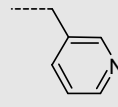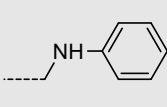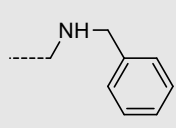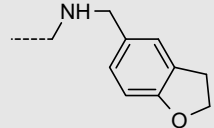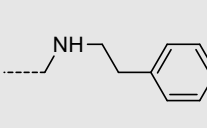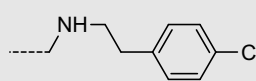
Table 2
R-groups (fragments, building blocks, reagents) used in the design of the initial diversity library of TCL analogs.

R-groups[a,b]							
1.	…H	2.	…F	3.	…Cl		
4.	…CH$_3$	5.	…OH	6.	…SH		
7.	…CN	8.	…NH$_3^+$	9.	(—CH$_3$ group)		
10.	(—NH–CH$_3$)	11.	(—N(CH$_3$)$_2$)	12.	(—CH$_2$OH)		
13.	(—CH$_2$SH)	14.	(—CH$_2$NH$_3^+$)	15.	(—CH$_2$NH–CH$_3$)		
16.	(—CH$_2$N(CH$_3$)$_2$)	17.	(—CH$_2$CH$_2$CH$_3$)	18.	(—CH(CH$_3$)$_2$)		
19.	(—CH(CH$_3$)CH$_2$CH$_3$)	20.	(—CH$_2$CH(CH$_3$)$_2$)	21.	(—CH(CH$_3$)CH(CH$_3$)$_2$)		
22.	(—CH$_2$CH$_2$CH(CH$_3$)$_2$)	23.	(phenyl)	24.	(4-methylphenyl)		
25.	(2-methylphenyl)	26.	(4-fluorophenyl)	27.	(4-chlorophenyl)		
28.	(4-hydroxyphenyl)	29.	(4-pyridyl)	30.	(3-pyridyl)		
31.	(cyclohexyl)	32.	(benzyl)	33.	(cyclohexylmethyl)		
34.	(4-pyridylmethyl)	35.	(3-pyridylmethyl)	36.	(—CH$_2$NH-phenyl)		
37.	(—CH$_2$NH-CH$_2$-phenyl)	38.	(—CH$_2$NH-CH$_2$-dihydrobenzofuran)	39.	(—CH$_2$NH-CH$_2$CH$_2$-phenyl)		
40.	(—CH$_2$NH-CH$_2$-(4-chlorophenyl))						

[a] R$_1$-groups: fragments 1–40, R$_2$-groups: fragments 1–23, R$_3$-groups: fragments 1–11, for the position of individual of R-groups see Fig. 1.
[b] Dashed bonds indicate the attachment points of the fragments.

the Cerius^2 program [22] (data not shown). The property upper limits for the R$_2$- and R$_3$-groups were augmented to permit the use of larger substituents due to somewhat limited diversity of the training set. Increased variation in the R$_2$- and R$_3$-groups is needed to design new TCL analogs. Available fragments whose descriptor values laid outside the optimum ranges and fragments with high similarity to other considered fragments were filtered out using

combined penalty and diversity score. Thus, 40 smaller polar and larger aliphatic and aromatic hydrophobic groups were selected as suitable and diverse building blocks for position 5 (R$_1$-group), Table 2. Out of them, 23 and 11 smaller aliphatic reagents were chosen as building blocks for positions 4′ (R$_2$-groups) and 2′ (R$_3$-groups), respectively, which are known to be able to accommodate smaller substituents [10–14]. Larger number of R$_1$ and in part also R$_2$-
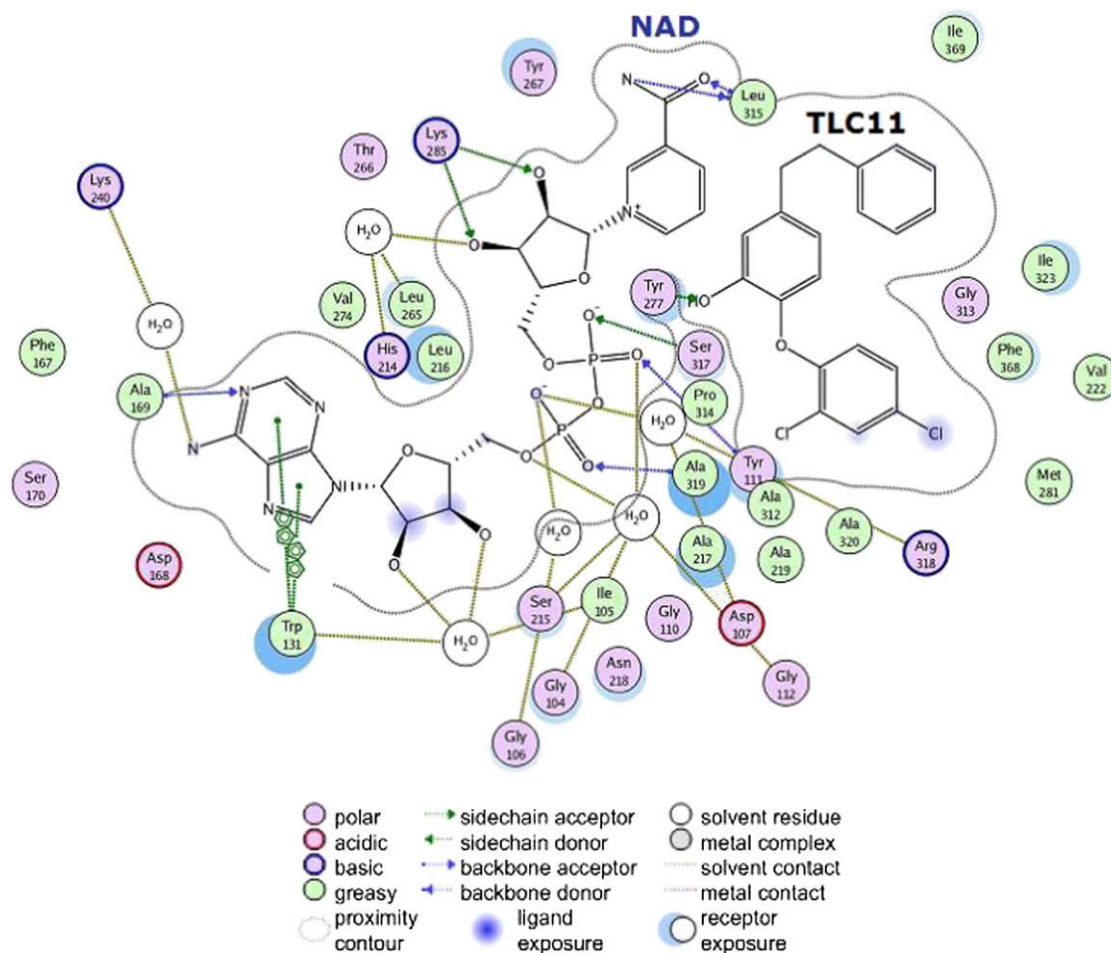
Fig. 2. TCL11 inhibitor–PfENR interaction scheme at the active site of the enzyme–ligand complex including the NAD⁺ co-factor. Interaction types are explained in the legend shown below.

groups was chosen in line with the elevated variability of the substituents filling the bulky R_1 cavity (Fig. 2) [8,13]. The size of the fragment-focused library of TCL derivatives was thus reduced to:

$$40(R_1) \times 23(R_2) \times 11(R_3) = 10,120 \text{ analogs}.$$

This virtual library was generated by attaching of the R-groups from the fragment-focused sets (Table 2) onto the diaryl ether scaffold of TCL (Fig. 1).

