

REPUBLIQUE DU CAMEROUN

Paix – Travail – Patrie

UNIVERSITE DE YAOUNDE I
FACULTE DES SCIENCES
DEPARTEMENT DE PHYSIQUE

CENTRE DE RECHERCHE ET DE
FORMATION DOCTORALE EN
SCIENCES, TECHNOLOGIE ET
GEOSCIENCES



REPUBLIC OF CAMEROUN

Peace – Work – Fatherland

UNIVERSITY OF YAOUNDE I
FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

POSTGRADUATE SCHOOL OF
SCIENCE, TECHNOLOGY □
GEOSCIENCES
LABORATORY OF NUCLEAR,
ATOMIC, MOLECULAR PHYSICS
AND BIOPHYSICS

**COMPUTER-AIDED DRUG DESIGN OF INHIBITORS OF
FALCIPAIN-3 (FP-3) OF PLASMODIUM FALCIPARUM,
WITH FAVORABLE PHARMACOKINETIC PROFILES.**

Thesis submitted in partial requirements for the award of a
Doctorat/PhD in Physics

Par : **BEKONO Boris Davy**
Master in Physics

Sous la direction de
MEGNASSAN Eugène E. Edjèm
Professor,
University Nangui Abrogoua, Côte d'Ivoire
Owono Owono Luc Calvin
Professor,
University of Yaoundé I, Cameroon

Année Académique : 2020





DEPARTEMENT DE PHYSIQUE
DEPARTMENT OF PHYSICS

**ATTESTATION DE CORRECTION DE LA
THESE DE DOCTORAT/PhD**

Nous, Professeur SAIDOU et Professeur ETOA François Xavier, respectivement Examineur et Président du jury de la Thèse de Doctorat/Ph.D de Monsieur BEKONO Boris Davy, Matricule 09W1303, préparée sous la direction des Professeurs OWONO OWONO Luc Calvin et MEGNASSAN Eugène, intitulée : «COMPUTER-AIDED DRUG DESIGN OF INHIBITORS OF FALCIPAIN-3 (FP-3) OF PLASMODIUM FALCIPARUM, WITH FAVORABLE PHARMACOKINETIC PROFILES», soutenue le Lundi 13 Juillet 2020, en vue de l'obtention du grade de Docteur/Ph.D en Physique, Spécialité Physique Nucléaire, Atomique, Moléculaire, et Biophysique Option Physique Moléculaire et Biophysique, attestons que toutes les corrections demandées par le jury de soutenance ont été effectuées.

En foi de quoi, la présente attestation lui est délivrée pour servir et valoir ce que de droit.

Fait à Yaoundé, le ...2...1...JUIL...2020...

Examineur

SAIDOU
Maître de Conférences

Le Président du jury

ETOA François Xavier
Professeur



Le Chef de Département de Physique

Nadia Jean-Marie
Bienvenu
Professeur

Dedication

This work is dedicate to

- my Mother Ondoua martine

- my beloved wife NGONO Estelle Constante

- my children Bekono Théandra and Bekono Jayden Daniel.

Acknowledgments

I am deeply indebted to the Almighty God for giving me life, strength and insight to be able to carry out this heavy scientific task.

I would like to express my sincere gratitude to my PhD supervisors: Prof. Luc Calvin OWONO OWONO, of the Laboratory for Simulations and Biomolecular Physics, Department of Physics, Higher Teacher Training College (Ecole Normale Supérieure) and Prof. Eugène MEGNASSAN ETCHRI EDJEM, Laboratory of Fundamental and Applied Biophysics, University of Nangui Abrogoua, Abidjan, Côte d'Ivoire. I am heavily indebted to you. Sirs, you both invested your relentless efforts in order to ensure that the basic concepts in computer-aided drug design today were well mastered and that the theme of this thesis was properly defined and understood. Kindly find here the expression of my profound gratitude.

A special thank to Doctor NTIE KANG Fidèle of the University of Buéa I would say that without his tireless commitment towards the realization of this work, the present thesis could not have been completed. He has provided to me computer programs, resources and much sound counsels that were very essential to the success of this work. He did a lot of works to obtain the fellowship which allowed me to finalize this thesis. Thank you Sir, you are a model for me.

I would thank all the members (teachers and PhD students) of the Laboratory of Fundamental and Applied Biophysics, University of Nangui Abrogoua of Nangui Abrogoua led by Prof Megnassan where a major part of the simulations was carried out. A special thank to Doctor Esmel Akori Elvice for his assistance in different tasks of the simulation during my stay in Côte d'Ivoire.

I also express my gratitude to the academic staff of the Department of Physics, The University of Yaounde I. I highly appreciate the quality of your teaching and scientific guidance.

I am tremendously indebted to all the members of the jury of the public defence of this thesis. Dear professors, I highly appreciate the fact that despite your tight schedules, you could find time to evaluate the scientific quality of the work presented in this thesis. THANK YOU Sirs.

I receive many encouragements. In this regard, I would thank a great deal the following persons: Prof. Moïse KWATO NJOCK, Coordinator of the Centre for Atomic Molecular Physics and Quan-

tum Optics (CEPAMOQ), University of Douala, Prof OWONO ANGUE, Department of Physics, Advanced Teachers' Training College (Ecole Normale Supérieure), The University of Yaounde I, Prof. SAÏDOU, Department of Physics, The University of Yaounde I, Prof. Lenta Bruno, Department of Chemistry, Higher Teacher Training College (Ecole Normale Supérieure), University of Yaoundé I, Prof. BOYOMO ONANA Marthe, Department of Physics, National Advanced School of Engineering (Ecole Nationale Supérieure Polytechnique), University of Yaoundé I and Dr. Brice DALI LOKOU, University of Abobo-Adjamé, Côte d'Ivoire.

To the students of the Department of Physics, University of Yaoundé I, I am very pleased to have you as my classmates.

I received financial and moral supports from the following persons who have always had a dream to see this thesis completed and defended. Without being exhaustive, I wish to heartily thank my parents (Mama Ondoua Martine and Prof OWONO ANGUE); my sisters; my brothers; You all deserve a big THANK YOU.

Computational resources used for this work were generously provided by the Laboratory of Fundamental and Applied Biophysics, University of Nangui Abougoar, Abidjan, Côte d'Ivoire, the Laboratory for Simulations and Biomolecular Physics, Department of Physics, Higher Teacher Training College, The University of Yaounde I.

I will remain eternally grateful to the following organisations for providing grants that allowed to achieve this work.

The Africa German Network Excellence in Science (AGNES), for granting a Mobility Grant in 2017; the Grant is generously sponsored by German Federal Ministry of Education and Research and supported by the Alexander von Humboldt Foundation, for a 2 months fellowship at the Laboratory of Fundamental and Applied Biophysics, University of Nangui Abougoar, Abidjan, Côte d'Ivoire, headed by Prof. Megnassan.

I express my deepest gratitude to all those who in one way or another, contributed to the achievement of this work.

Résumé

Les rapports successifs de l’OMS sur le paludisme, malgré leur caractère alarmant sur l’émergence d’une résistance du *Plasmodium falciparum* (*Pf*), la forme la plus virulente du parasite, mettent sous silence un danger majeur. En effet, depuis la mise en place de la thérapie combinée à l’artémisine (TCA) en 2006, des cas de résistance du *Pf* ont été enregistrés comme l’indiquent les publications scientifiques. Il apparaît ainsi que le traitement du paludisme dans le monde ne tient qu’à un fil, le TCA. Une résistance massive à cette combinaison en fera une pathologie sans traitement. Dans ce travail qui a pour titre “ Computer-aided drug design of inhibitors of falcipain-3 (FP-3) of *Plasmodium falciparum*, with favorable pharmacokinetic profiles”, notre but consiste à concevoir *in silico* des molécules biologiquement actives contre l’agent pathogène du paludisme, le *Pf*, qui par la suite pourront être développées en médicaments contre le paludisme. Pour ce faire, nous avons ciblé une enzyme vitale dans le processus de maturation du germe à savoir, la falcipain-3 (FP-3) qui est une protéase à cystéine impliquée en dernier dans la dégradation de l’hémoglobine dans la vacuole digestive de *Pf*. Pour conduire nos travaux, nous avons utilisé les techniques modernes de conception assistée par ordinateur de molécules à activité biologique. Il s’agit de la complexation et de la génération de pharmacophores. Toutes ces techniques ont pour objectif de déterminer l’énergie d’interaction entre le ligand (molécule bioactive) et le récepteur (enzyme cible de l’agent pathogène) à l’aide du schéma de Mécanique Moléculaire-Poisson Boltzmann (MM-PB). A partir d’un certain nombre de molécules (peptidomimétiques) dont l’activité a été testée contre l’enzyme cible, des modèles quantitatifs structure-activité (QSAR) ont été bâtis et utilisés pour prédire l’activité de nouveaux analogues (*de novo design*). L’analyse pharmacocinétique de ces nouvelles molécules s’est faite par le calcul des descripteurs moléculaires qui sont liés à l’absorption, la distribution, le métabolisme, l’élimination et la toxicité des molécules médicamenteuses. Ces nouvelles molécules serviront alors de squelette de base pour le développement de nouveaux médicaments contre le paludisme. Les composés les plus intéressants feront l’objet de synthèses chimiques et de tests biologiques.

Abstract

Malaria, a widespread disease caused by *Plasmodium falciparum* (*Pf*), has been declared a public health concern in many developing countries by the World Health Organization (WHO). Indeed since the implementation of artemisinin-combined therapy (ACT) in 2006, *Pf*-resistance cases were recorded as indicated by the scientific literature. Currently, the treatment of malaria depends mainly on ACT. Massive resistance to this combination may quickly render malaria into a pathology without treatment. In the current work entitled "Computer-aided drug design of inhibitors of falcipain-3 (FP-3) of *Plasmodium falciparum*, with favorable pharmacokinetic profiles", our aim was to discover new antimalarial compounds, which could be further developed into next generation antimalarials drugs, by using *in silico* methodologies. We have targeted falcipain-3 (FP3) which is a cysteine protease of *Pf* involved later in the process of hemoglobin degradation in the food vacuole of the parasite. To carry out our work, we used modern techniques of computer-aided drug design of molecules with biological activity. It is the complexation and the generation of pharmacophore. All these techniques aim to determine the interaction energy between the ligand (bioactive molecule) and the receptor (target enzyme of the pathogen) using the Molecular Mechanics-Poisson Boltzmann (MM-PB) scheme. From a number of molecules (peptidomimetics) whose activity has been tested against the target enzyme, quantitative structure-activity models (QSARs) have been constructed and used to predict the activity of novel analogues (*de novo design*). Pharmacokinetics analysis was done by the calculation of molecular descriptors, which are often used to predict the absorption, distribution, metabolism, excretion and toxicity (ADMET) of drug candidates. The best predicted new molecules will then serve as a scaffold for the development of new drugs against malaria. The most promising compounds discovered have been proposed for synthesis and biological tests.

List of Abbreviations and Acronyms

ACT: Artemisinin combination therapy

ADME-tox: Absorption, Distribution, Metabolism, Excretion and Toxicity

AIDS: Acquired Immune-deficiency Syndrome

AMBER: Assisted Model Building with Energy Refinement

CADD: Computer-aided drug discovery

CADDD: Computer-aided Drug Design/Discovery

CFF: Consistent Force Field

CG: Conjugate gradient

CHARMM: Chemistry at HARvard Molecular Mechanics

CVFF: Consistent-Valence Force Field

CQ: chloroquine

DS 2.5: Discovery Studio 2.5

FP: falcipain

FDM: Finite Difference method

HBA: Hydrogen bond acceptors

Hb: Hemoglobin

HBD: Hydrogen bond donors

HIV: human immunodeficiency virus

MD: molecular dynamics

MM: molecular mechanics

MMFF: Merck Molecular Force Field

MM-PB Molecular Mechanics with Poisson-Boltzmann

NP: natural product

PB: Poisson-Boltzmann

PDB: Protein databank

QM: quantum mechanics

QSAR: quantitative structure-activity relationship

RMSD: Root mean square deviation

SB-CADDD: Structure-based computer-aided drug design

SBD: structure-based drug design

SBVS: structure-based virtual screening

VL: virtual library

VS: virtual screening

WHO: World Health Organization

List of Symbols

Å: Angström ($1 \text{ \AA} = 10^{-10}$ meters)

T: Absolute temperature ($T = 298.13 \text{ K}$, at equilibrium conditions)

K_b : Boltzmann constant ($k = 1.380658 \times 10^{-23} \text{ J K}^{-1}$)

R: The ideal gas constant $8.31 \text{ J mol}^{-1} \text{ K}^{-1}$

J: Joule

IC_{50} : Concentration of a competitive tight binding inhibitor that causes 50% reduction of the rate of catalytic substrate conversion

K_i : Inhibitory constant

kcal mol^{-1} : kilo calories per mole ($1 \text{ kcal mol}^{-1} = 4186.8 \text{ J}$)

μM : Micromolar concentration ($1 \mu\text{M} = 10^{-6} \text{ mol dm}^{-3}$)

M_w : Molar weight

nM : Nanomolar concentration ($1 \text{ nM} = 10^{-9} \text{ mol dm}^{-3}$)

ns: nanosecond ($1 \text{ ns} = 10^{-9} \text{ s}$)

ps: picosecond ($1 \text{ ps} = 10^{-12} \text{ s}$)

ϵ_0 : Permittivity of a vacuum = $8.85 \times 10^{-12} \text{ AsV}^{-1} \text{ m}$

\vec{r} : Position vector

List of Figures

| | | |
|------|--|----|
| 1.1 | World distribution of malaria in 2016 [36]. | 6 |
| 1.2 | <i>Natural products and synthetic molecules currently used as antimalarial drugs</i> | 10 |
| 1.3 | <i>Geographical distribution of mortality due to artemisinin and artemisinin-combination therapy resistance in malaria-endemic areas, year 2014</i> [76]. | 11 |
| 1.4 | <i>The malaria parasite life cycle</i> [88]. | 13 |
| 1.5 | <i>The pharmaceutical drug discovery pipeline today</i> [108]. | 16 |
| 1.6 | <i>An example of workflow for in silico drug design and discovery</i> [109]. | 17 |
| 1.7 | <i>Chemical structures of example of marketed drugs involving use of structure-based-drug design</i> [114-118]. | 18 |
| 1.8 | <i>Summary of the degradation process of hemoglobin in the <i>P. falciparum</i> food vacuole</i> [129]. | 20 |
| 1.9 | <i>Schematic reaction mechanism of the cysteine protease-mediated cleavage of a peptide bond</i> [129]. | 21 |
| 1.10 | <i>Domains of falcipains</i> The active site residues (Cys, His, Asn) are conserved within the papain family, but falcipains have a unique N-terminal extension, which acts as a refolding domain, and a C-terminal inserted as a hemoglobin (Hb) binding domain. The prodomain has the ERFNIN and GNFD motifs, which are conserved in falcipains and papain[135]. | 22 |
| 1.11 | <i>The equilibrium of a drug being bound and unbound to its target.</i> | 23 |
| 1.12 | <i>Non-peptidic structures of some FP-2/3 inhibitors identified from the ZINC database and using molecular dynamics simulation.</i> | 24 |
| 1.13 | <i>Non-peptidic structures of some FP-2 inhibitors identified using virtual screening.</i> | 25 |
| 1.14 | <i>Active inhibitors of FP-2/3 identified from the SPECS database by SBVS.</i> | 26 |

| | | |
|------|---|----|
| 1.15 | <i>Some natural product-based inhibitors of FP-2/3.</i> | 26 |
| 1.16 | <i>The Binding pocket of chain A of FP-3 crystal structure (PDB code 3BWK, resolution 2.42 Å). K11017 (color in yellow) in stick representation, protein atoms in line representation (carbon atoms in their usual colours). This diagram was produced using the Discovery studio software.</i> | 31 |
| 1.17 | <i>3D Binding pocket of chain A of FP-3 crystal structure (PDB code 3BWK, resolution 2.42 Å).</i> | 31 |
| 1.18 | <i>Schematic view of interactions.</i> | 34 |
| 1.19 | <i>Elongation between two atoms..The Morse potential (blue) and harmonic oscillator potential (green). Unlike the energy levels of the harmonic oscillator potential, which are usually spaced by \hbar, the Morse potential level spacing decreases as the energy approaches the dissociation energy. The dissociation energy D_e is larger than the true energy required for dissociation D_0.</i> | 35 |
| 1.20 | <i>Electron cloud represented by the fixed-charge model and polarization models [180]</i> | 39 |
| 1.21 | <i>Representation of the energy of all possible conformers of a molecular system. The energy of the system is calculated using the forcefield. The conformation of the system is altered to find lower energy conformations through a process called minimization.</i> | 45 |
| 1.22 | <i>Steepest descent minimization [213]</i> | 46 |
| 2.1 | <i>Novel PEP analogues design methodology workflow.</i> | 49 |
| 2.2 | <i>The chemical structures of training (PEPx) and validation (VEPx) sets of inhibitors obtained from literature [221].</i> | 51 |
| 2.3 | <i>Forces acting on atoms of the molecule.</i> | 54 |
| 2.4 | <i>Steps of the molecular dynamics simulation procedure.</i> | 56 |
| 2.5 | <i>Desolvation of a drug and its target active site prior to binding.</i> | 57 |
| 2.6 | <i>The thermodynamic cycle of protein-ligand interactions.</i> | 58 |
| 2.7 | <i>Implicit solvation model</i> | 59 |
| 2.8 | <i>Cubic in finite difference method. A: Fragmentation of the medium in identical cubic cells of side ℓ (in red). B: Zoom on one of the cubic cells where the PB equation is discretized.</i> | 63 |
| 3.1 | <i>200 snapshots obtained during the molecular dynamics simulation of the most active ligand.</i> | 75 |

| | | |
|------|---|----|
| 3.2 | <i>A plot of the correlation equation between pIC_{50}^{exp} and relative complexation Gibbs free energies of the 8 training set of FP-3 inhibitors $\Delta\Delta G_{comp}$</i> | 80 |
| 3.3 | <i>Internal validation of QSAR model and calculation of cross validation coefficient (). Outlier (PEP 1) is highlighted in blue.</i> | 81 |
| 3.4 | <i>Enzyme-ligand interactions at the active site of FP-3 depicted in 3D for the most active PEP2.</i> | 85 |
| 3.5 | <i>2D schematic interaction diagram of the most potent inhibitor PEP2 at the active site of PfFP3.</i> | 86 |
| 3.6 | <i>Connolly surface of the active site of PfFP-3 with the most potent inhibitor PEP2. The binding site surface is colored according to residue hydrophobicity: red -hydrophobic, blue - hydrophilic and white – intermediate.</i> | 87 |
| 3.7 | <i>Breakdown of FP3-PEP interaction energy into each active site residue contribution at S2 pocket.</i> | 88 |
| 3.8 | <i>Plot of correlation between pIC_{50}^{exp} and P2, His 183 contribution to the intermolecular interaction energy of PfFP3-PEPs.</i> | 89 |
| 3.9 | <i>Breakdown of FP3-PEP interaction energy into each active site residue contribution at S1' pocket.</i> | 90 |
| 3.10 | <i>Bound conformations of inhibitors taken from E:I complexes</i> | 90 |
| 3.11 | <i>Coordinates (A,B) of the Hypo1 pharmacophore of FP-3 inhibition, (C) features of the pharmacophore of FP-3 inhibition, (D) pharmacophore mapping with PEP2 (with blue carbons) and PEP7 (with yellow carbons). Color legend of features: hydrophobic (blue), acceptor (green), donor (purple).</i> | 91 |
| 3.12 | <i>Plot of estimated and experimental activity for PH4.</i> | 92 |
| 3.13 | <i>R-groups (fragments, building blocks, substituents) used in the design of the initial diversity library of PEP analogues. Dashed bonds indicate the attachment points of the fragments.</i> | 94 |
| 3.14 | <i>Chemical structures towards Pf FP-3 of seven most potent PEP analogues.</i> | 98 |
| 3.15 | <i>The best virtual hit, analogue PEP-2-1-30-7-2 (with purple carbons), the inhibitor PEP7 (Figure 1) shown in yellow carbons, mapped a PH4.</i> | 98 |
| 3.16 | <i>2D schematic interaction diagram of the most potent analogue design PEP-2-1-30-7-2 at the active site of PfFP3.</i> | 99 |

| | | |
|------|--|-----|
| 3.17 | <i>Connolly surface of the active site of PffP-3 with the predicted most active PEP inhibitor PEP-2-1-30-7-2. The binding site surface is colored according to residue hydrophobicity: red - hydrophobic, blue - hydrophilic and white-intermediate.</i> | 100 |
| 3.18 | <i>2D interaction diagram of PEP-2-1-1-16-2 hydrogen bonds (HBs) (black, blue and green dashed line)</i> | 102 |
| 3.19 | <i>2D interaction diagram of PEP-2-1-12-21-30 hydrogen bonds (HBs) (black, blue and green dashed line).</i> | 102 |
| 3.20 | <i>2D interaction diagram of PEP-2-1-30-7-2 hydrogen bonds (HBs) (black, blue and green dashed line).</i> | 103 |
| 3.21 | <i>2D interaction diagram of PEP-2-6-12-21-30 hydrogen bonds (HBs) (black, blue and green dashed line).</i> | 103 |
| 3.22 | <i>2D interaction diagram of PEP-24-1-12-1-21 hydrogen bonds (HBs) (black, blue and green dashed line).</i> | 104 |
| 3.23 | <i>2D interaction diagram of PEP-3-1-12-1-13 hydrogen bonds (HBs) (black, blue and green dashed line).</i> | 104 |
| 3.24 | <i>2D interaction diagram of PEP-31-1-12-13-1 hydrogen bonds (HBs) (black, blue and green dashed line).</i> | 105 |