2.2. Analog-based focusing

In the next step, analog-based focusing was applied to further reduce the size of the generated library of TCL analogs. The analog-based focusing procedure relied on the optimum ranges of twenty molecular physicochemical descriptors including the Lipinski rule-of-five, derived for the 10 most potent inhibitors of the PfENR, which were selected from the training set (Table 1). A library of 1000 most diverse analogs with molecular characteristics compliant with the requirements for PfENR inhibitors and for bioavailable compounds was chosen from the fragment-focused library for a subsequent structure-based focusing step.

2.3. Structure-based focusing

To select a small highly focused combinatorial subset of TCL analogs with good predicted binding affinities to the PfENR,
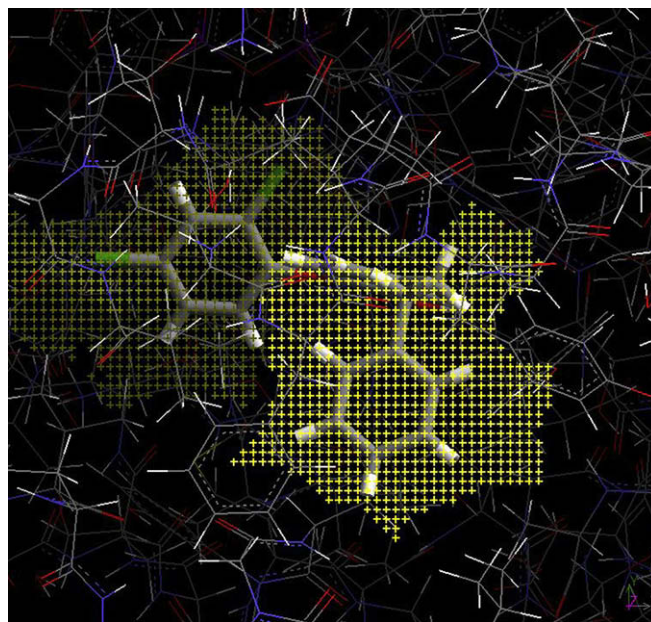
Fig. 3. Site model for the active site of the PfENR shown as yellow grid with the native ligand TCL11 (in stick representation) used for the binding site definition [13]. PDB entry code 2OOS. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
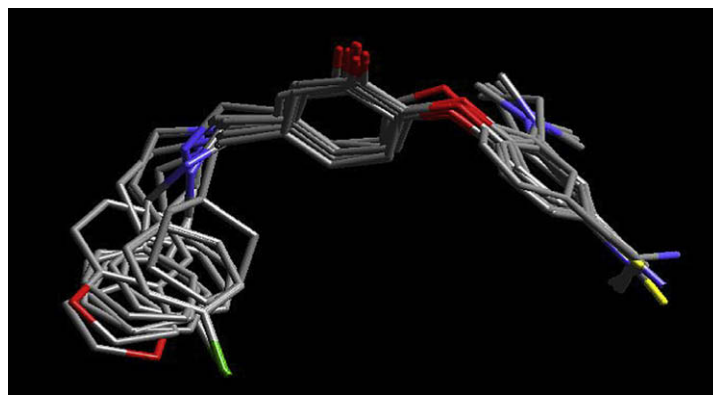
Fig. 4. Various TCL analogs docked to the binding site of the PfENR retaining the stacking interaction of the phenol ring with the nicotinamide ring and hydrogen-bonding with the 2′-hydroxyl group of the NAD$^+$ ribose [8,11,13]. The R_1, R_2 and R_3-groups replacing the chlorine atoms at positions 5, 2′ and 4′ fill the cavities of the binding site. Hydrogen atoms were omitted for a better clarity. Carbon atoms are shown in grey color, oxygens red, nitrogens blue, sulfur yellow and chlorine dark green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

structure-based focusing and *in silico* screening procedures have been further applied. Each analog was docked into the binding pocket of the PfENR represented by a binding site grid (Fig. 3) obtained from the crystal structure of the PfENR–TCL11 complex [13] by employing Monte Carlo ligand fitting algorithm of the LigFit module of Cerius^2 [22]. Comparison of the docked TCL analogs with the crystal structures of the TCL11 (Fig. 4) demonstrated that the binding mode for the designed analogs was very similar, showing the same stacking interaction with the nicotinamide ring of NAD$^+$ and comparable hydrogen-bonding pattern with the 2′-hydroxyl group of the NAD$^+$ ribose.

The docking procedure yielded 20 best binding conformers per analog, which were clustered into 10 conformational families depending on the mutual r.m.s. deviations. In each cluster, the conformer with the highest docking score was selected for subsequent virtual screening.

2.4. QSAR analysis of TCL analogs and parameterization of target-specific scoring function

To parameterize a scoring function specific for the PfENR target, we have correlated the scores implemented in the Cerius^2 program [22] with the experimental activities (IC_{50}^{exp}) of a training set of PfENR inhibitors TCL1–TCL16 (Table 1) [10,12,13]. The molecules of the training set were docked into the PfENR receptor model using the LigFit algorithm of Cerius^2 [22]. Various QSAR models relating the observed IC_{50} values to the computed LigScore, LUDI, PMF and PLP scores [22–25] were tested. From the set of available scoring functions, the LUDI score led to the best fit of the experimental activities to the computed enzyme–inhibitor binding energies predicted via scoring functions. The following QSAR equation was obtained by linear regression (Fig. 5):

$$pIC_{50} = -\log_{10} IC_{50}[PfENR] = -6.3473 + 0.0069\ LUDI \qquad (1)$$

(number of samples $n = 16$, correlation coefficient $R^2 = 0.83$, leave-one-out cross validated correlation coefficient $R_{xv}^2 = 0.78$, Fisher $F\text{-test} = 65.7$, statistical significance of the correlation $\alpha > 95\%$).

The cross validated correlation coefficient R_{xv}^2 of 0.78 indicates that a major portion of the variance of the training set data was well described by this QSAR model. In addition, the quality of the model and the scoring function parameterization was confirmed by predicting the PfENR inhibitory activities for a validation set of 4 TCL analogs VCL1–VCL4 (not included into the training set) with known IC_{50}^{exp} values (Table 1). The ratio of the predicted activities derived from Eq. (1) and observed activities $IC_{50}^{pre}/IC_{50}^{exp}$, which yielded

values close to 1, confirmed the predictive power of the QSAR model. The training and validation sets used in the QSAR model [10,12,13] display somewhat limited variation of the R_2- and R_3-groups' space due to restricted availability of experimental activity data. Prediction of inhibitory potencies by the trained target-specific scoring function that slightly exceed the activity ranges of the training set, is still possible, as QSAR models using ligand–receptor binding affinity estimate (LUDI score) of docked analogs are less sensitive to training set boundaries.