List of Tables

| | | |
|------|---|----|
| 1.1 | <i>Example of marketed drugs involving use of Structure-based drug design.</i> | 18 |
| 1.2 | <i>Description of the 5 Available Crystal Structures of FP-2 from the PDB.</i> | 28 |
| 1.3 | <i>Description of the 2 Available Crystal Structures of FP-3 from the PDB.</i> | 30 |
| 2.1 | <i>Training and validation sets of PEP inhibitors obtained from the literature.</i> | 51 |
| 3.1 | <i>In vitro Biological Activities of data set of PEP inhibitors obtained from the literature.</i> | 73 |
| 3.2 | <i>Results of Interaction Energies, Solvation Enthalpies and Entropic Terms of the Ligands in Complex with FP-3.</i> | 74 |
| 3.3 | <i>Results of Interaction Energies, Solvation Enthalpies and Entropic Terms of the Free Ligands.</i> | 76 |
| 3.4 | <i>Results of Binding Free Energies of Complexation of the Ligands in Complex with FP-3.</i> | 78 |
| 3.5 | <i>Comparison of Experimentally Measured and Theoretically Predicted Activities of the Training set.</i> | 82 |
| 3.6 | <i>Statistical information of correlation between computed $\Delta\Delta G_{comp}$ of training set inhibitors and their experimental activities IC_{50}^{exp} against FP-3.</i> | 83 |
| 3.7 | <i>Training and Validation Sets of PEP derivatives used in the QSAR Model of FP-3 Inhibition and in the Validated Complexation Model[221].</i> | 84 |
| 3.8 | <i>Interaction energy between some active site residue of FP-3 (S2) and PEPx (P2).</i> | 85 |
| 3.9 | <i>Interaction energy between P1 chains and HIS183 residues.</i> | 86 |
| 3.10 | <i>Output parameters of the 10 generated PH4 pharmacophoric hypotheses for FP-3 inhibitors after CatScramble validation procedure.</i> | 88 |

| | | |
|------|--|-----|
| 3.11 | Computed the ratio of predicted and observed activities ($\frac{pIC_{50}^{pred}}{pIC_{50}^{exp}}$) for the training set. | 92 |
| 3.12 | Statistical data on regression analysis of correlation for the training set between PH4 predicted activity (pIC_{50}^{exp} vs pIC_{50}^{pred}) and experimental one (pIC_{50}^{exp}) against FP-3. . . . | 92 |
| 3.13 | Molar masses, complexation Gibbs free energies and their components for the top 80 scoring virtually designed analogues. | 94 |
| 3.14 | Table 3.13 continued. | 95 |
| 3.15 | Table 3.13 continued. | 96 |
| 3.16 | Table 3.13 continued. | 97 |
| 3.17 | Predicted ADME-related properties of the best-designed PEP analogues. | 101 |

General Introduction and Objectives

Preamble

Diseases are nowadays one of the major problems facing world populations in general and those in sub-Saharan Africa in particular [1, 2]. To cope with diseases, more often chemical entities are used. Thus, several chemical entities have been discovered as drugs and are grouped into different categories based on their actions on the disease. Among them are antibiotics and anti-inflammatories [3-5]. Unfortunately, these drugs often date back many years, and as a result, resistance has been developed by pathogens that cause diseases. This is currently a major public health concern that needs to be taken into consideration by both governments and researchers [6]. The most promising approach to fighting with these resistances is the more than urgent development of new therapeutic formulas. These new drugs must interact with well-known target proteins as well as the search for new targets that might lead to new successful therapies [7, 8].

In addition to drug resistance problems currently used to control some of the major diseases affecting the African populations, such as Ebola, tuberculosis, malaria, trypanosomiasis, schistosomiasis, onchocerciasis, and HIV/AIDS, exist the problem of the high cost of research and development of new drugs in the process of overcoming the problem of resistance to the use of current drugs. The new drugs must be selective and less toxic, which requires huge investments to design and develop them into marketable drugs [9-11]. However, in due of the poverty of the African continent in general, the pharmaceutical industries have oriented their research on diseases that affect the "rich" countries and do not see the need to invest in the development of drugs against diseases that affect the poorest countries in the world [12]. The direct consequence is that these diseases have become endemic on the continent, and are known as neglected diseases, neglected tropical diseases or simply poverty-related diseases [13, 14]. This, therefore, suggests the urgent need to incorporate low-cost techniques or methods into the process of drug discovery for these diseases [15, 16]. Computer-aided drug design (CADD) also known as *in silico* methods by molecular modeling (MM) and combinatorial chemistry (CC) methods, for lead discovery and lead optimization represents one of such strategies [17, 18]. This makes use of bioinformatics [19] and cheminformatics tools [20] in goal to respectively identify potential drug targets against parasitic diseases and propose new molecules which interact favorably

against these targets and which can be developed into drugs. This involves docking, complexation methods, the virtual screening (VS) of huge compound virtual library or databases against validated drug targets and evaluating the interactions between the potential binders and the receptor active site by the use of mathematical scoring functions methods, followed by the careful selection of virtual hit compounds to be screened by biological assays. The evaluation of the effectiveness of these methods often rests on the ability of the scoring method to clearly discriminate the active compounds from inactive ones. When this objective is achieved, the number of compounds to undergo biological tests is considerably narrowed down and hence the cost of discovery of a drug drastically reduced [21].

The rationale of the work

Malaria, a widespread disease caused by *Plasmodium falciparum* (*Pf*), has been declared a public health concern in many developing countries by the World Health Organization (WHO). Indeed since the implementation of artemisinin-combined therapy (ACT) in 2006, *Pf*-resistance cases were recorded as indicates the scientific literature [22-25]. Currently, the treatment of malaria in the depends mainly on the ACT. Massive resistance to this combination may quickly render malaria into a pathology without treatment. The above observations strongly suggest the need for industry-academia partnerships for the search of new antimalarials with alternative modes of action. In the search for new drugs against malaria, two strategic approaches are distinguished; one based directly on the parasite, in which only two known antimalarials are known, viz quinine and artemisinin. The second approach engages the inhibition of a therapeutic target of the parasite. Unfortunately, for the first approach, for example, the mechanisms of action of quinine and artemisinin are not known so far with precision. The second approach, which involves the indirect inhibition of the parasite by the inhibition of its vital processes of life requires designing new molecules which bind specifically to known drug targets in the parasite. The identification of therapeutic targets against *Pf* has increased tremendously over the past two decades [22-25]. This favors the use of the second approach in the rational drug discovery and design. The work presented here is part of this second approach. His aims is to use CADD methods (structure-based and ligand-based *in silico* methods) to design and screen new antimalarials against cysteine protease falcipain-3 (FP-3) one of the enzymes most involved in the degradation of hemoglobin and which is totally different of the cysteine protease of the human [26, 27] and multidrug resistant strains, in order to cut down the cost required to design and develop next generation antimalarials drugs. It wants to benefit from the strong scientific production that ultimately allowed designing inhibitors of proteases of *Pf* in the image of those of HIV [28, 29]. In addition to identifying the structural information at the level of each pocket of the active site of the cysteine protease FP-3. Thus, the pathway may be open to the easy design of FP-3 inhibitors which have the potential to be developed as drugs against malaria, particularly against resistant forms. This will reduce the costs of designing and developing new generation antimalarial drugs.

Goals and Objectives

Our main aim will be to design and screen novel antimalarial compounds, which are likely to be lead for the discovery of drugs against malaria, particularly against drug-resistant malaria, by computer-aided methodologies.

Specific goals will include :

1. Elaboration of a model of 3D-quantitative structure activity relationship (3D-QSAR) :

From a series of Inhibitors (I_x) with biological activities (IC_{50}^{exp}) that have been determined experimentally against FP-3, we will construct the FP-3: I_x complexes by *in situ* modification of the X-ray crystal structure of FP-3 in complex with K11017 (PDB entry code: 3BWK [30]). We will calculate the Gibbs free energies of ligand-receptor complex formation $\Delta\Delta G_{comp}$ according to the *in silico* approach of Molecular Mechanics Poisson-Boltzmann (MM-PB) in order to establish a correlation $pIC_{50}^{exp} = -\log_{10} IC_{50}^{exp} = a \times \Delta\Delta G_{comp} + b$ (*) which explains the variation of (IC_{50}^{exp}) by that of ΔG . This task will require a relevant use of the thermodynamics of Enzyme-ligand interactions.

2. Elaboration of an inhibition pharmacophore from 3D-QSAR :

The QSAR model will produce the inhibition conformation in the active site of FP-3. From this conformation, the pharmacophore (PH4) model will be constructed. This model will reveal the relative positions of the active centers (hydrophobic, acceptor, hydrogen bond donor) required to inhibition.

3. Building virtual library (VL) containing 3D structures analogues of I_x :

The skeleton of the I_x will allow substitutions that, once combined, populate a part of the chemical space, reaching several million compounds. Filtering based on the pharmacokinetic and bioavailability properties of absorption, distribution, metabolism, excretion and toxicity (ADMET) will allow the size of this VL to be reduced to a few hundred thousand compounds.

4. Screening the virtual library (VL) and identifying the «leads» :

The filtered VL will be screened with the inhibition PH4. The best leads will be evaluated by complexation and their predictive inhibitory activity (pIC_{50}^{pre}) will be deduced after calculating ΔG and using the previously established QSAR (*) equation. The best analogues will then be proposed for chemical synthesis and biological evaluation.

Results expected

We will be able to develop a QSAR predictive model for FP-3 inhibiting peptidomimetic derivatives, from the crystal structure of the enzyme (3BWK) in complex with K11017 by the use of the complexation model. The conformations of the inhibitors and structural information from the previously

established QSAR model will be used in the development of the PH4 model and the construction of the VL. Filtering based on the pharmacokinetic and bioavailability properties of absorption, distribution, metabolism, excretion, and toxicity (ADMET) will allow the size of this VL to be reduced to a smaller number of peptidomimetic analogues. These peptidomimetic analogues retained will be screened with the inhibition pharmacophore (PH4), but just few of these analogues will perfectly map the PH4. The 3D-QSAR model will be used to predict their activities, and will led to the discovery of a smaller subset of inhibitors that will be declared provisionally active against FP-3, with activities are in the nano molar range. From the point of view of ADMET, these inhibitors should have a profile that allows them to be developed as potential drugs in the fight against resistant forms of malaria and will be proposed for synthesis and biological evaluation. It should be recalled that the present work has the goal to propose new molecules likely to be potent drugs against malaria and which will be far from resistances in the future.

Outline of the thesis

This thesis has been subdivided into three chapters. The first chapter describes the bibliographical study of malaria and the search for antimalarial drugs. A structural study of cysteine proteases of *Pf* (falcipain-2 and falcipain-3) is also made (with their known inhibitors). The use of molecular modelling in the drug discovery process is also defined. The second chapter describes the computational methodology used to carry out the task. It focuses mainly on molecular mechanics (MM), implicit solvation of Poisson Boltzmann (PB), thermodynamics of protein-ligand interactions. In the third chapter, we presented the results and their discussion.

OVERVIEW

1.1 Malaria as a burden in the world

The malaria parasite was discovered in 1878 by the French Alphonse Laveran [31]. The incidence of the disease has increased 2- to 3-fold over the last 35 years [32]. In 1982, it was estimated that approximately 215 million people were chronically infected and that some 150 million new cases were reported each year [33-35]. Globally, death cases related to malaria grew rapidly from 1990 and reached a peak of 232 million cases in 2003, with 1.2 million deaths in 2004 alone. In the year 2000, with the adoption of the Millennium Development Goals (MDGs) to improve the state of the world by 2015, one of the eight goals was to reduce the number of malaria cases by 50% by 2015 [34]. Malaria incidence has indeed declined in recent years in all regions of the world [35]. This is because, between 2000 and 2015, an expansion of interventions against malaria contributed to a 37% decline in malaria mortality rates globally, which reduced an estimated 4.3 million deaths. It is, however, reported that more than 3.2 billion people worldwide are still at risk for malaria, considered high risk for 1.2 billion, with an estimated 214 million malaria cases, leading to approximately 438,000 malaria deaths [35]. According to WHO, at the beginning of 2016, nearly half of the world's population was at risk of malaria, Fig. 1.1, [36]. Malaria is known to have a large global burden, with strong poverty-promoting effects. It persists as chronic infections, in spite of the available effective medical treatments. Thus it could be classified as a neglected tropical disease [37, 38]. Diseases like malaria, caused by tropical parasites, represent a major cause of mortality and morbidity in the developing world [38]. Malaria being one of the deadliest diseases the world is facing. Most of the morbidity cases have been recorded in Africa, South East Asia and South America, the majority being pregnant women and children under five years of age [39]. The situation has been worsened by the resistance of the malaria parasites to conventional drugs, particularly in *P. falciparum*, added to the fact that these poverty-stricken populations hardly have access to commercialized drugs ¹ [40-42]. *P. falciparum* causes hundreds of millions of infections annually and is regarded as the most virulent human malaria pathogen [43-45]. This species is responsible for one death in five children dying in Africa [46].

¹A typically approach to resolve this problem is to identify key pathways in *P. falciparum* physiology in which enzymes might be novel drug targets. Thus, the present challenge of most scientists working on novel antimalarial drugs is to design, synthesize and test more potent and specific drugs that inhibit new targets than those inhibited by currently used drugs.

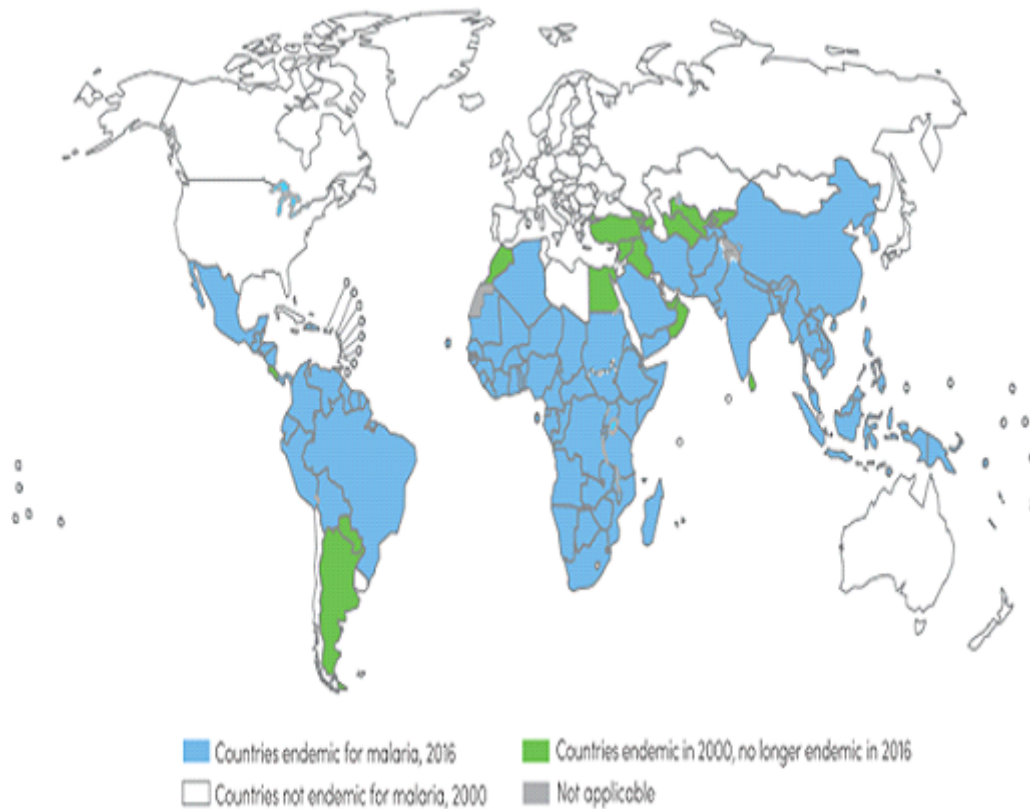


Figure 1.1: World distribution of malaria in 2016 [36].

Moreover, it is reported that an African child has between 1.6 and 5.4 episodes of malaria every year and every 30 seconds a child dies of malaria in Africa [46].

1.2 The case of Africa

According to the WHO report in 2017 [47], 90% of malaria-related deaths in 2016 were recorded in Africa, making a total of 206 Africans dead from mosquito bites². The incidence of malaria has increased in Africa between 2014 and 2016. *Pf* is the most prevalent malaria parasite in sub-Saharan Africa where it is responsible for 99% of estimated malaria cases in 2016 [47].

1.3 Malaria in Cameroon

Malaria is the most prevalent disease in Cameroon and remains the leading cause of morbidity and mortality. It is recognized as a public health problem and is at the center of health concerns by the public authorities. Indeed, this disease occurs in the ten regions of the country, and nearly two-thirds of the population is exposed (MINSANTE/PNLP³, 2006). Moreover, in 2017 according to the

²Referred to as High-burden continent

³PNLP: Programme National de Lutte contre le Paludisme. Cameroonian Ministry of Health.

Cameroon National Program for Malaria Control of the Cameroonian Ministry of Health, and the WHO report 2017, Cameroon recorded 4000 deaths from malaria, 70% of which were constituted children under five. The northern region is the most affected area with a morbidity rate of 29.3%. The Northwest has the lowest morbidity rate at 13.06%. The WHO report on malaria in 2017 in Cameroon shows that the curve of malaria infection has risen, contrary to the downward trend observed since 2010. And in this report Cameroon is ranked among the top 10 most countries affected by malaria in the world, contributing 3% of the mortality rate of the entire planet. These statistics show that malaria remain a major challenge for public authorities and researchers in Cameroon.

1.4 The different species of *Plasmodium*

Malaria is a parasitosis caused by a protozoan of the genus *Plasmodium* of the apicomplex family, transmitted by the bite of a female anophele. Four species were known to infect humans but recently a fifth long-confounded species was discovered [48, 49]. This brings to five the number of species that can infect humans. Thus we have:

- ▷ *Plasmodium falciparum* (*Pf*): it is the species responsible for the most serious form of the often fatal disease. It is present in the tropical zones of Africa, Latin America and Asia. This species develops resistant forms recorded against existing antimalarial treatments.
- ▷ *Plasmodium vivax* (*Pv*): it co-exists with *Pf* in many parts of the world, and is widespread in South America and Asia. It is responsible for mild benign fever. The *Pv* is not as innocuous as one can imagine, because serious cases even fatal, due to *Pv* were reported [50-52]. And recently in 2016 Shaw et al. have reported serious cases related to *Pv* infection [53]. This species is also a serious threat to many peoples in the world.
- ▷ *Plasmodium oval* (*Po*): it is found in intertropical Africa, as well as in certain regions of the Pacific (mainly present in endemic areas where *Pv* is absent). *Po* does not kill but can cause reviviscence for 4 to 5 years.
- ▷ *Plasmodium malariae* (*Pm*): it has a scattered geographical distribution. *Pm* is not lethal but can cause relapses 20 years after the first infection.
- ▷ *Plasmodium knowlesi* (*Pk*): infections in humans due to *Pk* are recent. This is explained because this species has long been confused with *Pm* [48, 49]. This species is only present in Asia [54].

Of these five species, *Pf* and *Pv* appear to be responsible for high rates of malaria-related morbidity and mortality around the world [55].

1.5 Malaria vectors

Anopheles mosquitoes are vectors of malaria. It is estimated that there are about 400 species of Anopheles. Of these species, 25 are known by their vector criteria. The main characteristics of these vectors are:

- ▷ a long life;
- ▷ resistance to environmental change; and
- ▷ a tropism for the man.

Funestus anopheles and *gambiae anopheles* remain the predominant vectors found in wetlands, while *Anopheles arabiensis* is the most dominant vector in dry (or arid) areas. The intensity of malaria transmission in sub-Saharan Africa depends on the abundance of these species in these areas [56]. In sub-Saharan Africa, *Anopheles gambiae* is the vector of *Pf* [57].

1.6 Malaria co-infected with HIV

Coupling of malaria co-infection with HIV/AIDS is nowadays a real public health problem. Indeed, it is becoming increasingly clear that the coupling of malaria co-infection with HIV/AIDS causes an interaction between infections. Malaria worsens HIV infection in adults and pregnant women by increasing viral load. It is also possible that malaria accelerates progression to AIDS and increases the risk of HIV transmission from adult to adult and from mother to child. HIV-infected people at increased risk of complicated and severe malaria and death. Reports also suggest that failure of malaria treatment may be more common among people with HIV who have low CD4 counts compared to HIV-free patients. Further research is suggested in goal to investigate the impact of malaria on the natural course of HIV, potential therapeutic implications, interactions at the cellular and molecular levels, and drug interactions that may arise between antimalarial and drug treatments ARVs.

1.7 Means of control of malaria and the limits of existing treatments

1.7.1 Fight against malaria

The development of an effective vaccine against malaria remains a major challenge for researchers in the eradication of this disease [58, 59]. Indeed, vaccine attempts remain limited because of the complexity of the parasite's biology and its antigenic diversity [60]. Clinical trials of vaccine candidates

to date have not shown good results in terms of protection and duration of action [60]. RTS, S/AS01, also called Mosquirix™, is an injectable vaccine developed by GlaxoSmithKline since the 1980s in partnership with PATH. It has been proven that this vaccine has been shown to provide young children with partial protection against malaria. This product is now evaluated in sub-Saharan Africa as a complementary tool for control that can be added (and not replaced) to the core set of prevention, diagnostic and treatment measures recommended by the WHO. Currently, the means to fight against malaria are: Anti-vector prevention and curative treatment

a) Anti-vector prevention

It involves the control of the vector agent through the use of long-lasting pyrethroid insecticides, long-lasting pyrethroid insecticide-treated mosquito nets [61, 62] and sometimes by a preventive treatment for travelers. Besides this, there are other means, especially for personal protection. Domestic insecticides are used here, clothing is provided for body parts exposed to mosquito bites, a fine mesh is recommended for house openings, air conditioning and the use of repellents. Remember that these measures are to clean up the environment and the living environment. This allows among other things to eliminate some potential breeding sites [61, 62]. In addition, chemical and biological methods are used to limit the mosquito population [62]. In fact, the larvae of the vector mosquito can be destroyed by larvicides (Temephos) and by biological control by introducing larvivorous fish (such as *Aphanius dispar*) or bacteria (such as *Bacillus thurengiensis*) into larval breeding sites [61].

b) Curative treatment

The drug can be defined as any substance that has curative or preventive properties against diseases [63]. By extension, it is considered as any product that can be administered to humans for the purpose of establishing a medical diagnosis, or restoring, correcting or modifying an organic function. It also defines as a small organic molecule that either activates or inhibits the function of a biomolecule like proteins which are involved in disease and hence results in a therapeutic benefit to the patient. Antimalarials are drugs active against the infection of the body by four species of haematozoa of the genus *Plasmodium* [64]. The history of antimalarial therapy includes the use of a large number of small molecules, e.g. quinine, chloroquine, artemisinin and atovaquone, are the treatment of choice [65]. Artemisinin, which is the most efficient currently used anti-malarial drug, is only a single mutation away from emerging drug resistance [66]. Artemisinin combination therapy (ACT), the first-line treatment, was suggested by WHO since 2006 in order to decrease the transmission of *Pf* and delay the onset of resistance against artemisinin alone in malaria-endemic regions [67]. Since 2012, WHO has recommended for countries in the phase-out phase of *falciparum* malaria and in countries with artemisinin-resistant strains, the addition of a single low dose of primaquine to treatment ACTs in order to reduce gametocyte carriage. Which will limit vector transmission [68]. Many of the currently used antimalarial drugs were not, however, developed on the basis of rationally identified targets. On the contrary, they were discovered by the identification of the antimalarial activity of natural products (NPs) (e.g. quinine and artemisinin, Fig. 1.2), compounds chemically related to NPs (e.g. chloroquine and artesunate), or compounds active against other infectious pathogens (e.g. antifolates and tetracyclines [69]).

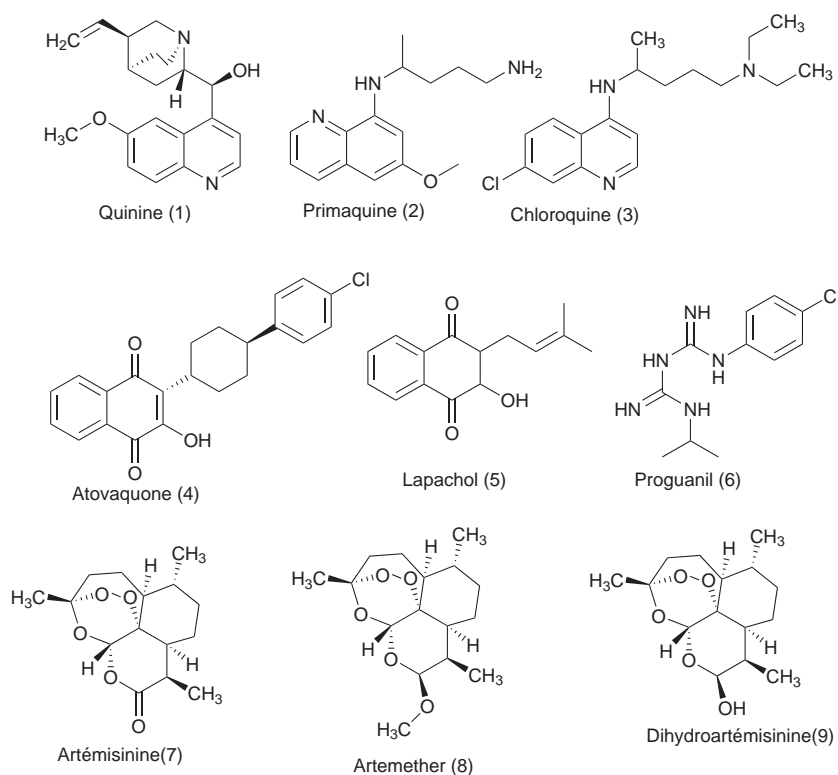


Figure 1.2: Natural products and synthetic molecules currently used as antimalarial drugs

1.7.2 Limitations of existing treatments

By the mid-1950s, it was confidently expected that malaria would be eradicated worldwide. However, by the mid-1960s these expectations were cut off by the problems of resistance [70]. The vector mosquito developed resistance to potent insecticides such as dichlorodiphenyltrichloroethane [71] and certain strains of *Pf* became resistant to chloroquine treatment so that by the early 1980s, several strains of *Pf* had become multi-drug resistant [33]. Also, the existence of *Pf* strains resistant to ACT was already reported in South East Asia almost a decade ago [72] (Fig. 1.3). This raises concerns about malarial intervention in Africa, where virulent and resistant strains of *Pf* are prevalent. Malaria drug resistance is often associated with genetic mutations in genes encoding for target proteins. This may result in an unusual expression and folding of these proteins, thus modifying the usual binding site of the drug. This is sometimes due to massive and systematic use of the available drugs. To date, one of the major problems in the curative fight against malaria is the emergence and spread of resistant strains of parasites to available drugs [73-75] and the lack of an effective vaccine. It is therefore very needed to develop other generations of antimalarial drugs, acting against new biochemical targets.

1.8 Discovery of drugs against malaria

The increasing resistances of the mosquitoes to insecticides on the one hand and that of existing anti-malarial drugs, on the other hand, show the need for rapid discoveries of innovative and inexpensive

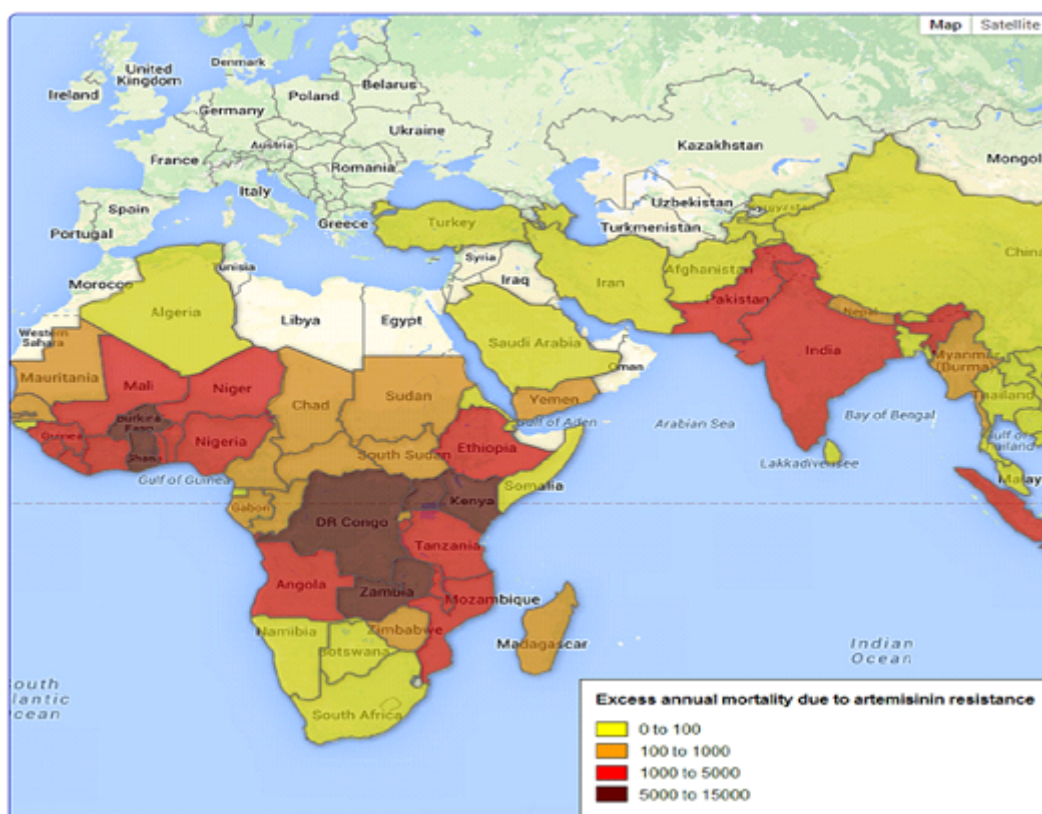


Figure 1.3: Geographical distribution of mortality due to artemisinin and artemisinin-combination therapy resistance in malaria-endemic areas, year 2014 [76].

drugs. It is, therefore, more than urgent to seek new therapeutic targets at the level of the parasite. This requires a better understanding of its development process that is to say the biology of the parasite. Thus, many malaria prevention and treatment programs involving various actors have been put in place. We can cite :

- ▷ Roll back malaria (RBC) launched in November 1998 jointly by WHO, the World Bank, the United Nations Children’s Fund (UNICEF) and the United Nations Development Program (UNDP). Its main objective was to halve the death rate of the disease by 2010.
- ▷ Malaria Vaccine Initiative (MVI), launched by the Bill and Melinda Gates Foundation in 2000, aimed at improving prevention and developing a vaccine against malaria.
- ▷ Other programs have emerged such as Medicines for Malaria Venture (MMV), Multilateral Initiative on Malaria (IMI).

In the year 2002, a flagship program called *Pf* Genome Sequencing Consortium was launched. His goal was to sequence the entire genome of the human malaria parasite. And nowadays, the sequence of almost all *Pf* genes will have been determined [77]. This result opens the way for genetic, biochemical and immunological research aimed at developing new drugs and vaccines for malaria.

1.9 *Plasmodium falciparum*: life cycle of the parasite

To rationally develop new anti-malaria agents, it is essential to study the genetics and physiology of *P. falciparum*. It is equally important to understand the *P. falciparum*-host interaction in order to learn how the parasite circumvents host defences and causes disease. The life cycle of *Plasmodium* parasites is completed through many morphological changes, which alternate between invertebrate and vertebrate hosts. *P. falciparum* has three critical points in its life cycle, during which a small number of parasites multiply rapidly in order to generate much larger populations [78]: the mosquito vector stage, liver and blood stages (Fig. 1.4) [78, 79]. The two last processes take place in the vertebrate (human) host. All malaria types have a similar life cycle. The female and male gametocytes (the only stages within the parasite's life cycle able to mediate the transition from the human host to the insect host [80]) mate inside the gut of the mosquito vector and go through meiosis and thereafter migrate to the midgut wall of the vector mosquito and produce oocyst [79], in which many sporozoites develop [81]. Sporozoites, the infectious form of the parasite, are injected into a human host through the saliva of an Anopheles mosquito when it bites the human host to feed on blood. These sporozoites travel to the liver, where they invade the human liver cells within a few minutes, thus taking a new form and multiplying (exo-erythrocytic forms). During the process of liver cell rupture, blood stage parasites (also known as merozoites), are released, thereby constituting the erythrocytic cycle/ asexual blood stage. Each merozoite invades a red blood cell. After two days the parasite starts repeating mitotically. This continues through some stages (e.g. ring, trophozoite and schizont), thus producing more merozoites [82, 83]. This is the stage of the life cycle which causes disease, often leading to death if nothing is done [41, 84]. Some merozoites change into gametocytes, a form which does not cause disease but remains in the blood until cleared by drugs or the immune system, or taken up by a mosquito bite. In the mosquito's midgut, "male" gametocytes fertilize the "female" to form oocysts [85], which mature into thousands of sporozoites, eventually swimming to the mosquito's salivary glands to initiate another life cycle (Fig. 1.4). The form of the parasite best studied is the asexual blood stage, largely due to the fact that this is the disease-causing stage. Its experimental accessibility is relatively easy [82]. Thus, most Plasmodium proteins that have been well characterized are expressible during the erythrocytic cycle. Although *P. falciparum* proteases may be involved in many processes, only three functions (hydrolysis of host hemoglobin, erythrocyte invasion, and erythrocyte rupture) in erythrocytic parasites are well defined. Therefore, potential anti-malarial drug targets could be broadly classified into three mainly categories: hemoglobin (Hb) degradation enzymes, signaling and transport protein factors [40]. Thus *Pf* enzymes involved in the degradation of hemoglobin have been validated as potential targets of antimalarials after the unveiling of the human genome in 2002 [40, 86] because the degradation of hemoglobin is a vital process for *Pf* [87].