The derived scoring function (Eq. (1)), which is specific for the ENR of *P. falciparum*, was subsequently used for *in silico* screening and scoring of the conformers with the highest docking score of the library of TCL analogs docked to the active site of the PfENR in the preceding structure-based focusing step.

2.5. In silico screening

The best conformers of TCL analogs were screened *in silico* by using the LUDI score [23]. The predicted PfENR inhibitory activities
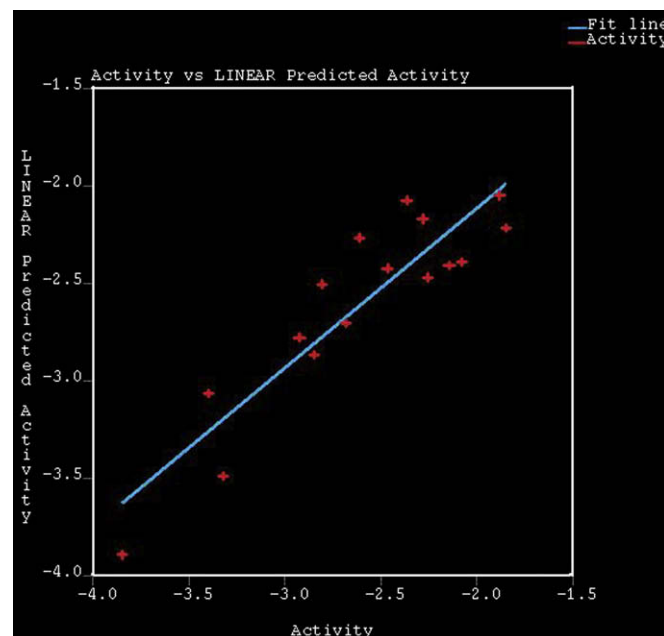

Fig. 5. Plot of QSAR regression equation derived for the training set of 16 TCL derivatives, which was used as the scoring function specific for the PfENR target.
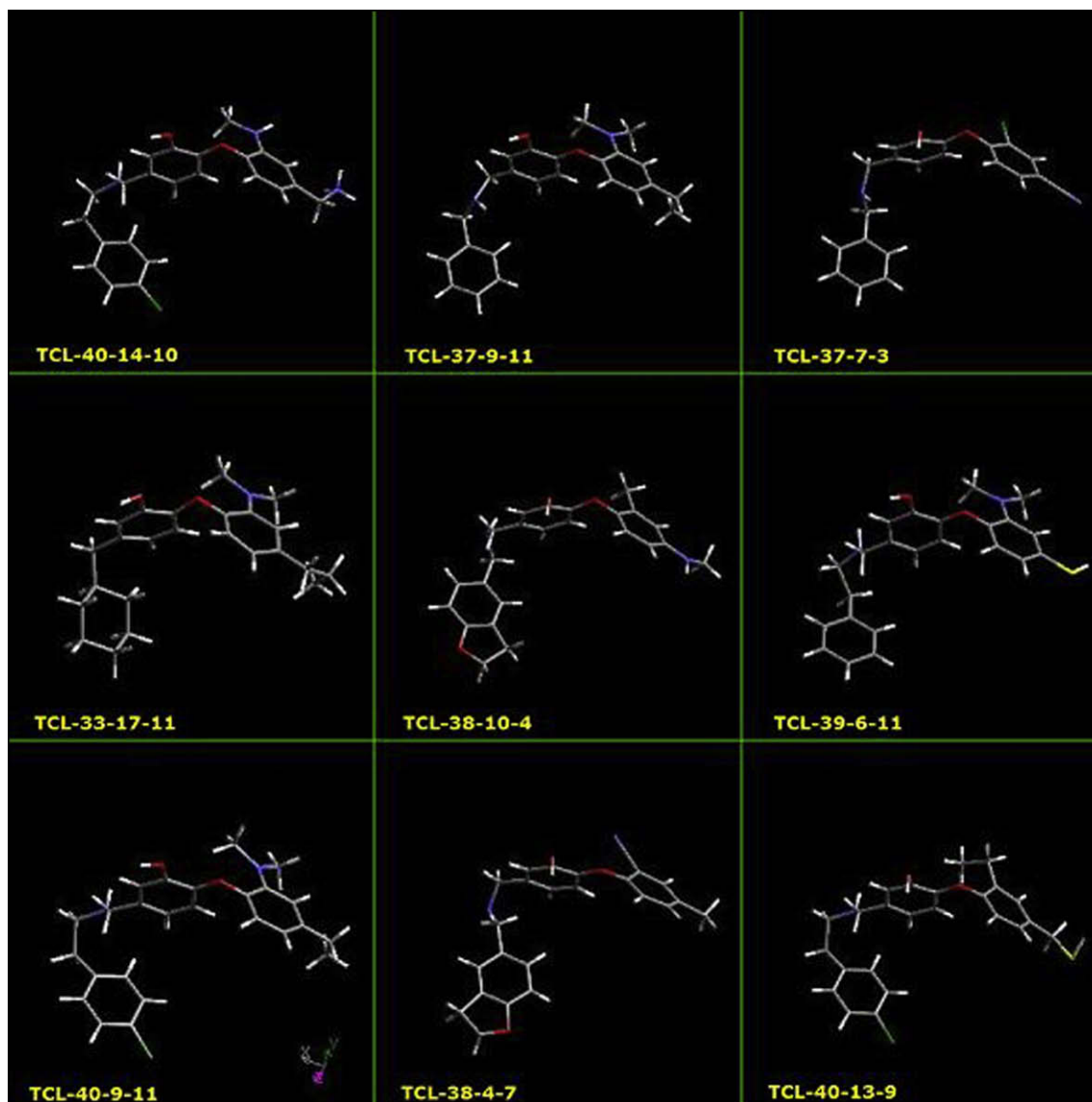
Fig. 6. 3D-structures of the nine best designed TCL analogs.

(IC_{50}^{pre}) of the designed analogs were calculated from the target-specific scoring function, Eq. (1). The analogs were then rank-ordered according to the estimated IC_{50}^{pre}. Nine analogs with the highest predicted potencies in the low nanomolar range (virtual hits) are shown in Figs. 6 and 7.

An analog whose structure corresponds to the TCL was also included in the designed library. The predicted IC_{50}^{pre} for the TCL was equal to 295 nM, which is significantly higher than the observed activity $IC_{50}^{exp} = 73$ nM. Over 260 analogs displayed activities predicted from the LUDI score lower than 200 nM.