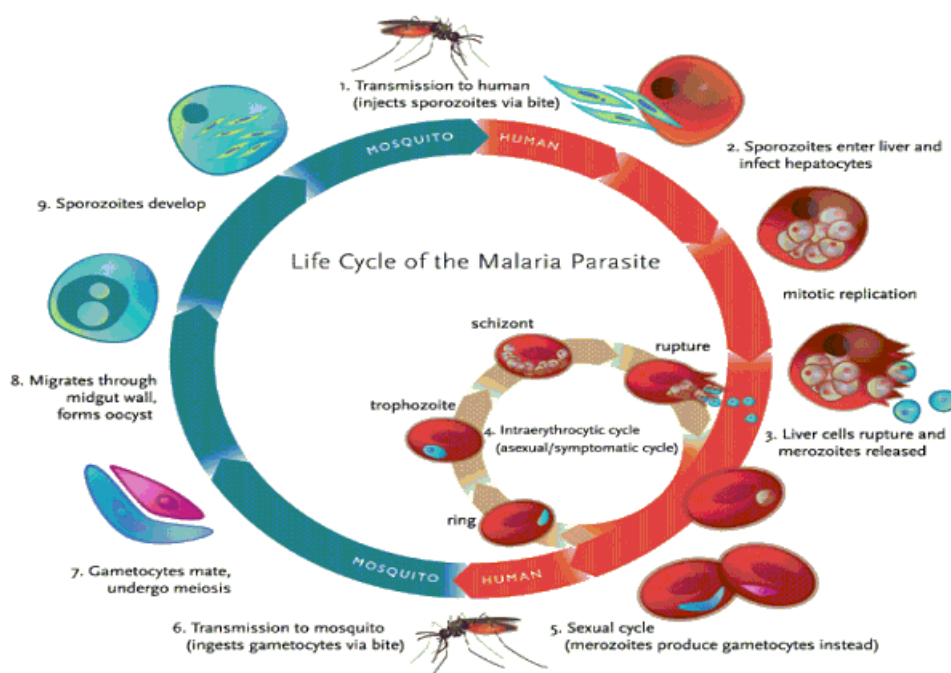


Figure 1.4: The malaria parasite life cycle [88].

1.10 The proteome of *Plasmodium falciparum*

According to the works of Gardner et al. [77], *Pf* has approximately 5268 predicted proteins, about these, 60% (3 208 hypothetical proteins) did not have sufficient similarity to proteins in other organisms to justify the provision of functional assignments. In order to better know the proteins involved in different stages of the *Pf* development cycle, other studies have been led. Thus, the analysis of four stages of the cycle (sporozoite, merozoite, trophozoite and gametocyte) by Florens et al. [89] has led to identify 152 proteins (6%) present in these four stages. At the sporozoite stage, about 49% of the proteins were identified specifically at this stage. Similarly, merozoites, trophozoites, and gametocytes are present at about 20 to 33% of specific proteins.

1.11 Mode of action of antimalarial drugs

The majority of antimalarials target the erythrocyte cycle of the parasite. It is a little more than 60 years since the research focuses on this erythrocyte cycle of the parasite in order to deepen the biochemical knowledge to understand the mechanisms of action of antimalarials [90]. Thus, antimalarials can be classified into two categories according to their modes of action and their chemical structures:

- ▷ Erythrocyte schizonticides; and
- ▷ Inhibitors of metabolite synthesis.

a) Erythrocyte schizonticides

The schizonticides are antimalarial which have a quinoline or endoperoxide core. Among these we have:

- ▷ Aminoalcohols (quinine, mefloquine, lumefantrine) ;
- ▷ Amino-4-quinoline (chloroquine, amodiaquine,) ;
- ▷ Artemisinin and its derivatives

The product of degradation of hemoglobin in the digestive vacuole of the parasite called heme (Ferriprotoporphyrin IX), is targeted by these antimalarials [91]. Heme is crystallized by the parasite in order to protect itself from its toxic effects. The mechanism of action of quinine can be explained by its accumulation in lysosomes (acidic digestive vacuoles) of blood schizonts. Here, hemoglobin is digested, subsequently inhibiting the enzyme that polymerizes hemozoin (a degradation product of hemoglobin) that is toxic to the parasite. Chloroquine and amodiaquine have the advantage of having a fast action. Their mechanism of action is almost similar to that of quinine. Andchloroquine was the most studied, although its mechanism of action is not fully understood. Many studies have suggested that chloroquine would act by forming a complex with hematin (oxidized heme) by electrostatic bonds and thus inhibit the crystallization of heme into hemozoin [92-94]. This chloroquine-hematin complex is thought to have toxic effects on parasites. Artemisinin and its derivatives have a fast but brief activity. They act on multidrug-resistant and chloroquine-resistant *Pf* strains. Their antimalarial activity is based on their peroxide (trioxane) structure and the adverse effects are few and relatively safe. The mechanism of action of artemisinin remains the subject of much debate in the scientific community. It has been reported that artemisinin and its derivatives are pro-drugs, and are administered in inactive form [95]. In the presence of the ion Fe^{2+} , the endoperoxide bridge of artemisinin is cleaved. This cleavage leads to the activation of artemisinin and the release of free radicals having toxic properties on parasite development [95, 96]. Others studies have showed that artemisinin would have an inhibitory activity on cysteine proteases (falcipains) [96, 97].

b) Inhibitors of metabolite synthesis

During the erythrocyte phase, *P. falciparum* uses a nucleotide synthesis pathway to synthesize its nucleic acids [90]. This pathway of synthesis of folic acid by the parasite is inhibited by antifolics and antifolinics. Thus, the main antimalarial inhibitors of parasite nucleic acid biosynthesis are:

- ▷ antifolinics (proguanil, pyrimethamine); and
- ▷ antifolics(sulfonamides and sulfones)

1.12 The cause of resistance to existing treatments

Drug resistance⁴ can be defined as the reduction in effectiveness of a drug in curing a disease [98]. Since the appearance of chloroquine-resistant strains, many studies have been led in order to understand the molecular mechanisms involved in antimalarial drug resistance. Thus it is reported that two transport proteins called *P. falciparum* multidrug resistance 1 (*Pfmdr1*) and *P. falciparum* chloroquine resistance transporter (*Pfcrt*) located in the food vacuole of the parasite are involved in the mechanism of resistance [99].

1.13 The pharmaceutical drug discovery process

From the previous sections, it stands out clear that the eradication of malaria will require the design and development of new therapeutically active agents against, particularly, the resistant strains of *Pf* that is responsible of the lethal cases of malaria [45]. In traditionally drug discovery approach, on average, it takes 10-15 years and US \$500-800 million to introduce a drug into the market [100, 101].

Over the years there have been numerous efforts in the development of new drugs against Malaria [102, 103]. Drug discovery today is an interdisciplinary undertaking, involving enormous costs⁵. A recent study leaded by Lubell et al. [76] compares the projected health and economic consequences if widespread malaria resistance to ACTs spreads or emerges in many more countries. The authors have developed an economic model to explore a scenario where ACTs failed at a rate of 30% and treatment for severe cases of malaria is reverted to quinine. The model projects that the scenario of widespread antimalarial resistance will result in each year in an increase of more than 116,000 deaths, an excess of \$32 million in health care costs, and an estimated \$385 million from productivity losses due to extended patient illness. This is usually because the release of a new drug on the market involves many years of expensive laboratory work⁶ and clinical testing (which represents the greatest cost in the drug development process). Thus, pharmaceutical companies expect to require a significant payback from each new drug in order to compensate for the heavy investments, which may be the reason why the cost of treatment for malaria-resistant cases is estimated this high [104, 105].

A typical procedure for the design and development of a drug involves several crucial steps, beginning with the identification and validation of a drug target⁷ (receptor), followed by the identification and

⁴In any population of infectious agents, some have a mutation that helps them resist the action of a drug. The drug then kills more of the nonresistant microbes, leaving the mutants without competition to multiply into a resistant strain. This situation is more likely if the drug is not taken properly (e.g., a course of antibiotics not completed, anti-HIV drug doses missed) or not prescribed properly (e.g., an antibiotic given for a viral disease). Resistance factors can also be transferred between species that infect the same body. The over prescription of antibiotics in humans has accelerated the evolution of resistant strains of bacteria making it increasingly difficult to fight off certain disease-causing organisms.

⁵It is estimated that \$US 49.3 billion was invested globally in drug development in 2004.

⁶It often takes at least 12 years to conceive and develop a single new drug.

⁷This involves the identification of a disease related biochemical (pathological) process and the molecular structures that can be used to interfere with this process, so as to eventually cure the disease.

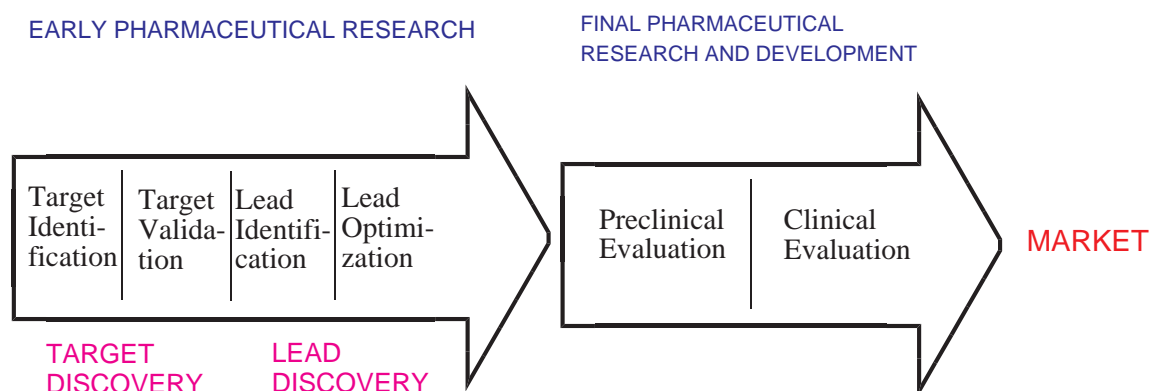


Figure 1.5: The pharmaceutical drug discovery pipeline today [108].

validation of a lead compound ⁸ which should be active at concentrations of 10 μM or less [106], then preclinical and clinical trials before the eventual release into the market [107] as one can see in the modern day drug discovery pipeline outlined in Fig. 1.5. The overall success of a drug discovery program depends on the choice of the biological target, the correct disease association for the chosen target and the correct choice of a drug development candidate for successful human clinical trials [104, 105].

Practically this whole process involves five major steps:

- (i) selection of a biological target of interest,
- (ii) identification of a lead chemical compound, often referred to as the hit ⁹,
- (iii) optimization of the lead structure (hit to lead¹⁰),
- (iv) animal tests in the goal to ensure that the drug is efficacious in animal disease models and therefore safe for human clinical trials, and
- (v) human clinical trials that prove the drug to be efficacious and safe for use in the general patient population.

Computer-aided drug design/discovery (CADD)

It is generally recognized that drug discovery and development are very time and resources consuming processes¹¹. Computer-aided drug design and discovery is also known as rational drug design use

⁸Defined here as a compound which will interact with the target in some way, so as produce a relevant pharmaceutical intervention.

⁹A compound that inhibits or has a high binding affinity for one or more targets. Often, hits are initially virtual entities, i.e., coming from computational calculations.

¹⁰A lead is a hit that has been well characterized experimentally, e.g., one that has been shown to have a high dissociation constant (Kd) for the target of interest such that the functional activities of the target are decreased on binding and/or has demonstrated effectiveness of treatment against disease in an animal model. A potential lead is a computationally predicted hit that has been shown to work experimentally against *in vitro* (cell culture) disease models of the organism.

¹¹12 to 15 years.

computational approaches to discover, develop, and analyze drugs. CADD approaches make use of the structural knowledge of either the target (structure-based) or known ligands with biological activity (ligand-based) to facilitate the determination of promising candidate drugs. When a target is selected for the design of new lead compounds three different situations can be faced regarding the amount of information of the system that is available:

- i) the structure of the receptor is well known and the bioactive conformation of the ligand is not known,
- ii) only the bioactive conformation of the ligand is known and
- iii) the target structure and the bioactive conformation of the ligand are unknown.

For the first situation, techniques such as structure-based computer-aided drug design and discovery (SB-CADD) can be used while techniques as ligand-based computer-aided drug design and discovery (LB-CADD) are often used for the second case. These two techniques are summarized in Fig. 1.6. For the last case, the technique named *De novo* design is often used.

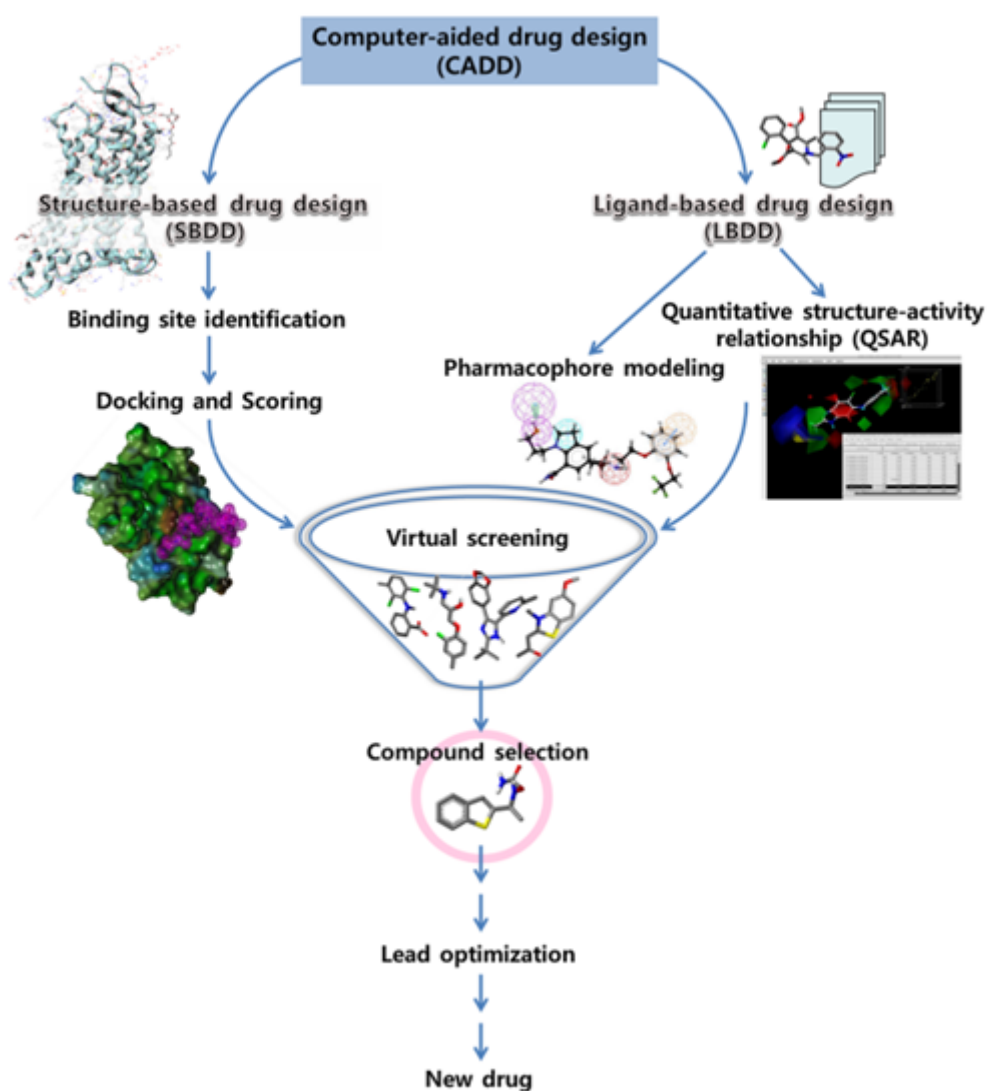


Figure 1.6: An example of workflow for in silico drug design and discovery [109].

1.13.1 Structure-based computer-aided drug design and discovery (SB-CADDD)

Structure-based computer-aided drug design (SB-CADDD) relies on the ability to determine and analyze 3D structures of biological molecules. This technique is based on the fact that the ability of molecule (often call ligand) to interact with a specific protein and exert a desired biological effect depends on its ability to favorably interact with a particular binding site on that protein [110, 111]. Thus, molecules that share favorable interactions will exert similar biological effects. Compounds like-drug have been identified through this approach [110, 111]. Nowadays, the coupling of SBDD with *in silico* drug design approaches has played a key role in the development of several marketed drugs. [112, 113]. Table 1.1 gives examples of marketed drugs obtained from the structure-based drug design while Fig. 1.7 displays its structures.

Table 1.1: Example of marketed drugs involving use of Structure-based drug design.

| Year | Generic name | Brand name | Manufacturer | Against/Inhibits | Ref |
|------|--------------|------------|------------------|------------------|-------|
| 1989 | Zanamivir | Relenza | GlaxoSmithKline | Neuraminidase | [114] |
| 1997 | Nelfinavir | Viracept | Hoffman-La Roche | HIV protease | [115] |
| 1998 | Raltitrexed | Tomudex | AstraZeneca | Thymidylate | [116] |
| 1999 | Amprenavir | Agenerase | GlaxoSmithKline | HIV protease | [117] |
| 2007 | Raltegravir | Agenerase | Merck | HIV integrase | [118] |

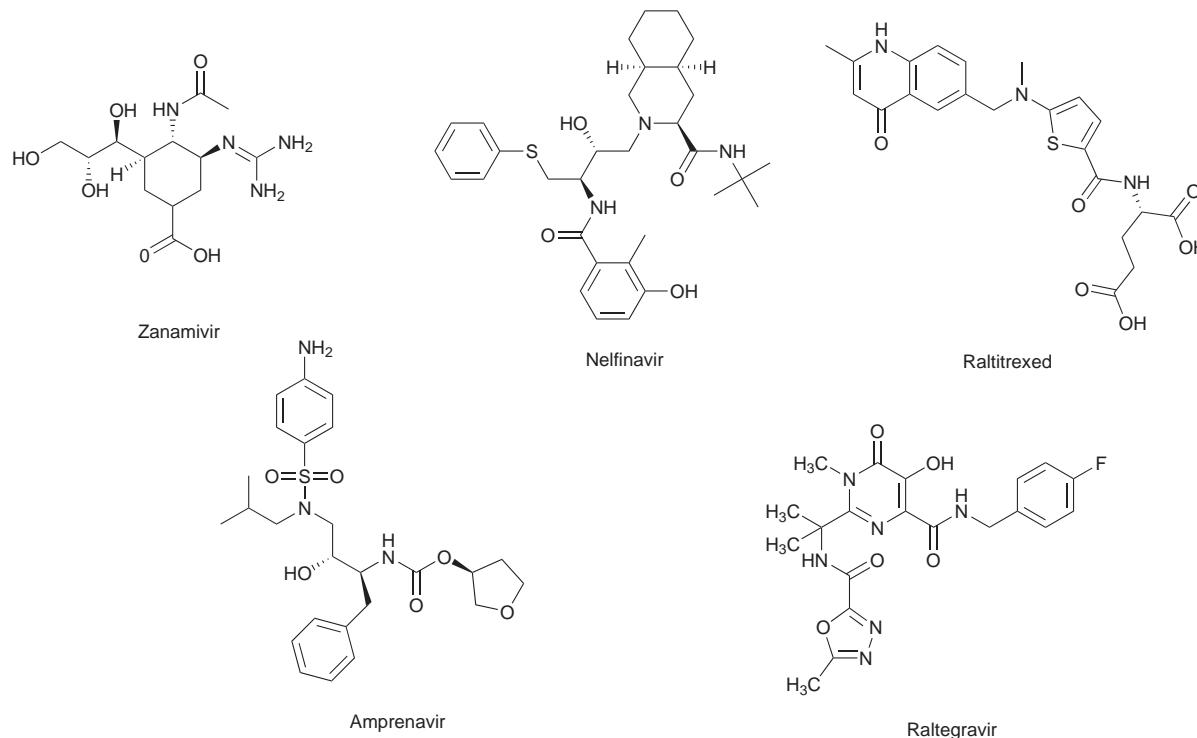


Figure 1.7: Chemical structures of example of marketed drugs involving use of structure-based-drug design [114-118].

1.13.2 Ligand-based computer-aided drug design and discovery (LB-CADDD)

When the 3D structure of the target is not available, information taken from a set of ligands against this target can allow identifying significant structural and physicochemical properties responsible for the observed biological activity. The ligand-based computer-aided drug discovery (LB-CADDD) approach, therefore, consists of analyzing of known ligands which interact with a validated target but which 3D structure is not already available. The goal of this technique is to predict the binding affinity between a ligand and its target based on the fact that similar compounds display similar interactions (often similar biological responses) with the same target. [119]. *Ab initio* quantum physics methods, or density functional theory, are frequently used to predict the conformation of the small molecule and to model conformational changes in the biological target that may occur when the small molecule binds to it. These methods provide an estimate of the electronic properties (electrostatic potential, polarizability, etc.) of the ligands which will affect the affinity with its target [120].

1.14 Antimalarial target pathways in this study

A promising target for drug development should be essential for growth, metabolism and survival of the pathogen and also absent from the human host or structurally dissimilar from the human homologue, this allows for developing non-toxic therapeutic agents [121]. Thus, several drug targets have been identified and validated against *Pf*. The design of anti-malarial compounds by structure-based drug design (SBDD) approaches have received much attention. Earlier studies indicated that *P. falciparum* has a limited capacity for *de novo* amino acid synthesis. The hemoglobin hydrolysis is a source of amino acids necessary for the life of a parasite [122].

1.14.1 Hemoglobin degradation process by *Plasmodium falciparum*

To obtain the nutrients required for growth during the intra-erythrocytic stage, the parasite has to break down a great amount of hemoglobin from the host cell [123]. The Hb degradation process occurs in food vacuoles of the parasite where 25-75% hemoglobin is digested, thereby releasing the required amino acids for incorporation into parasite proteins [40, 124]. Several proteases are involved in this degradation [125] and they can be divided into two major functional groups:

- i). proteases involved in the invasion and rupture of erythrocytes, and
- ii). proteases involved in hemoglobin hydrolysis [126].

In the latter, two relevant protease families: aspartic proteases and cysteine proteases, called plasmepsins and falcipains (FPs), respectively [87], one metalloprotease (falcilysin) [127], and one dipeptidyl aminopeptidase [128] are involved. The degradation process follows an ordered pathway (Fig.1.8)

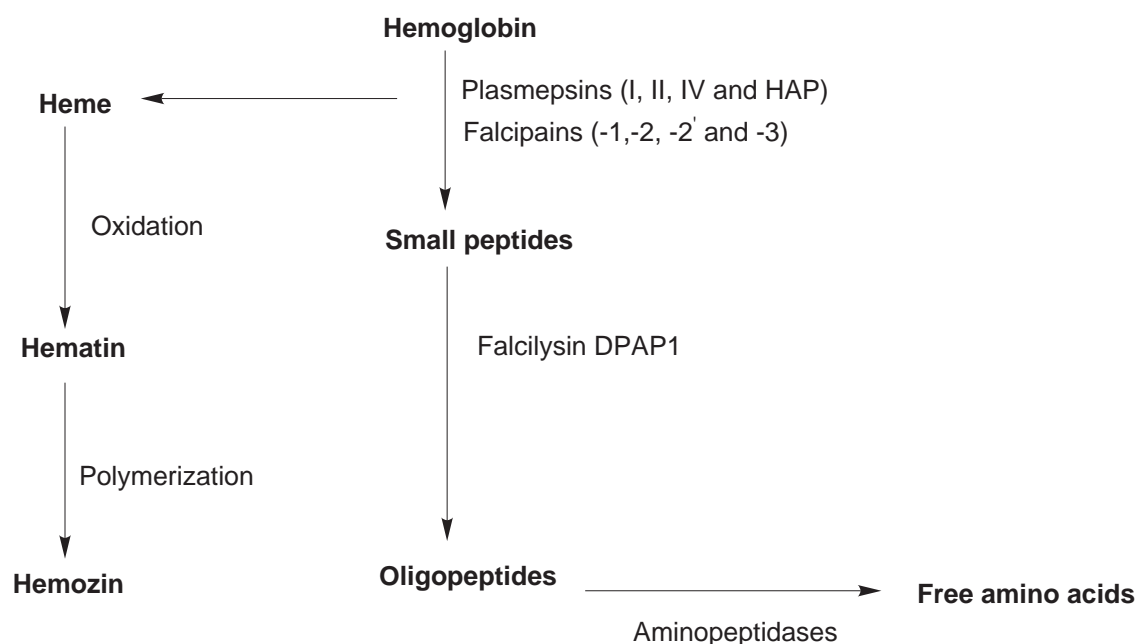


Figure 1.8: Summary of the degradation process of hemoglobin in the *P. falciparum* food vacuole [129].

[129]. Therefore, blocking the process (or inhibiting any of these enzymes) would be lethal to the parasite. We can therefore conclude that the inhibition of hemoglobin degradation by cysteine protease inhibitors would lead to parasites containing abnormal food vacuoles full of undegraded host hemoglobin.

As display in figure Fig.1.8, Hemoglobin is initially degraded by aspartic proteases (e.g. plasmepsins I, II, and possibly III and IV) located in the digestive vacuole [40, 127, 130] following by cysteine proteases. There are a total of 30 predicted cysteine proteases [43], which have been named according to the function of a catalytic cysteine mediating protein hydrolysis via the nucleophilic attack on the carbonyl carbon of a susceptible peptide bond [86].

1.15.2. Catalytic mechanism of the cysteine protease

The cysteine proteases are most commonly exemplified by papain, a well-described plant enzyme isolated from the latex of *Carica papaya* fruit. The papain molecule consists of two subdomains forming the active site cleft of the enzyme [131, 132]. The proteolytic activity of all cysteine proteases arises from the presence of the catalytic Cys and His residues in the enzyme active site [133]. In the case of papain-like cysteine proteases, the catalytic site is complemented with Asn that ensures an orientation of the His imidazole ring optimal for successive stages of hydrolysis. The crucial step of the catalytic process involves the formation of a reactive thiolate/imidazolium ion pair (Cys-S⁻/His-Im⁺), which results from proton transfer between Cys-25 and His-159 (papain numbering). In principle, the thiolate anion attacks the carbonyl carbon of the scissile peptide bond, and the double bond between the carbon and the oxygen converts into a single on Fig. 1.9. The oxygen assumes a

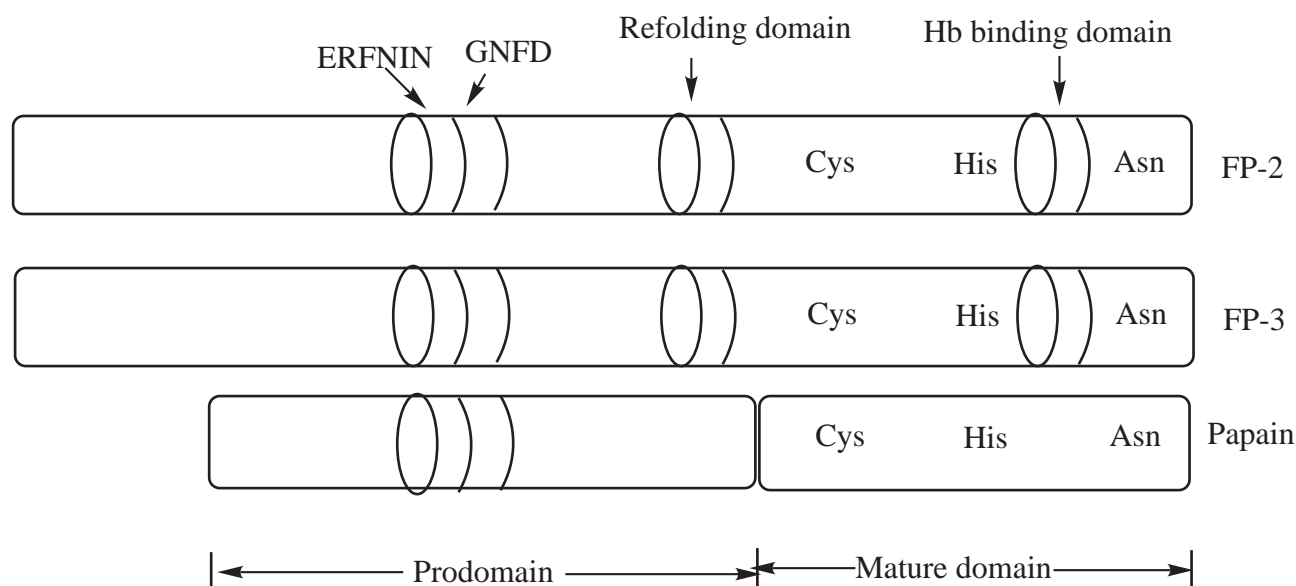


Figure 1.10: Domains of falcipains The active site residues (Cys, His, Asn) are conserved within the papain family, but falcipains have a unique N-terminal extension, which acts as a refolding domain, and a C-terminal inserted as a hemoglobin (Hb) binding domain. The prodomain has the ERFNIN and GNFD motifs, which are conserved in falcipains and papain [135].

ized papain-family enzymes [136]. It has been recently demonstrated that the first insertion is located at the N-terminus of the mature enzyme [136]. It has been shown from previous functional studies that the N-terminal extension of FP-2 and FP-3 plays a crucial role in folding of the mature protein into its active conformation [137]. FP-2 and FP-3 efficiently hydrolyze human hemoglobin in the acidic food vacuole of parasite [138, 139]. Pandey et al. [135] showed that FPs have a hemoglobin-binding domain near the C-terminus end of the mature domain (Fig. 1.10).

1.15.1 Key aspects of FP-2 and FP-3

The interaction of a drug with a macromolecular target involves a process known as binding. There is usually a specific area of the macromolecule where this takes place, known as the active site or binding site Fig.1.11.

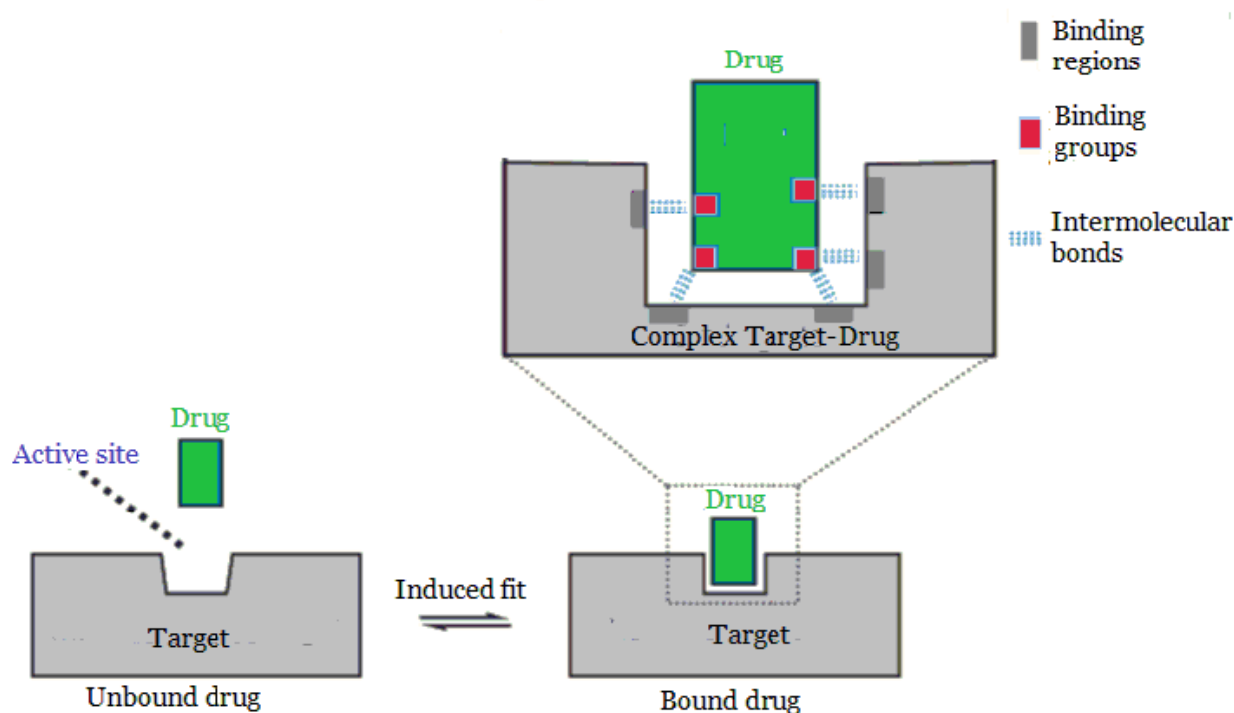


Figure 1.11: The equilibrium of a drug being bound and unbound to its target.

It has been reported that the active sites of FP-2 and FP-3 are located within a cleft between the structurally distinct domains of the papain-like fold and three amino acids residues Cys42/51, His174/183, Asn204/213, respectively (His159, Cys25 and Asn175 following papain convention, Fig. 1.10), are responsible of the catalytic activity [135]. The binding site of FP-2 and FP-3 is large enough to be subdivided into pockets S1', S1-S4 for a better analysis of the interactions between the amino acids of each pocket and inhibitors [136].

1.16 Currently known FP-2 and FP-3 inhibitors

Nowadays, many compounds have been identified as inhibitors of FP-2 and FP-3, which are able to block the enzyme's activity by forming a reversible or irreversible covalent bond with the amino acids of active site cysteine. These FP inhibitors can be broadly divided into three categories [32]:

- i). peptide-based inhibitors,
- ii). non peptidic inhibitors and
- iii). peptidomimeticinhibitors.

Most of the FP inhibitors identified so far are peptide-based inhibitors [32]. Peptide-based inhibitors include: peptidylfluoromethyl ketones [140], peptidyl vinyl sulfones [141], peptidyl aldehydes and ketoamide derivatives [142], epoxysuccinyl derivatives [143] and peptidylazirines [144]. Peptidyl-aldehydes and -ketoamides are reversible inhibitors, while the other inhibitors are irreversible. The drug design approaches discussed here include, rational approaches for the discovery of non-peptidic and peptidic inhibitors against cysteine proteases, along with examples of FPs, SBDD and CADD approaches for the design of peptidomimetic. We also include NP-based drug discovery and provide insight for the understanding of drug discovery process against both the targets. We will then review steps towards the discovery of the inhibitors against FP-2 and FP-3 using rational approaches. Many significant computer-aided drug design/ discovery approaches targeting FP-2 and FP-3 has been previously carried and have been recently fully described in the first part of this thesis [145]. Some of these inhibitors is display in Figs. 1.12, 1.13, 1.14, and 1.15.

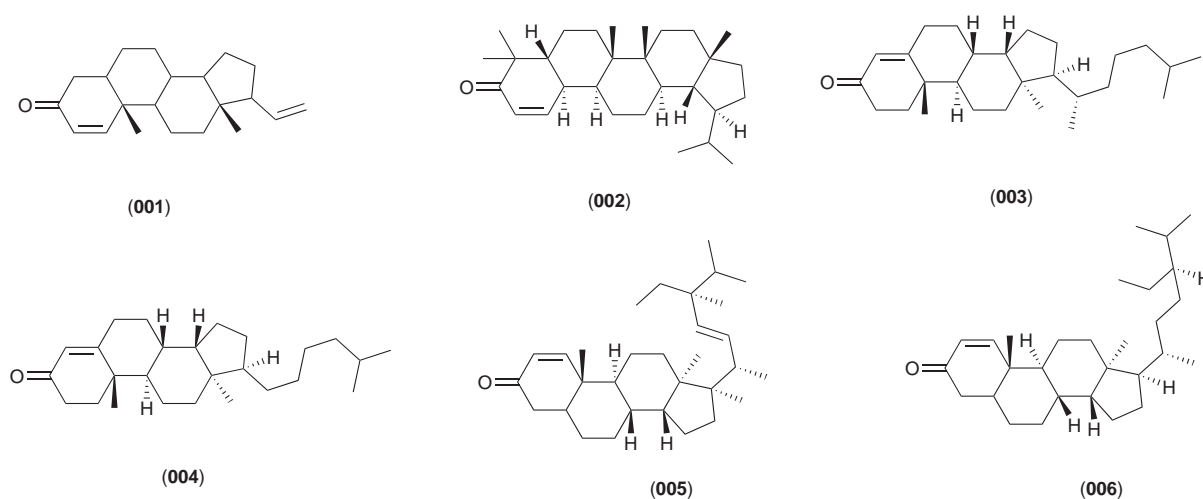


Figure 1.12: Non-peptidic structures of some FP-2/3 inhibitors identified from the ZINC database and using molecular dynamics simulation.

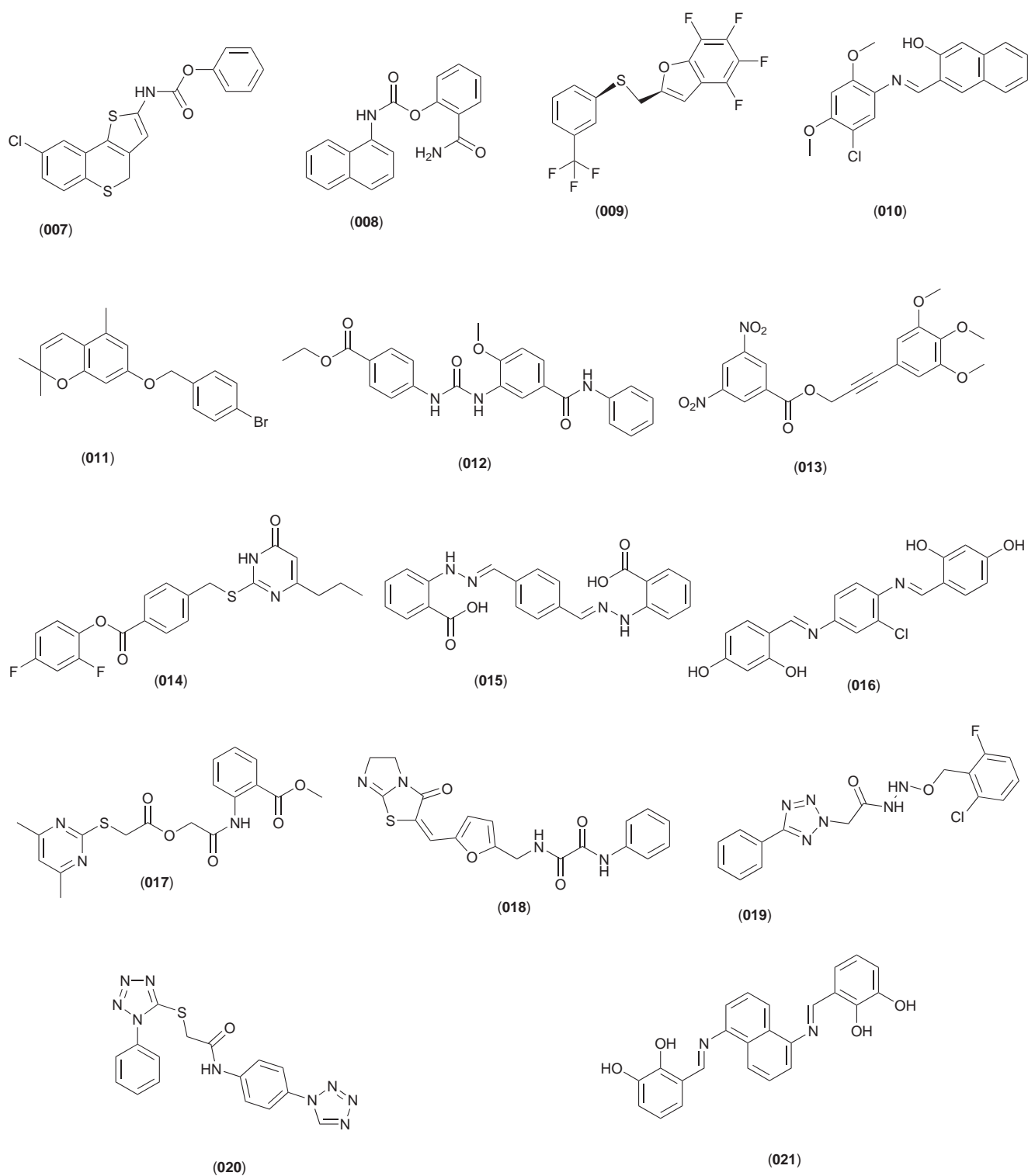


Figure 1.13: Non-peptidic structures of some FP-2 inhibitors identified using virtual screening.