The designed analog with the highest predicted inhibitory potency, TCL-40-9-11 (Fig. 7), contains an aromatic side chain (fragment 40, Table 2) in position 5 (R_1-group) that fills the hydrophobic pocket formed mainly by Tyr 267, Val 274, Tyr 277, Ile 323, Phe 325 and Ile 326 (Fig. 8A). The overlap of the fragment 40 of TCL-40-9-11 with the aromatic ring of the substituent in position 5 of the ligand TCL11 present in the crystal structure of the inhibitor–PfENR complex used for the docking is rather close (Fig. 8B). The longer linker chain connecting the benzene ring of fragment 40 to the phenol moiety of the TCL scaffold finds its own alternative conformation. The substituent (fragment 9) in the 4′ position

(R_2-group) of the TCL-40-9-11 is surrounded by residues Asn 218, Val 222, Met 281 and the substrate binding loop residues Ala 319 and Ile 323. The tertiary amine in the 2′ position (R_3-group) interacts mainly with the NAD^+ co-factor and residues of the binding loop Ser 317, Ala 319, Ala 320 and solvent (Fig. 8A).

2.6. Inhibitor–enzyme interactions

We have analyzed inhibitor–enzyme interactions of the best 9 virtual hits docked to the PfENR in terms of the individual R-groups and scaffold contributions (Table 3). The virtual hits displayed improved total inhibitor–enzyme non-bonding interaction energies (E_{int}) with the PfENR including the NAD^+ co-factor with respect to TCL and the inhibitor TCL11 containing ethylbenzene group in R_1 position ($IC_{50}^{exp} = 76$ nM) [13]. The scaffold of the analogs showed significant contribution to the E_{int} coming from the interaction with the NAD^+ co-factor. The protonated primary amine group 14 (Table 2) in R_2 position of the analog TCL-40-14-10 interacted strongly with the pyrophosphate moiety of the NAD^+. The largest R_1 substituents contributed by up to 1/3 of the total inhibitor–enzyme interaction energy. The following R-groups exhibited
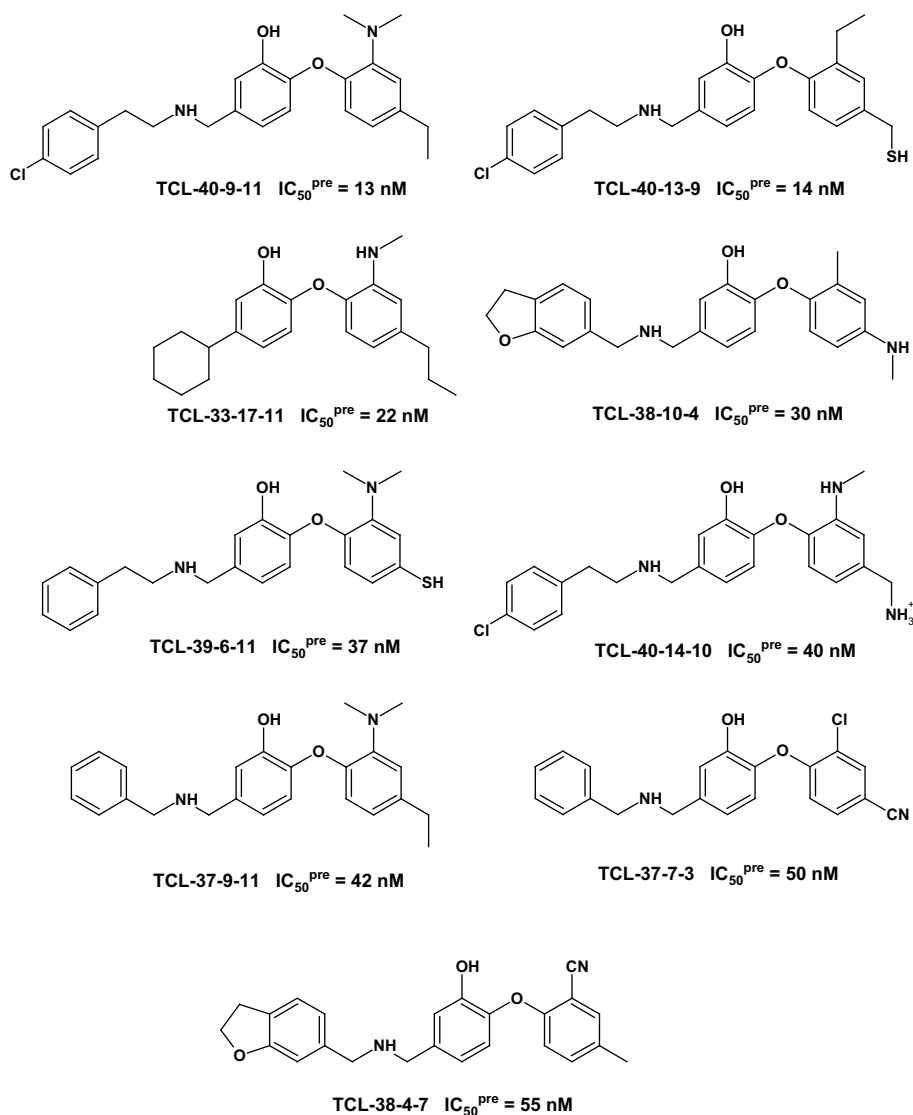
Fig. 7. Chemical structures and predicted inhibitory potencies towards the PfENR of the nine best designed TCL analogs.

the strongest interaction with the PfENR enzyme–co-factor complex: R_1: 38 and 40; R_2: 9 and 13; R_3: 11 and 9.

The virtual hits demonstrated also approximately two times stronger interactions with the residues Tyr 277 and Lys 285, which has been proposed to be important to the catalytic mechanism of the ENR [8], compared to TCL (data not shown). It is namely likely that the hydrogen bond between the inhibitors and Tyr 277 of the ENR prevents this residue from interacting with its natural substrate [26]. The increase in the E_{int} especially with Tyr 277 can be attributed mainly to the bulkier R_1-groups.

2.7. Combinatorial subset selection

Relative frequency of occurrence of the individual R-groups was monitored in the 266 TCL analogs which displayed predicted IC_{50}^{pre} of PfENR inhibition lower than 200 nM (Fig. 9). A clear preference for bulkier fragments in all the R-groups is evident. The fragments that displayed the highest frequency of occurrence were selected to constitute a highly focused combinatorial subset (Fig. 10). This subset displays greatly increased probability to contain potent TCL analogs with inhibition activities towards the PfENR in the low

nanomolar range. The size of the resulting combinatorial subset was narrowed down to only:

$$8(R_1) \times 5(R_2) \times 3(R_3) = 120 \text{ analogs}.$$

which permits its rapid synthesis and testing for the PfENR inhibitory activity.

2.8. ADME-related properties and hit prioritization

Numerous drugs at a late stage of pharmaceutical development and many more lead compounds fail due to adverse pharmacokinetic properties [27]. It is therefore important to incorporate ADME properties' prediction into the lead compound selection.