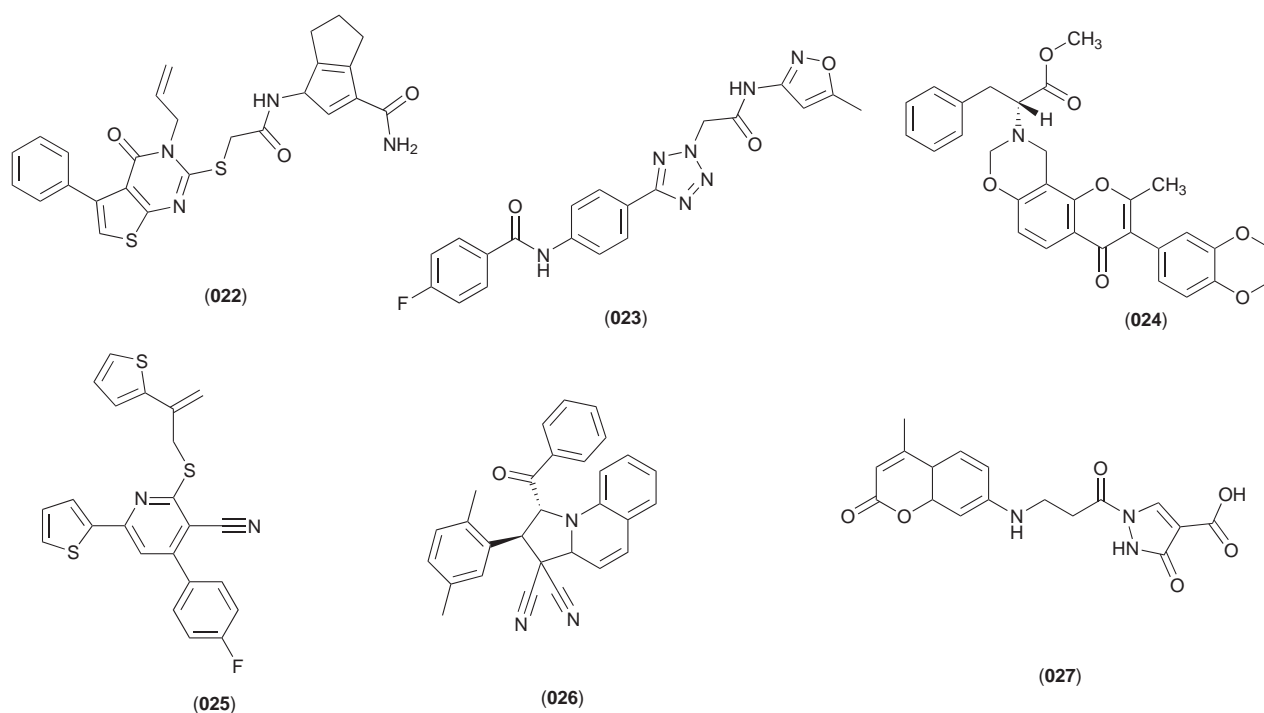


Figure 1.14: Active inhibitors of FP-2/3 identified from the SPECS database by SBVS.

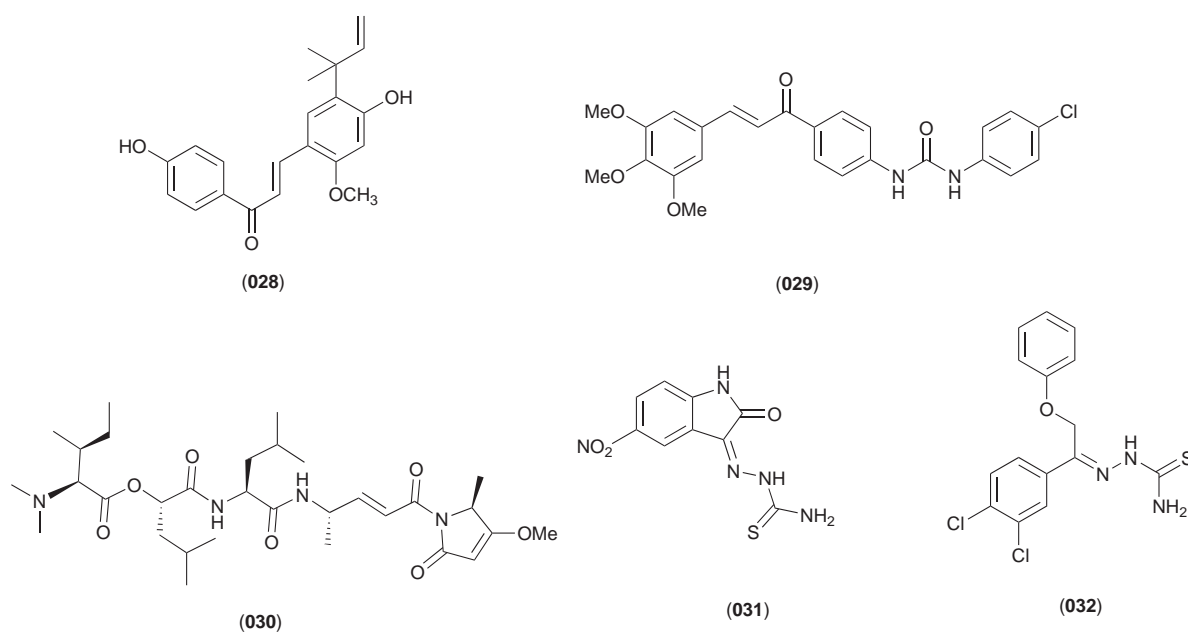


Figure 1.15: Some natural product-based inhibitors of FP-2/3.

1.17 Available crystal structures of FP-2

Traditionally, the first consideration before embarking on a CADD project is whether the detailed 3D structure of the drug target is known. SBDD of antimalarial compounds targeting FP inhibition is

possible due to the availability of several 3D-structures of these enzymes. Several crystal structures of FP-2 from *P. falciparum* have been published and stored in the protein databank (PDB) [146]. A summary of the available crystal structures of FP-2 so far is given in Table 1.2.

Table 1.2: Description of the 5 Available Crystal Structures of FP-2 from the PDB.

| PDB code | Resolution (Å) | # of Chains | Missing residues | Co-crystallized Ligands (of substrates) | Source | Ref ^a |
|----------|----------------|-------------|------------------|---|-----------------------|------------------|
| 1YVB | 2.7 | 1 | 0 | none | X-ray crystallography | [147] |
| 2GHU | 3.1 | 4 | 2 | none | " | [148] |
| 2OUL | 2.2 | 1 | 3 | none | " | [149] |
| 3BPF | 2.9 | 4 | 12 | ligE64 ^b | " | [150] |
| 3PNR | 2.6 | 4 | none | none | " | [151] |

^aReferences, ^bligE64=((1S,2S)-2-(((S)-1-((4-Guaminobutylamino)-4-methyl-1-oxopentane-2-yl) carbamoyl)cyclopropanecarboxylic acid

Although the majority of previous studies was devoted to the inhibitors search of FP-2 and FP-3, it has been reported that FP-3 is expressed later in the life cycle and appears to be a more efficient haemoglobinase than FP-2 [139]. And the literature indicates that the inhibition of FP-3 is lethal for the parasite [145]. FP-3 is, therefore, a well-known drug target of *Pf*. Thus, from the above due, we have focused in the design of FP-3 inhibitors in the current work.

1.18 Available crystal structures of FP-3

For FP-3, two crystal structures from *P. falciparum* have been deposited so far. A summary of the available crystal structures of *P. falciparum* FP-3 is given in Table 1.3.

Table 1.3: Description of the 2 Available Crystal Structures of FP-3 from the PDB.

| PDB code | Resolution (Å) | # of Chains | Missing residues | Cocrystallized Ligands (of substrates) | Source | Ref ^a |
|----------|----------------|-------------|------------------|--|-----------------------|------------------|
| 3BPM | 2.5 | 2 | I1 | Leupeptin ^b | X-ray crystallography | [149] |
| 3BWK | 2.42 | 4 | I5 | ligK11017 ^c | " | [30] |

^aReferences, ^bligLeupeptin=N-Acetyl-leucyl-leucyl-N-[(5-[(diaminomethylidene)amino]-1-oxopentan-2-yl)]-leucinamide, ^cclig K11017= leucine-homophenylalanine-phenylvinyl sulfone (Mu-Leu-hPh-VSPH).

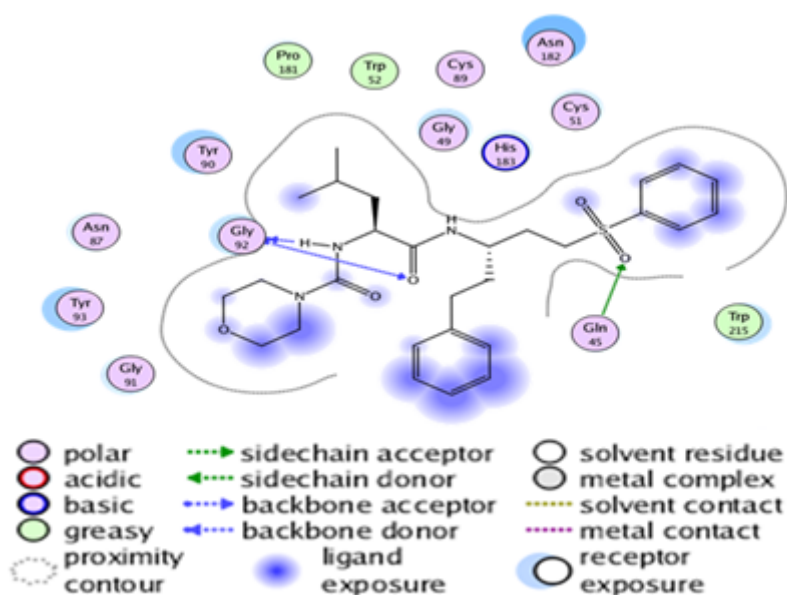


Figure 1.16: The Binding pocket of chain A of FP-3 crystal structure (PDB code 3BWK, resolution 2.42 Å). K11017 (color in yellow) in stick representation, protein atoms in line representation (carbon atoms in their usual colours). This diagram was produced using the Discovery studio software.

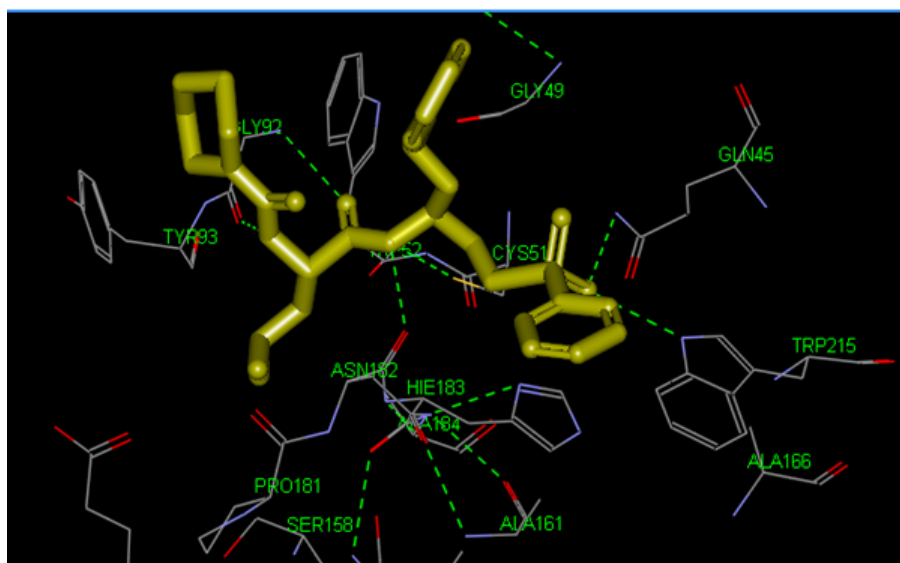


Figure 1.17: 3D Binding pocket of chain A of FP-3 crystal structure (PDB code 3BWK, resolution 2.42 Å).

In view of the crystallographic structures of FP-3 available in the PDB and presented above, the 3BWK complex presents a ligand similar to the series of molecules that we used (cf chapter 2 Fig. 3.2). This is one of the reason why this complex was chosen for the calculations performed in the current work.

1.19 Molecular modeling

Molecular modeling is a general term which covers all the theoretical and computational methods, used to model or simulate the behaviour of molecules. These techniques are often used to build, display, manipulate, simulate, and analyze molecular structures, and to calculate the properties of these structures [152] in order to guide the researcher to draw and design new molecules, which is called drug design. Molecular modeling is used in a number of different research areas as biophysics, chemistry, bioinformatics, etc., thereby the term does not have a rigid definition. In a general way, it describes the generation, manipulation or representation of three-dimensional (3D) structures of molecules and their associated physico-chemical properties. So it is a tool for researchers concerned about the structure and reactivity of molecules. And it involves a range of computerized techniques based on theoretical chemistry methods and experimental data in order to predict molecular and biological properties. Molecular modeling methods can be classified in the following categories [153]:

- ▷ Quantum methods; also known quantum physics, quantum theory, including quantum field theory, is a fundamental theory in physics which describes nature at the smallest scales of energy levels of atoms and subatomic particles [154-156]. QM is the correct mathematical description of the behavior of electrons [157, 158].
- ▷ Molecular mechanics;
- ▷ The hybrid method combining quantum mechanics (QM) and molecular mechanics (QM/MM);
- ▷ Molecular dynamics; and
- ▷ Pharmacophore

Here we most describe the methods used in this work.

1.19.1 Molecular mechanics

a) Rationale of molecular mechanics methods

Appeared in 1930 by Andrews [159], the term Molecular mechanics (MM) is a formalism which was developed in order to evaluate molecular systems (geometries and properties). It is an empirical method which allows representing the potential energy of small molecules as well as large molecular systems (as the case of biological systems or material assemblies having thousands of atoms) as a function of geometric variables [152]. This is done by adjusting bond lengths, bond angles and torsion angles to equilibrium values that are dependent on the hybridization of an atom and its bonding scheme. MM methods are less complicated, fast, and are able to handle very large systems including enzymes. The MM calculations are used to determine [152]:

- (i) which conformations (or geometry) are stable;
- (ii) which conformations are the most preferred and;
- (iii) the energy differences between stable conformations.

b) Underlying principles of MM

The advanced developments in MM started in the sixties when computers were becoming more accessible and more efficient [152]. The MM calculations are based on classical mechanic principles (Itself based on the Born-Oppenheimer approximation where the motion of electrons is much faster than nuclei) [159]. In this method, a potential energy function is associated with each degree of freedom of the molecule viz: elongation of the bonds, variation of valence angles, dihedrals (rotation around a bond). The optimization of all the parameters by minimizing the energy leads to the equilibrium geometry of the various conformers and their relative energies. For molecules with a large number of conformers, there are automatic search procedures for local energy minima. In MM techniques, the molecule is treated essentially as a set of charged point masses (the atoms) which are coupled together with springs having different constants of stiffness [152]. The values of these force constants come from: experimental data or values obtained by *ab initio* methods [152]. A MM calculation results give a disposition of nuclei such that the sum of all energy contributions is minimized [152, 159]. These results mainly concern the geometry and energy of the system. In MM approach, one establishes a mathematical equations and parameters that relates a molecule's potential energy to its geometry. These equations are called "forcefield" [160], and is composed of a sum of terms that represent different types of energy contributions. The principle of the MM is therefore, to calculate the potential energy of a molecule (or a system of molecules) as a function of the coordinates of the atoms [161] by using the forcefield.

$$E_P = f(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) \quad (1.1)$$

Where \vec{r}_i is the vector position of the i^{th} atom.