We have selected a number of relevant molecular properties that help to predict the pharmacokinetic behavior of PfENR inhibitors and prioritize selection of virtual hits for further development. A set of 13 ADME-related properties out of total of 36 descriptors calculated by the QikProp program [28] is given in Table 4. We have used an overall ADME-compliance score – drug-likeness parameter (#stars, Table 4), to assess the pharmacokinetic profiles of the hits. The #stars parameter indicates the number of property descriptors

Fig. 8. A. Virtual hit, analog TCL-40-9-11, shown in yellow color docked to the binding site of the PfENR overlaid with the inhibitor TCL11 (Table 1) shown in green, found in the crystal structure of the ligand–receptor complex [13] (PDB code 2OOS). The phenol rings of the inhibitors interact face-to-face with the nicotinamide ring of NAD$^+$ by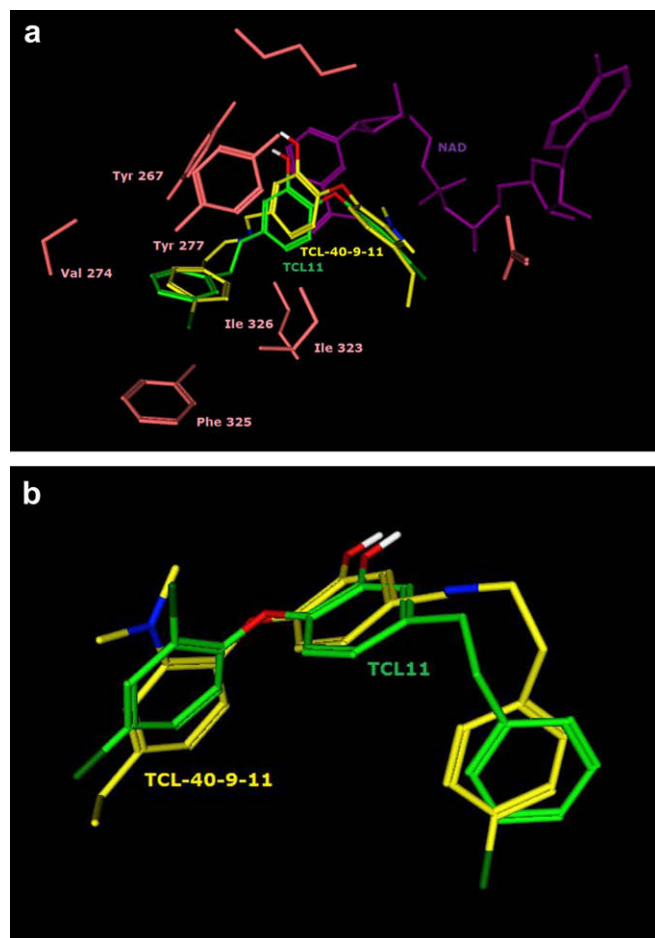 π–π stacking interactions. The co-factor NAD$^+$ is shown in purple. Only heavy atoms and residue side chains are displayed for better clarity. Selected pink colored PfENR residues are marked by residue name and number. B. The overlap of the aromatic rings of the native ligand TCL11 and the virtual hit TCL-40-9-11 docked to the binding site of PfENR is rather close. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

computed by QikProp that fall outside the optimum range of values for 95% of known drugs. The drug-likeness parameter can thus serve as a secondary compound selection criterion.

Eight out of the 9 best virtual hits with the highest predicted PfENR inhibitory potencies are predicted to comply with the pharmacokinetic profiles of drug-like compounds better than TCL (Table 4), which itself is not orally bioavailable [29]. Only one compound (TCL-33-17-11) violates the optimum ranges of the ADME-related properties in 3 descriptors. Therefore, we can expect that some of the virtual hits can be further developed into lead compounds.

3. Conclusions

The presented study yielded a small highly focused virtual combinatorial subset of new TCL analogs, selective FAS-II antimalarials, which contains compounds with predicted inhibitory potencies against the *P. falciparum* ENR in the low nanomolar range. The predicted potencies of this combinatorial subset are lower than the IC$_{50}$ values of the TCL and of the training set compounds. The best virtual hits are drug-like molecules endowed with favorable ADME-related properties.

This study can thus help to direct the attention of synthetic chemists working on the preparation of a next generation of anti-malarial agents towards a particular subset of the chemical space, which is predicted to contain TCL analogs with high PfENR inhibition potencies.

4. Methods

4.1. Virtual library generation

The analog model building was performed with Cerius^2 program [22] using class II consistent force field CFF91 [30] and Rappé and Goddard equilibrated charges [31]. The library of analogs was enumerated by attaching the R-groups (fragments, building blocks) onto the TCL scaffold using the CombiChem module of Cerius^2 program [22]. Chemical reagents considered in this study were taken from the directories of chemicals available from the commercial suppliers of chemicals [21]. Each analog was built either as a neutral molecule or as a molecular ion containing a protonated primary amine group and its molecular geometry was refined by molecular mechanics optimization using smart minimizer with high convergence criteria (energy difference of

Table 3
PfENR–inhibitor interaction energy (E_{int}) and its contributions from the R-groups and the scaffold for 9 best virtual hits (Fig. 7).

E_{int}[a] [kcal mol^{-1}]										
Virtual hit[b] TCL-R$_1$-R$_2$-R$_3$	R$_1$-group ENR + NAD$^+$[c]	NAD$^+$[d]	R$_2$-group ENR + NAD$^+$	NAD$^+$	R$_3$-group ENR + NAD$^+$	NAD$^+$	Scaffold ENR + NAD$^+$	NAD+	Total ENR + NAD$^+$	NAD$^+$
TCL-40-9-11	−24.7	−3.0	−4.1	−0.1	−10.1	−3.3	−29.6	−6.5	−68.5	−12.9
TCL-40-13-9	−24.9	−2.9	−3.9	−0.5	−8.3	−4.3	−30.3	−6.7	−67.5	−14.5
TCL-33-17-11	−16.4	−0.4	−7.5	−0.3	−9.8	−3.0	−31.6	−8.9	−65.3	−12.6
TCL-38-10-4	−26.3	−1.2	−6.2	1.6	−4.8	−2.5	−34.6	−10.3	−72.0	−12.4
TCL-39-6-11	−23.3	−2.2	−1.6	−0.6	−9.8	−2.8	−31.6	−7.6	−66.3	−13.2
TCL-40-14-10[e]	−24.4	−1.6	−2.4	−20.6	−6.1	−1.1	−33.5	−10.3	−66.5	−33.9
TCL-37-9-11	−20.0	−1.2	−4.8	−0.2	−9.6	−2.8	−32.2	−8.9	−66.7	−13.1
TCL-37-7-3	−20.7	−1.2	−4.6	1.1	−5.7	0.5	−33.8	−13.2	−64.9	−12.7
TCL-38-4-7	−25.8	−1.3	−3.2	−0.2	−4.5	1.5	−33.8	−10.2	−67.3	−10.1
TCL1 triclosan	−5.1	1.4	−4.6	1.6	−5.0	1.3	−34.7	−16.7	−49.3	−12.5
TCL11	−19.2	−1.0	−4.8	1.8	−5.5	0.7	−35.0	−15.0	−64.6	−13.4