1.19.2 Forcefields for MM analysis of protein and macromolecular systems

1.19.2.1 Forcefield terms

In molecular modeling field, the term forcefield (used in molecular mechanics and molecular dynamics simulations) refers to mathematical equation describing the dependence of the energy (potential energy) of a system of atoms on the coordinates of its particles [162]. This mathematical equation is an analytical form of interatomic potential energy, $U(r_1, r_2, \dots, r_N)$, with a set of parameters [163]. These

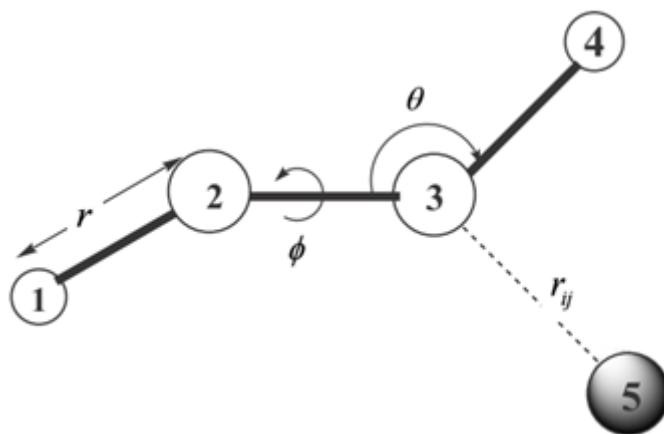


Figure 1.18: Schematic view of interactions.

parameters come from ab initio or semi-empirical quantum mechanical calculations, experimental data (X-ray, neutron, and electron diffraction, NMR, Raman, infrared, and neutron spectroscopy...) [164-168]. The optimization (i.e. accurate forcefield) of a forcefield can be done by fitting these parameters. A forcefield must be simple in order to be evaluated quickly, but also detailed to reproduce the properties (experimental data) of the molecular system studied [168, 169]. An expression of a FF is composing of:

- the bonded terms that relate to atoms that are linked by covalent bonds and;
 - nonbonded or non-covalent terms that describe the long-range electrostatic and van der Waals forces.
- See Fig. 1.18

Nowadays, many forcefields are available [163]. These FF have different degrees of complexity, and usually oriented to treat different types of molecular systems. However a typical expression of a FF may be look like as follow:

$$E_{total} = E_{bonded} + E_{non-bonded} \quad (1.2)$$

or explicitly as follow:

$$E = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2} [1 + \cos(n\phi - \delta)] + \sum_{improper} V_{imp} + \sum_{LJ} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{elec} \frac{q_i q_j}{r_{ij}} \quad (1.3)$$

where the first four terms refer to bonded terms or local contributions to the total energy (bond stretching, valence angle bending, and dihedral and improper torsions), and the last two terms (nonbonded interactions) serve to describe the repulsive and Van der Waals interactions (in this case by means of a 12-6 Lennard-Jones potential) and the Coulombic interactions. The symbols r , θ and ϕ are the bond lengths, angles and torsions respectively; r_0 and θ_0 are the reference or equilibrium values; n and δ

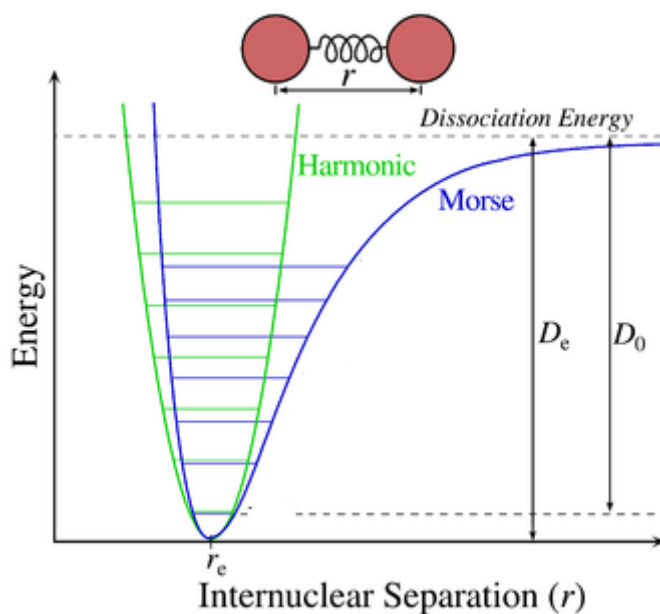


Figure 1.19: Elongation between two atoms. The Morse potential (blue) and harmonic oscillator potential (green). Unlike the energy levels of the harmonic oscillator potential, which are usually spaced by $\hbar\omega$, the Morse potential level spacing decreases as the energy approaches the dissociation energy. The dissociation energy D_e is larger than the true energy required for dissociation D_0 .

are the torsional multiplicity and phase respectively. The phases δ are usually constrained at 0° or 180° . k_b and k_a are the bonded force constant. V_n determines the height of the potential barrier. r_{ij} is the nonbonded distance between atoms i and j . ϵ_{ij} and σ_{ij} are the Lennards-Jones parameters, q_i and q_j are atomic partial charges. σ_{ij}^{12} and σ_{ij}^6 are repulsion and attraction parameters, respectively, which depend on the types of the atoms i and j .

a) Description of the bonded terms

The first term of Equation 1.3 (bond stretching) represented with a simple harmonic function that controls the length of covalent bonds. But this harmonic function supposes that the bond must not be broken. Consequently, no chemical processes can be studied. Additionally, this potential is a bad approximation for bond displacements greater than 10% of the equilibrium value. This can limit the accuracy of FF during the MD simulations for example compared to ab initio MD approach. Sometimes, other functional forms are employed to improve the accuracy of this term (usually Morse potential Fig. 1.19).

$$E_{bondstretching-Debye} = D_e \left(1 - e^{-\alpha(r_{ij}-r_0)}\right)^2 \quad (1.4)$$

D_e is the well depth; α a well width $\alpha = \sqrt{\frac{k_e}{2D_e}}$ where k_e is the force constant at the minimum of the well.

Unfortunately, these functional forms are more expensive in terms of computing time and in most cases, the harmonic approximation is reasonably good. To this, most of the existing potentials use the simple harmonic function.

The second term of Equation 1.3 which constitutes the deformation energy of the valence angles is also represented by a harmonic potential, but in some circumstances a trigonometric potential is preferable:

$$E_{anglebinding} = \frac{1}{2}k_a(\cos\theta - \cos\theta_0)^2 \quad (1.5)$$

For optimizing the fitting to vibrational spectra, other terms are often added, and the most common addition is the introduction of the Urey-Bradley (UB) potential:

$$E_{UB} = \sum_{angles} \frac{1}{2}k_{UB}(S - S_0)^2 \quad (1.6)$$

where S is the distance between the two external atoms forming the angle (for example distance between atom 2 and atom 4 in Fig. 1.18), and S_0 is equilibrium distance.

In any molecule which contains more than four consecutive atoms (atom 1, 2, 3, and 4 in Fig. 1.18), the torsional or dihedral term must be also included. Torsional motions are necessary to reproduce the major conformational changes of a molecule due to rotations around bonds. The third term of equation 5 is the energy corresponding to the deformation of the dihedral angles. Torsional motions play an important role in the determination of the local structure of a molecule and can also give an insight into the relative stability of different molecular conformations. As display in the Equation 1.3, torsional energy is usually represented by a cosine function. In this function, ϕ is the torsional angle, δ is the phase, n defines the number of minima (or maxima) between 0 and 2π , and V_n represents the height of the potential barrier. The torsional potential have others alternative representations which can be found in the scientific publications. The normal torsion terms described above are not sufficient to maintain the planarity of some particular groups (sp² hybridized carbons in carbonyl groups or in aromatic rings for example) [168]. The fourth term "improper torsion" It (called improper torsions because the four implicated atoms are not linearly linked) is needed to ensure it. This extra component describes the contribution to the energy of those out-of-plane motions. Typical expressions for the improper torsion term are given as follow:

$$E_{imp} = \sum_{impropers} \frac{k_{imp}}{2} [1 + \cos(2\omega - \pi)] \quad (1.7)$$

Or

$$E_{imp} = \sum_{impropers} \frac{k_{imp}}{2} (2\omega - \omega_0) \quad (1.8)$$

where ω_0 is the improper angle corresponding to the deviation from planarity.

b) Description of the non-bonded terms

In molecular physics, Van der Waals forces (or interactions) act between any pair of atoms belonging to different molecules or between atoms of the same molecule that are sufficiently separated. These interactions coming from the equilibrium between repulsive and attractive forces between two atoms. The repulsion component modeled by the function form $\frac{b}{r_{ij}^n}$ or $be^{-\alpha r_{ij}}$ is due to the overlap of the electron clouds of both atoms, while the attractive component that varies as $-\frac{a}{r_{ij}^6}$, result by interactions between induced dipoles. Various formulations of the van der Waals' interaction were proposed. They use different power (e.g., 9 or 10) for the repulsive part of the potential in order to give a less steep curve. The Lennard-Jones (LJ) potential (12-6 ie when n is equal to 12) is very often used to represent these interactions (term five in equation 5). Other functions used occasionally are the 9-6 LJ potential used in the COMPASS FF or even a buffered 14-7 LJ term, as employed in MMFF. A set of parameters is often define (e.g. σ_{ij} and ϵ_{ij}) for each different pair of atoms, but for practical reasons, most forcefields give individual atomic parameters (i.e. σ_i and ϵ_i), together with some rules to combine them. In the case of LJ potential for example, the well depth for the interaction between two atoms i and j is given by the geometric mean, $\epsilon_{ij} = (\epsilon_i \epsilon_j)^{\frac{1}{2}}$, while the value at which the potential is equal to zero can be given by the geometric, $\sigma_{ij} = (\sigma_i \sigma_j)$, or the arithmetic mean, $\sigma_{ij} = \frac{1}{2}(\sigma_i + \sigma_j)$, depending on the FF chosen. The last term in equation (5) serves to describe the electrostatic interactions by assigning a partial atomic charge to each nucleus and use Coulomb's law to compute their contribution to the total energy [169].

c) Special terms of forcefields

The terms described above to the contributions of potential energy appear in almost all forcefields. However, in certain cases, some other additional terms can be encounter. These additional terms describe the coupling effect between stretching, bending, and torsion and are known as cross terms. Certain force fields (such as Allinger's MM2, MM3 and MM4 [170, 171]) introduce even cross terms that involve up to three internal coordinates (eg. bond-angle-bond and angle-torsion-angle). The goal of these cross terms is to bring corrections to the intramolecular energy and allow to well reproduce the subtle physical phenomena. The following equations are examples of such cross-terms.

1.19.2.2 General classification of forcefields

The development of molecular mechanics (appeared in 1930) allowed in 1960 the apparition of first forcefields. Their main goal was to predict molecular structures, vibrational spectra, and enthalpies of isolated compounds [170]. They were, therefore, primary oriented to treat small organic molecules. Certain of them continue to be developed and used nowadays. The MM potentials developed by

Allinger's group [170-174] are the best example. The molecular mechanics application at the more complex systems (biological molecules for example) was lead to develop more widely applicable forcefields. The Assisted Model Building with Energy Refinement AMBER forcefield [175], are the good example. This forcefield contain parameters for all the atoms in the periodic table. CHARMM, AMBER , CFF, MMFF, GROMOS, and OPLS [175-179] for example are other popular forcefields. Let's mention that many of these forcefields are continuously developping and different versions are available. These forcefields was developed by academic research groups.

According to the form of Equation 1.3 and the add of the cross term above describe, forcefields are usually classified (class I, class II and class III).

a) Class I forcefields

The forcefields using an energy expression such as in Equation 1.3 i.e the internal terms in this equation are primarily harmonic or sinusoidal in nature are often called first-generation or class I forcefields.

b) Class II forcefields

Class II forcefields are referred to the forcefields that include the cross terms above mentioned. They contain cubic and/or quartic terms in the potential energy for bond and angles (anharmonic terms). These higher-order terms allow for a more accurate reproduction of quantum mechanic potential energy surfaces and such as experimental properties (vibrational spectra for example). They also introduce more parameters in the forcefield (k'_b, k'_a, \dots).

b) Class III forcefields

Accuracy of modeling of drug target systems (biomolecular systems) requires an explicit treatment of electrostatics, including electronic polarization. Thus with the advances in quantum physics, simulation algorithms and computers hardware, class III forcefields also known as polarizable forcefields were develop. As the second-generation forcefields, Class III forcefields contain also cubic and/or quartic terms in the potential energy for bond and angles. They were appeared in 1990. As above mention, many efforts to include the explicit treatment of induced electronic polarization in biomolecular forcefields. In spite of the fact that QM/MM methods have often used to treat the electronic explicit effect, now, there are variety of methods to include explicit electronic polarizability in FFs. Here, the goal is not to discuss in detail the theoretical methods by which polarizability can be included into forcefields because there are a number of scientic publications that have previously described these theoretical methods in some detail. Basic concepts of these models are briefly described. Thus, an explicit treatment of electrostatics can be performed by two manners: the charge redistribution within

each atom, either by induced dipole or by Drude oscillator (also called the charge-on spring model or shell model), and the other is based on charge flow between atoms, such as the fluctuating charge model (FQ model) which is also known as charge equilibration or chemical potential equilibration (see Fig. 1.20) [180].

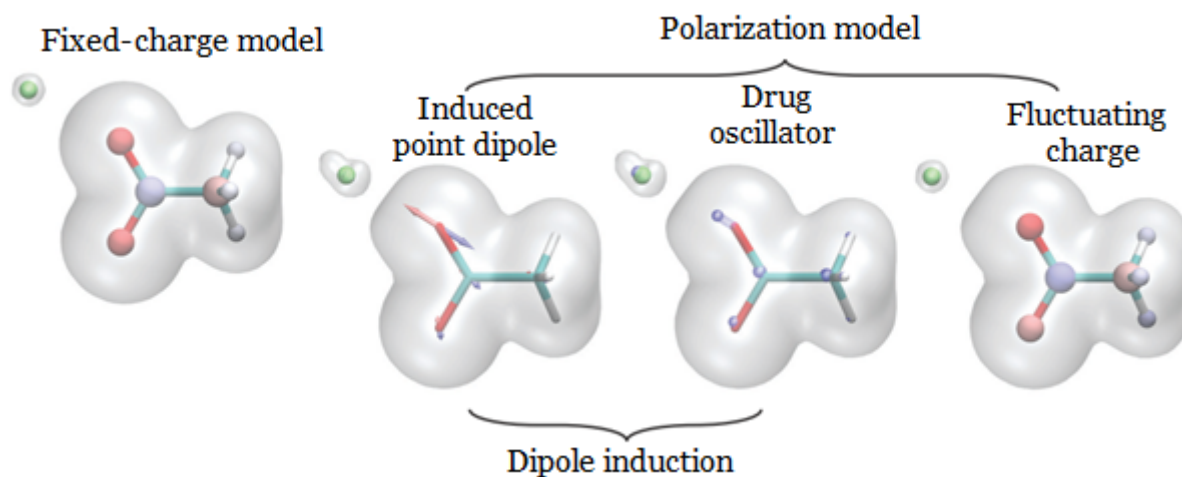


Figure 1.20: Electron cloud represented by the fixed-charge model and polarization models [180]

The induced dipoles are often determined iteratively via the self-consistent field (SCF) method which uses isotropic atomic polarizabilities following by the charge-charge, charge-dipole, and dipole-dipole interactions are computed in order to obtain the electrostatic energy of the system. AMOEBA [181] is a good example of forcefield which use these methods. In the classical Drude oscillator model [182], a Drude particle is attached to the atomic core of its parent atom via a harmonic spring with force constant kD . The displacement of the Drude particle creates an induced dipole moment. The positions of the Drude particles are solved iteratively via SCF to ensure that the Drude particles are at the ground state. Drude FF is a good example which is based on fluctuating charges. The FQ model is based on the electronegativity equalization principle. The atomic charges are redistributed to equalize the electronegativity/chemical potential at each site [181, 183]. CHEQ is a good example of forcefield which is based on fluctuating charges. Owing to the nature of polarization, these three polarization models have similar functional forms. In addition, some of the general forcefields mentioned above have also developed polarizable versions. Thus polarization has been introduced into CHARMM using either fluctuating charges or the shell model, and AMBER, OPLS and GROMOS have also been extended to include polarization [184]. A more comprehensive list of polarizable forcefields is available in science publications.

1.19.3 Description of forcefields used in the current work

1.19.3.1 The Consistent Forcefield (CFF91)

The CFF forcefield [185, 186] is one of the forcefields implemented to read protein structures and macromolecules in the accelrys software packages [187]. CFF is a class II force field [188]. The CFF91 has been developed to yield consistent accuracy of results for conformations, vibrational spectra, strain energy, and vibrational enthalpy of proteins. CFF forcefield has been implemented in the Discovery Studio software package [189] in order to read the atomic coordinates of the proteins and complexes prior to solvation. Its variant, the consistent forcefield 91 (CFF91) [190] was generally employed in the Insight II software package [191]. The CFF91 is an all-atom forcefield for proteins, nucleic acids (organic molecules in general). CFF91 describes an intramolecular part, comprising bonds and bond angles with harmonic functions and also includes a torsional potential. CFF91 is useful for hydrocarbons, proteins, protein-ligand interactions. For small models it can be used to predict: gas-phase geometries, vibrational frequencies, conformational energies, torsion barriers, crystal structures. For liquids: cohesive energy densities. For crystals: lattice parameters, atomic coordinates, sublimation energies and for macromolecules: protein crystal structures.

CFF91 proposes potential energy functions including terms of anharmonicity (cubic and quartic terms) in order to increase their accuracy. The energy associated with the elongation or compression of the bond length is given as follows in CFF91:

$$E_r = \sum_r \left[K_2 (r - r_0)^2 + K_3 (r - r_0)^3 + K_4 (r - r_0)^4 \right] \quad (1.9)$$

The bond angle bending energy in CFF91 is given as follows:

$$E_\theta = \sum_\theta \left[H_2 (\theta - \theta_0)^2 + H_3 (\theta - \theta_0)^3 + H_4 (\theta - \theta_0)^4 \right] \quad (1.10)$$

The torsional energy of dihedral angle modification energy in CFF91 is shown by equation 1.11 :

$$E_\phi = \sum_\phi V_1 [1 - \cos(\phi - \phi_{01})] + V_1 [1 - \cos(2\phi - \phi_{02})] + V_1 [1 - \cos(3\phi - \phi_{03})] \quad (1.11)$$

The out-of-plane energy in CFF91 is illustrated below:

$$E_\chi = \sum_\chi K_\chi (\chi - \chi_0)^2 \quad (1.12)$$

E_{cross} : represents the coupling between the first four terms (Equation 1.4). In CFF91 is give as follow ;

$$E_{cross} = E_{rr'} + E_{r\theta} + E_{\theta\theta'} + E_{\theta\theta'\phi} + E_{r\phi} + E_{l\phi} \quad (1.13)$$

with $E_{rr'}$ represents potential energy between two consecutive bonds of respective lengths r and r'

$$E_{rr'} = \sum_r \sum_{r'} K_{rr'} (r - r_0) (r' - r'_0) \quad (1.14)$$

$E_{r\theta}$ represents potential energy between a valence angle θ and one of the length links r defining this angle:

$$E_{r\theta} = \sum_r \sum_{\theta} K_{r\theta} (r - r_0) (\theta - \theta_0) \quad (1.15)$$

$E_{\theta\theta'}$ represents potential energy between two consecutive measurement valence angles θ and θ'

$$E_{\theta\theta'} = \sum_{\theta} \sum_{\theta'} K_{\theta\theta'} (\theta - \theta_0) (\theta' - \theta'_0) \quad (1.16)$$

$E_{\theta\theta'\phi}$ represents potential energy between two consecutive valence angles θ θ' and ϕ the central dihedral angle.

$$E_{\theta\theta'\phi} = \sum_{\theta} \sum_{\theta'} \sum_{\phi} K_{\theta\theta'\phi} (\theta - \theta_0) (\theta' - \theta'_0) \cos\phi \quad (1.17)$$

$E_{r\phi}$ represents potential energy between a dihedral angle ϕ and one of its peripheral links of length r

$$E_{r\phi} = \sum_r \sum_{\phi} (r - r_0) [V_1 \cos\phi + V_2 \cos 2\phi + V_3 \cos 3\phi] \quad (1.18)$$

$E_{l\phi}$ represents potential energy between a dihedral angle ϕ and the central link of length l :

$$E_{l\phi} = \sum_l \sum_{\phi} (l - l_0) [F_1 \cos\phi + F_2 \cos 2\phi + F_3 \cos 3\phi] \quad (1.19)$$

while the non-bonded interactions are respectively represented by a Lennard-Jones potential and a Coulomb potential as follows:

$$E_{vdw} = \sum_i \sum_{j>i} \left[2 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^9 - 3 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (1.20)$$

$$E_{nonbonded} = \sum_i \sum_{j>i} \frac{q_i q_j}{R_{ij}} \quad (1.21)$$

In the above equations, K_x , H_y , and V_z are force constants; the r_0 and θ_0 respectively represent the equilibrium bond length and angle; ϵ represents the dielectric constant, q_i is the partial charge on atom i and r_{ij} is the interatomic distance between atoms i and j . The parameters of the bonded and nonbonded potential parts of the standard CFF91 MM forcefield (ϵ_{ij} and σ_{ij}) are defined in the Discover user guide [192]. CFF91 is particularly adapted to MM calculations in the gas phase and gives interesting results for vibrational spectra of a molecule, when compared to other forcefields.