[a] Enzyme–inhibitor non-bonding interaction energy (1-6-9 potential) as defined in CFF91 force field [30].
[b] For chemical structures of the virtual hits see Fig. 7.
[c] Interaction with the PfENR enzyme including the NAD$^+$ co-factor (Fig. 8A).
[d] Interaction with NAD$^+$ only.
[e] R$_2$-group of TCL-40-14-10 carries net positive charge (Table 2).
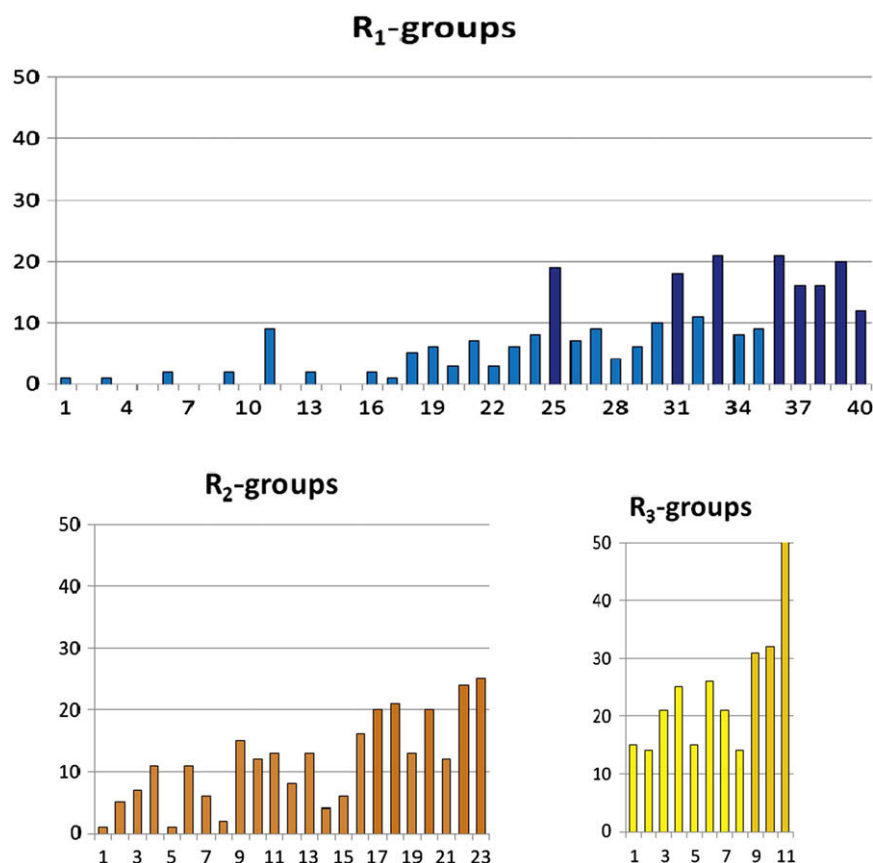
Fig. 9. Frequency of occurrence of individual R-groups (fragments) in the 266 analogs for which IC_{50}^{pre} towards the PfENR $\leq$ 200 nM (for the numbering of fragments see Table 2). Most frequently occurring fragments were selected to form highly focused combinatorial subset of the size: $8(R_1) \times 5(R_2) \times 3(R_3) = 120$ analogs.

10^{-4} kcal mol^{-1}, r.m.s. displacement of 10^{-5} Å) and a dielectric constant of 4.

4.2. Fragment- and analog-based library focusing

Twenty molecular descriptors available in Cerius^2 [22] that characterize a wide spectrum of physicochemical properties related to composition, structure and receptor binding, such as total, polar and hydrophobic molecular surfaces, volume, molecular mass, total charge, dipole moment, sum of atomic polarizabilities, number of hydrogen bond donors and acceptors, partitioning coefficient log $P_{o/w}$, number of rotatable bonds, principal moment of inertia, radius of gyration, desolvation free energy in water and octanol, molar refractivity, topological indices of Balaban, Wiener and Hosoya of the building blocks and molecules of the training set of PfENR inhibitors – analogs of TCL [10,12,13], were computed. For descriptor calculation, the building blocks were considered as real molecules by filling any open valences by hydrogen atoms. Optimum ranges of the properties (descriptors) were defined in terms of upper and lower bounds,

Fig. 10. R-groups (reagents) forming the highly focused combinatorial subset of virtual library of TCL analogs, which is predicted to contain nanomolar PfENR inhibitors.

Table 4
Predicted ADME-related properties of the 9 best virtual hits (Fig. 7) computed by QikProp program [28] used to prioritize the hit selection.

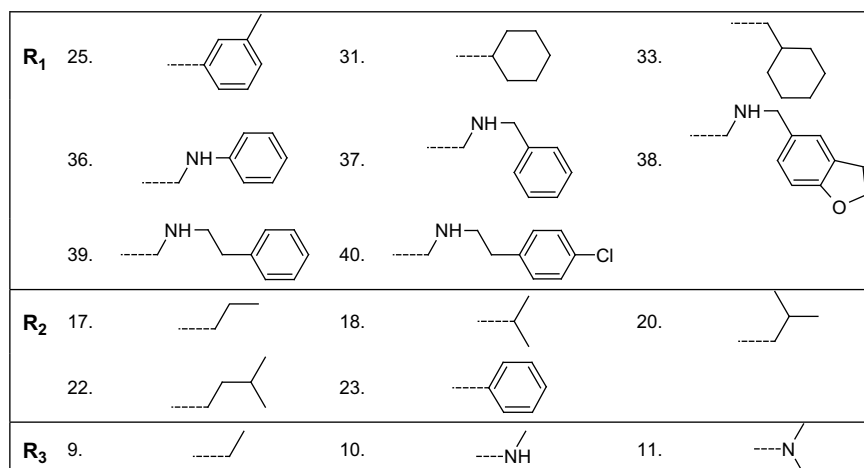| Virtual hits[a]	#stars[b]	M_W[c] [g mol^{-1}]	S_{mol}[d] [Å^2]	$S_{mol,hfob}$[e] [Å^2]	V_{mol}[f] [Å^3]	RotB[g]	HB$_{don}$[h]	HB$_{acc}$[i]	log $P_{o/w}$[j]	log S_{wat}[k] [g dm^{-3}]	log K_{HSA}[l]	log B/B[m]	BIP$_{caco}$[n] [nm s^{-1}]	#metab[o]	IC$_{50}^{pre}$ [p] [nM]
TCL-40-9-11	0	425	750	346	1356	10	2	4	5.7	−5.8	1.1	−0.4	68	8	13
TCL-40-13-9	0	428	727	247	1311	11	3	4	5.6	−5.5	1.0	−0.5	37	8	14
TCL-33-17-11	3	368	686	516	1265	8	1	3	6.3*	−6.7*	1.3*	−0.4	1808	6	22
TCL-38-10-4	0	391	704	345	1262	8	3	5	4.1	−4.4	0.7	−0.4	65	8	30
TCL-39-6-11	0	394	701	219	1265	10	3	4	4.9	−4.7	0.8	−0.4	59	8	37
TCL-40-14-10[q]	0	412	692	211	1269	11	5	5	3.1	−2.3	0.4	−0.6	2	9	40
TCL-37-9-11	0	377	722	318	1286	9	2	4	5.1	−5.1	1.0	−0.3	130	7	42
TCL-37-7-3	0	364	655	77	1147	8	2	4	3.9	−5.2	0.5	−0.8	14	4	50
TCL-38-4-7	0	387	691	282	1232	8	2	5	3.7	−5.2	0.6	−0.7	25	6	55
TCL1 triclosan[r]	2	290	447	0	757	3	1	2*	4.0	−4.3	0.3	0.1	249	1	73[s]

[a] For the chemical structures of the virtual hits see Fig. 7.
[b] Drug likeness – number of property descriptors (from the full list of 36 descriptors of QikProp, ver. 2.0 [28]) that fall outside the range of values for 95% of known drugs are marked by *.
[c] Molecular weight, in Da (range for 95% of drugs: 130–725 Da) [28].
[d] Total solvent-accessible molecular surface, in Å^2 (probe radius 1.4 Å) (range for 95% of drugs: 300–1000 Å^2).
[e] Hydrophobic portion of the solvent-accessible molecular surface, in Å^2 (probe radius 1.4 Å) (range for 95% of drugs: 0–750 Å^2).
[f] Total volume of molecule enclosed by solvent-accessible molecular surface, in Å^3 (probe radius 1.4 Å) (range for 95% of drugs: 500–2000 Å^3).
[g] Number of rotatable bonds (range for 95% of drugs: 0–15).
[h] Number of hydrogen bonds donated by the molecule (range for 95% of drugs: 0–6).
[i] Number of hydrogen bonds accepted by the molecule (range for 95% of drugs: 2–20).
[j] Logarithm of partitioning coefficient between *n*-octanol and water phases (range for 95% of drugs: −2 to 6).
[k] Logarithm of aqueous solubility (range for 95% of drugs: −6.0 to 0.5).
[l] Logarithm of predicted binding constant to human serum albumin (range for 95% of drugs: −1.5 to 1.2).
[m] Logarithm of predicted blood/brain barrier partition coefficient (range for 95% of drugs: −3.0 to 1.0).
[n] Predicted apparent Caco-2 cell membrane permeability in Boehringer–Ingelheim scale, in nm/s (range for 95% of drugs: <5 low, >100 high).
[o] Number of likely metabolic reactions (range for 95% of drugs: 0–15).
[p] Predicted IC$_{50}^{pre}$ were estimated from the QSAR equation (1) and the computed LUDI score.
[q] For the QikProp calculation the neutral (nonprotonated) form of TCL-40-14-10 was considered.
[r] For comparison we list also the ADME-related properties of TCL. The other* of TCL (descriptor that falls outside the range of values for 95% of known drugs) is with a property not included in this table.
[s] Experimental activity of TCL [13].

average values and standard deviations. The penalty–diversity algorithm of the library design module of Cerius^2 [22] was used to select the best and diverse candidates. High penalty scores were then assigned to the considered fragments or analogs whose descriptor values laid outside the optimum ranges to filter those with unsuitable properties. The diversity between the fragments was evaluated via distance based Max–Min function applied to all descriptors used.

4.3. Structure-based library focusing

Crystal structure of the PfENR complexed with one of the most active analogs of triclosan TCL11 (Fig. 1, Table 1) [13] (*Protein Data Bank* [32] entry code 2OOS) was used as the receptor for docking of generated TCL analogs. The shape and size of the ligand binding site of rigid PfENR receptor were defined from the bound native ligand (TCL11) and were mapped onto a 3D energy grid of the size (82 × 77 × 75) Å with a resolution of 0.25 Å using the LigFit module of the Cerius^2 [22]. The binding site model was enlarged by 3 grid layers. Subsequently, the analogs were docked into the binding site model as flexible molecules by generating conformers of each analog via randomization of its dihedral angles (10,000 Monte Carlo steps). The generated conformers were fitted to the site model via flexible fit algorithm by comparing principal moments of inertia of the site and the analog after 50 rigid body minimization steps of 4 ligand orientations [33]. The docking score (ligand–receptor interaction energy) was computed as the non-bonding molecular mechanics energy term using CFF91 force field [30] for each generated conformer after 500 energy minimization steps for the best ligand pose at the receptor binding site. During the docking score calculation a grid representation of rigid PfENR receptor, a cut-off distance of 20 Å applied to the non-bonded interactions and a dielectric constant of 4 were used. Twenty best-fitting conformers were saved and clustered into ten conformational families according to their mutual r.m.s. deviations by means of Jarvis–Patrick complete linkage clustering method [34]. The best representative of each cluster was considered in the virtual screening of the analogs.

4.4. QSAR model

Training set of 16 TCL derivatives – compounds with known inhibitory potencies towards the PfENR [10,12,13], were docked to the binding site model of the PfENR receptor model using the LigFit docking procedure of Cerius^2 [22]. Various scores implemented in the Cerius^2 (such as LigScore, LUDI, PMF, PLP1 and PLP2 [22–25]) were computed and several QSAR models, which relate the IC$_{50}^{exp}$ to the computed scores, were prepared by linear regression analysis encoded in the QSAR module of the Cerius^2 program [22]. The regression equation (1) $-\log_{10}$ IC$_{50} = f$(LUDI), which correlates the computed LUDI score (energy estimate type 2) with the experimental inhibitory potencies towards the PfENR, displayed the highest statistical significance. The predictive power of Eq. (1), which was then used as the target-specific scoring function for the *in silico* screening of the designed virtual library TCL analogs, was verified by applying it to a validation set of 4 similar PfENR inhibitors with known IC$_{50}^{exp}$ values, which were not included into the training set. The ratio of predicted activities IC$_{50}^{pre}$ obtained from the regression equation (1) and observed IC$_{50}^{exp}$ was used to evaluate the performance of the QSAR model.

4.5. In silico screening

The conformer with the highest docking score in each cluster was selected for virtual screening using the LUDI scoring function (energy estimate type 2) [23]. The LUDI function was selected since it performed best in the QSAR model of the PfENR inhibitory potencies of the training set. The LUDI score is a sum of five ligand–receptor interaction contributions derived from ideal hydrogen bonds, perturbed ionic interactions, lipophilic interactions, contributions due to the freezing of internal degrees of freedom and due to the loss of translational and rotational entropy of the ligand [35]. The LUDI score was then used for prediction of PfENR inhibitory potencies (IC_{50}^{pre}) of the focused virtual library of TCL analogs by employing this parameter in a target-specific scoring function. The scoring function, specific for the ENR receptor of *P. falciparum*: $pIC_{50}^{pre}[PfENR] = a + bLUDI$, was parameterized using the QSAR model described above.

4.6. Inhibitor–enzyme interaction energy decomposition

The inhibitor–enzyme interaction energy was computed for fully minimized analog:PfENR:co-factor complexes containing the 9 best virtual hits docked to the active site. Interaction energy was computed as the 1-6-9 non-bonding pairwise interatomic interaction potential including the coulombic, dispersion and repulsion terms, as defined in the CFF91 force field [30]. Original net atomic charges, van der Waals and repulsion atomic parameters of CFF91 were used together with a dielectric constant of 4 to account partially for the dielectric shielding effect within the protein receptor.

4.7. ADME-related properties

Prediction of descriptors related to adsorption, distribution, metabolism and excretion of inhibitors was carried out by the QikProp program [28] based on the method of Jorgensen [36–38]. The program computes pharmacokinetics relevant properties such as octanol/water partitioning coefficient, aqueous solubility, brain/blood partition coefficient, Caco-2 cell permeability, serum protein binding, number of likely metabolic reactions, and others. Drug likeness (#stars) - the number of property descriptors from the full list of 36 descriptors computed by the QikProp, that fall outside the range of values determined for 95% of known drugs, was used as additional compound selection filter.

Acknowledgement

References

[1] Malaria. Fact Sheet No. 94, World Health Organization, Geneva, May 2007 (http://www.who.int/mediacentre/factsheets/fs094/en/index.html).
[2] B. Greenwood, T. Mutabingwa, Nature 415 (2002) 670–672.
[3] S.A. Ralph, G.G. van Dooren, R.F. Waller, M.J. Crawford, M.J. Fraunholz, B.J. Foth, C.J. Tonkin, D.S. Roos, G.I. McFadden, Nat. Rev. Microbiol. 2 (2004) 203–216.
[4] C.O. Rock, J.E. Cronan, Biochim. Biophys. Acta 1302 (1996) 1–16.
[5] A.J. Fulco, Prog. Lipid Res. 22 (1983) 133–160.
[6] N. Surolia, A. Surolia, Nat. Med. 7 (2001) 167–173.
[7] R. McLeod, S.P. Muench, J.B. Rafferty, D.E. Kyle, E.J. Mui, M.J. Kirisits, D.G. Mack, C.W. Roberts, B.U. Samuel, R.E. Lyons, M. Dorris, W.K. Milhous, D.W. Rice, Int. J. Parasitol. 31 (2001) 109–113.
[8] R. Perozzo, M. Kuo, A.b.S. Sidhu, J.T. Valiyaveettil, R. Bittman, W.R. Jacobs Jr., D.A. Fidock, J.C. Sacchettini, J. Biol. Chem. 277 (2002) 13106–13114.
[9] H.N. Bhargava, P.A. Leonard, Am. J. Infect. Control 24 (1996) 209–218.
[10] J.S. Freundlich, J.W. Anderson, D. Sarantakis, H.-M. Shieh, M. Yu, J.-C. Valderramos, E. Lucumi, M. Kuo, W.R. Jacobs Jr., G.A. Schiehser, D.A. Fidock, D.P. Jacobs, J.C. Sacchettini, Bioorg. Med. Chem. Lett. 15 (2005) 5247–5252.
[11] M. Chhibber, G. Kumar, P. Parasuraman, T.N.C. Ramya, N. Surolia, A. Surolia, Bioorg. Med. Chem. 14 (2006) 8086–8098.
[12] J.S. Freundlich, M. Yu, E. Lucumi, M. Kuo, H.-C. Tsai, J.-C. Valderramos, L. Karagyozov, W.R. Jacobs Jr., G.A. Schiehser, D.A. Fidock, D.P. Jacobs, J.C. Sacchettini, Bioorg. Med. Chem. Lett. 16 (2006) 2163–2169.
[13] J.S. Freundlich, F. Wang, H.-C. Tsai, M. Kuo, H.-M. Shieh, J.W. Anderson, L.J. Nkrumah, J.-C. Valderramos, M. Yu, T.R.S. Kumar, S.G. Valderramos, W.R. Jacobs Jr., G.A. Schiehser, D.P. Jacobs, D.A. Fidock, J.C. Sacchettini, J. Biol. Chem. 282 (2007) 25436–25444.
[14] T.J. Sullivan, J.J. Truglio, M.E. Boyne, P. Novichenok, X. Zhang, C.F. Stratton, H.-J. Li, T. Kaur, A. Amin, F. Johnson, R.A. Slayden, C. Kisker, P.J. Tonge, ACS Chem. Biol. 1 (2006) 43–53.
[15] J.L. Lavandera et al., Unpublished Results.
[16] W.A. Warr, J. Chem. Inf. Comput. Sci. 37 (1997) 134–140.
[17] H. Kubinyi, Curr. Opin. Drug Discov. Devel. 1 (1998) 16–27.
[18] V. Frecer, E. Burello, S. Miertus, Bioorg. Med. Chem. 13 (2005) 5492–5501.
[19] V. Frecer, F. Berti, F. Benedetti, S. Miertus, J. Mol. Graph. Model. 27 (2008) 376–387.
[20] T. Rungrotmongkol, V. Frecer, W. De-Eknamkul, S. Hannongbua, S. Miertus, Antivir. Res., in press.
[21] Available Chemicals Directory, Version 95.1, MDL Information Systems, San Leandro, CA.
[22] Cerius2 Life Sciences Software, Version 4.6, Accelrys, San Diego, CA.
[23] H.J. Böhm, J. Comput. Aided Mol. Des. 8 (1994) 243–256.
[24] I. Muegge, Y.C. Martin, J. Med. Chem. 42 (1999) 791–804.
[25] G.M. Verkhivker, K. Appelt, S.T. Freer, J.E. Villafranca, Protein Eng. 8 (1995) 677–691.
[26] G. Nicola, C.A. Smith, E. Lucumi, M.R. Kuo, L. Karagyozov, D.A. Fidock, J.C. Sacchettini, R. Abagyan, Biochem. Biophys. Res. Commun. 358 (2007) 686–691.
[27] F. Darvas, G. Keseru, A. Papp, G. Dorman, L. Urge, P. Krajcsi, Curr. Top. Med. Chem. 2 (2002) 1287–1304.
[28] QikProp, Version 2.0, Schrödinger, Inc., New York, NY.
[29] F. Maestrelli, P. Mura, M.J. Alonso, J. Microencapsul. 21 (2004) 857–864.
[30] J.R. Maple, M.J. Hwang, T.P. Stockfish, U. Dinur, M. Waldman, C.S. Ewing, A.T. Hagler, J. Comput. Chem. 15 (1994) 162–182.
[31] A.K. Rappé, W.A. Goddard III, J. Phys. Chem. 95 (1991) 3358–3363.
[32] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucleic Acids Res. 28 (2000) 235–242.http://www.rcsb.org.
[33] K.P. Peters, J. Fauck, C. Frommel, J. Mol. Biol. 256 (1996) 201–213.
[34] P. Willett, in: P.M. Dean (Ed.), Molecular Similarity in Drug Design, Chapman and Hall, Glasgow, 1994, pp. 110–137.
[35] S. Fischer, J.C. Smith, C. Verma, J. Phys. Chem. B 105 (2001) 8050–8055.
[36] E.M. Duffy, W.L. Jorgensen, J. Am. Chem. Soc. 122 (2000) 2878–2888.
[37] W.L. Jorgensen, E.M. Duffy, Bioorg. Med. Chem. Lett. 10 (2000) 1155–1158.
[38] W.L. Jorgensen, E.M. Duffy, Adv. Drug Deliv. Rev. 54 (2002) 355–366.