1.19.3.2 The Merck Molecular Forcefield (MMFF94)

The potential energy of a molecular system is usually expressed as a function of atomic coordinates. In the present work, the MMFF forcefield that was implemented in the MOE [193] software packages has been used to minimize the molecules of the built virtual library. MMFF was developed by Halgren [194-196] and has been aimed more at drug-like organic compounds than at proteins. The potential energy of this forcefield may be written as:

$$E(x) = E_{str} + E_{ang} + E_{stb} + E_{oop} + E_{tor} + E_{vdw} + E_{ele} + E_{sol} + E_{res} \quad (1.22)$$

Each energy term itself, being a sum of several kinds of atomic interactions, such that; E_{str} : represents the bond stretching energy; E_{ang} : represents the bond angle bending energy; E_{stb} : represents the stretch-bending energy; E_{oop} : represents the out-of-plane energy; E_{tor} : represents the torsional energy of dihedral angle modification energy; E_{vdw} : represents the van der Waals energy; E_{ele} : represents the electrostatic interaction energy; E_{sol} : represents the implicit solvation energy, and E_{res} : represents the restraint energy. Among the MMFF, exist MMFF94 which is a particular all-atom forcefield parameterized for small organic molecules in which partial charges are based on bond-charge increments and is particularly adapted for use with the Generalized Born solvation models, with conjugated nitrogen atoms considered to be tetrahedral [194]. In the MMFF94s variant of this forcefield, conjugated nitrogen atoms are less tetrahedral than in MMFF94, while in MMFF94x conjugated nitrogen atoms are planar [195, 196]. Like CFF91, MMFF94 is a class II forcefield.

MM is presently the only practical method for the calculations on very large molecules or for conformational searching on highly flexible molecules. Although MM calculations are extremely useful, they consider essentially the position of the nuclei and therefore cannot fully represent chemical reality. Also, MM provides no information on electronic structure, and furthermore cannot be used when the molecule is not in its ground state, or when covalent bonds are being broken or formed. Thus, in other projects of research, MM techniques is often coupling by QM techniques [197, 198]. In this case the method is called the Hybrid quantum mechanics/molecular mechanics (QM/MM) method.

This Hybrid method can be used in order to study well the molecular systems [199, 200] and have become the method of choice for modeling reactions of biomolecular systems in condensed phases [201, 202].

1.20 Conformational analysis

Conformational analysis can be used to define the study of the conformations of a molecule and their influence on its properties. Le Bel [203] and Hoff [204] are considered as the founding fathers of the conformational analysis. Modern MM programs often have a graphical interface that allows the user to enter the structure of a molecule and minimize its energy in that conformation [205]. Conformational analysis allows for the calculation of the relative energy associated with the conformation of a molecule [206]. To perform a conformational search, one find a method to systematically generate all possible structures (all conformations) determines the energy of each conformation as described by the potential energy function. The energies of conformations can be calculated by using QM methods or any one of the MM forcefields previously described. Often, the conformational energy calculated is function of the applied forcefield. i.e. for the same conformation, the value of energy can change [207]. Many techniques to carry out a conformational search are described in the literature [208, 209]. In all of them, the energy evaluation of the molecule via the potential energy function (forcefield) is an important step [210]. In a nonlinear molecular system without constraint and having a number of atoms $n > 2$, the number of degrees of freedom is equal to the number of internal coordinates i.e. $3n-6$. In the conformational search, the structures are generated by varying the most flexible degrees of freedom, namely the dihedral angles [206]. In conformational research based on the rotation of the group of atoms around single bonds (torsion drive), one vary systematically all the dihedral angles of a molecule. Thus, for example, a molecule with three dihedral angles generates 216 conformations if the increment of angle is of 60° . 46656 conformers with six dihedral angles and more than 10 million for 9 dihedral angles. The increment of 30° is often used. The goal of the conformational analysis is to find as much minima as possible and to calculate the Boltzmann population.

$$N_i = N_0 \exp\left(-\frac{E_i}{K_b T}\right) \quad (1.23)$$

$$N_j = N_0 \exp\left(-\frac{E_j}{K_b T}\right) \quad (1.24)$$

so

$$\frac{N_i}{N_j} = \exp\left(-\frac{\Delta E_{ij}}{K_b T}\right) \quad (1.25)$$

The "Boltzmann formula" gives the ratio between the number of particles, respectively N_i and N_j , occupying the energy states E_i and E_j :

E_j : energy of the molecule;

K_b : Boltzmann constant;

T : Temperature (Kelvin).

Conformational analysis is, therefore, an effective approach to determine the stable conformations of a molecular structure corresponding to the minima of its intramolecular energy. The calculated energy measures the difference between the energy of the molecular structure under consideration and that of a hypothetical structure of which all coordinates would take their reference values. The values of this energy are all very useful when it comes to comparing the relative stability of conformations or stereoisomers of the same molecule [211].

1.20.1 Energy minimization

In any drug design study, the protein and ligand interactions are carried out in their lowest energy state. Usually the proteins are selected from Protein Data Bank and the ligands can be either from literature or a database. In such cases, where ligands are derived from literature, certain steps need to be undertaken to perform ‘protein-ligand’ interactions studies. Of all the steps involved in these studies, the first and foremost major step is often energy minimization and conformation search analysis. Also known as better known as geometry minimization, energy minimization is used to determine the lowest or the minimum energy conformation for a molecule under study. When a molecular structure is built, it usually needs to be refined to bring it to a stable, sterically acceptable, conformation. This is all the more true after building in three dimension of some structures when the process of adding fragments generates serious atom clashes. The refinement process is also refer to minimization (or optimization), and is an iterative procedure in which the coordinates of the atoms are adjusted so that the energy of the structure is brought to a minimum. i.e. start with an initial geometry and then change that geometry to find a lower-energy shape. This often results in finding a local minimum of the energy as depicted in Fig. 1.21.

The structure with the lowest energy is considered to have the most stable arrangement, and by definition the optimum geometry. Minimization generally results in a modeled structure with a close resemblance to a real physical structure. The ability to compute the energy of a structure is a necessary part of the minimization process, and is an extremely important aspect of a modeling system. To facilitate calculations, it is generally considered that the variable term of this energy depends on the construction of the molecule and the arrangement of its atoms: this is the principle of empirical methods (MM, MD). In most of these methods, the interactions with the solvent are not taken into account. Only interactions between the constituent atoms of the molecule are evaluated [212]. The total energy function has many minima and maximas. There is no general mathematical method to find the overall minimum of this function. Numerical analysis methods are used to find local minima.

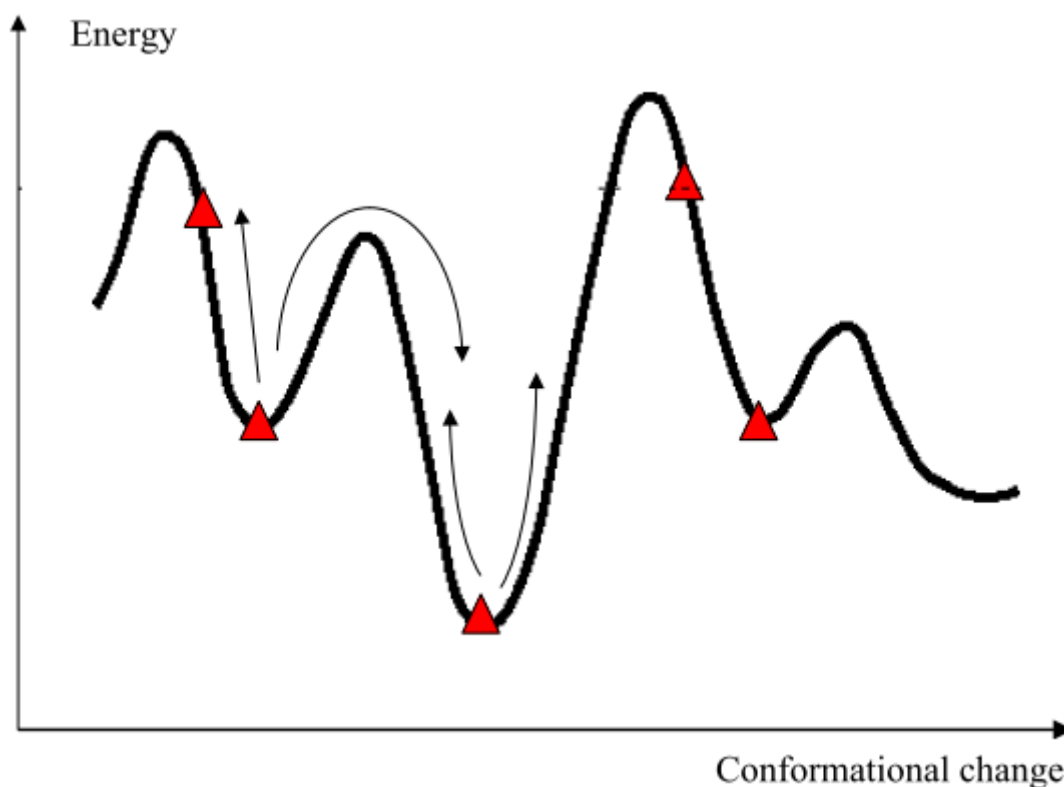


Figure 1.21: Representation of the energy of all possible conformers of a molecular system. The energy of the system is calculated using the forcefield. The conformation of the system is altered to find lower energy conformations through a process called minimization.

1.20.2 Energy minimization methods

Conformation search algorithms are an automated means for generating many different conformers and then comparing them based on their relative energies. In drug design and discovery studies some algorithms are often used.

1.20.2.1 The steepest descent approach

The steepest descent (SD) method is also called the biggest slope method. In this method, the gradient vector \vec{g} points in the direction where the function increases most, i.e. the function value can always be lowered by stepping in the opposite direction [213]. The descent direction, i.e. the direction in which atoms are moved down towards the minimum, in step k in SD is dictated by the gradient g_k . When the function starts to increase, an approximate minimum may be determined by interpolation between the calculated points. At this interpolated point, a new gradient is calculated and used for the next line search. If the line minimization is carried out accurately, it will always lower the energy function value. It is therefore guaranteed that this function approaches a minimum. The steepest descent path therefore oscillates around the minimum path, as illustrated in Fig. 1.22

As show in Fig. 1.22, when the minimum is approached, the rate of convergence slows down based

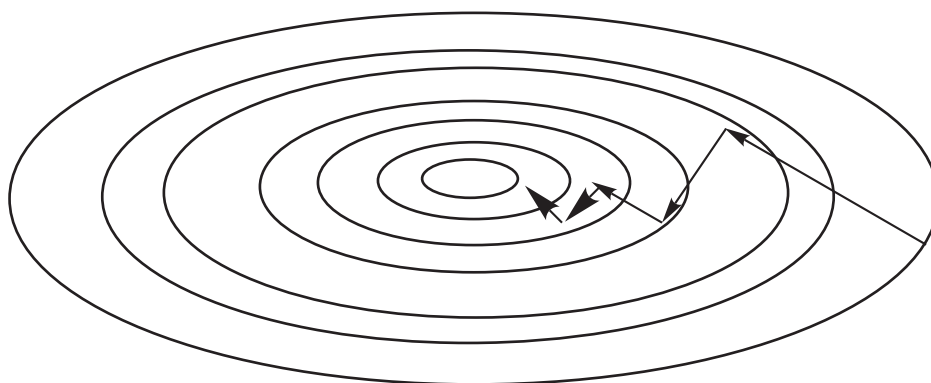


Figure 1.22: Steepest descent minimization [213]

on the fact that the steepest descent will actually never reach the minimum, it will crawl towards it at an ever decreasing speed. SD is very simple, requires only storage of a gradient vector and it is one of the few methods that is guaranteed to lower the function value.

1.20.2.2 The conjugate gradient method

The Conjugate Gradient (CG) method is an improvement of SD [214]. The pitch is adjusted at each cycle to obtain a better decrease in energy. It achieves a minimum in fewer steps than the SD method and requires significant computation time when the starting structure is far from the minimum. The initial step in the CG algorithm is equivalent to a steepest descent step, but subsequent searches are performed along a line formed as a mixture of the current negative gradient and the previous search direction. Often, the minimization beginning with Steepest Descent followed by Conjugate Gradient. This protocol is implemented in Accelrys Discovery Studio software [189].

1.20.2.3 The Newton-Raphson method

Newton's method Raphson evaluates the second derivatives of molecular energy with respect to geometric parameters and thus converges more rapidly. The programming of this procedure is more difficult than that of the SD and CG methods. Only the gain in computation time and accuracy is so important that almost all molecular mechanics programs have adopted it. This method is effective only when the geometry of the initial conformation is close to the minimum [215].

1.20.2.4 The simulated annealing method

Developed by Kirkpatrick et al. [216] Simulated Annealing (SA) [217] works are based on the non-linear cooling process of glass where various steps of changes precede the formation of a crystal. Here it is assumed that different rates of cooling lead to different formations of glass. The system is

first ‘melted’ at a high effective temperature. Annealing is the lowering of the temperature via slow steps to ensure the system ‘freezes’ so that there are no other changes. This technique performs a check at every interval of cooling where if the energy is lower, the step of cooling is accepted.

1.20.2.5 The simplex method

The simplex is referring as a geometric figure defined by a number of points higher by one compared with the number of factors (or dimensions of the, factor space). In the 2D-dimensional factor space the simplex is a triangle and in the 3D-dimensional space it is a tetrahedron. In this method, the ”simplex” is moved in the factor space depending on the results of experiments performed for the factor values corresponding to the simplex vertices. There is always a possibility that the optimum found is a local optimum. This method is interesting when the starting structure is close to the global minimum [218].

MATERIALS AND METHODS

2.1 Data collection

Before embarking any calculations in the current work, we have described rational and computer-aided drug discovery approaches for the design of promising FP-2 and FP-3 inhibitors, with a focus on a variety of structure-based and ligand-based modeling approaches [145]. Moreover, the key features of ligand recognition against these targets have been emphasized [145].

2.2 The complexation model employed

Scheme 1 display the workflow of different steps involved for the computer-aided drug design of novel PEP analogues

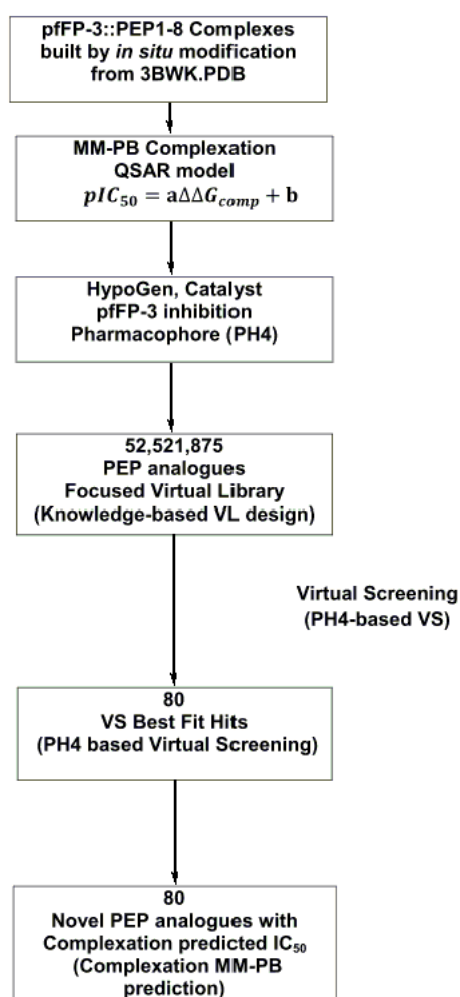


Figure 2.1: Novel PEP analogues design methodology workflow.

2.2.1 Preparation of the structure 3BWK

Structures downloaded in Protein Data Bank have often many errors [219, 220]. The correction of these structures (or preparation) is a first step in the field of computer-aided drug design. Due to this, the high-resolution X-rays crystal structure of the native inhibitor K11017 (Mu-Leu-Hph-VSPH where VSPH: phenyl vinyl sulfone; Hph: homophenylalanyl; Mu: morpholino urea:) bound to the plasmodial FP-3 (Protein Data Bank [220] entry code 3BWK) has been prepared by using an algorithm that consisted in reading the protein and ligand atoms with the CFF91 forcefield, with no Morse and cross terms. In this step wise process, the SD method is employed initially, followed by the CG (smart minimizer method) as follows:

Step 1: For heavy atoms of both protein and ligand, 500 iterations with the SD method until the maximum derivative is less than $3.00 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, followed by 1000 iterations using the CG method until the maximum derivative is less than $1.00 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$.

Step 2: For heavy atoms of ligand only, 500 iterations with the SD method until the maximum deriva-

tive is less than $1.00 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, followed by 2000 iterations using the CG method until the maximum derivative is less than $0.20 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$.

Step 3: For heavy atoms of ligand only, 1000 iterations with the SD method until the maximum derivative is less than $1.00 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, followed by 1000 iterations using the CG method until the maximum derivative is less than $0.20 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$.

Step 4: For heavy atoms of ligand only, 500 iterations with the SD method until the maximum derivative is less than $1.00 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, followed by 3000 iterations using the CG method until the maximum derivative is less than $0.20 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$.

Step 5: For general relaxation of the entire system, 500 iterations with the SD method until the maximum derivative is less than $1.00 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, followed by 2000 iterations using the CG method until the maximum derivative is less than $0.50 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, then 2000 iterations using the CG method until the maximum derivative is less than $0.10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, 2000 iterations using the CG method until the maximum derivative is less than $0.05 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ and finally 10000 iterations using the CG method until the maximum derivative is less than $0.001 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$. The vibrational energy of this final structure is then calculated.

This methodology has been applied in order to preserve the three dimensional structure of the enzyme (3BWK) use as start point for the calculations.

2.2.2 Biological activities of compounds included in the training and test sets

The chemical structures and biological activities (IC_{50}^{exp}) of training and validation sets of peptidomimetic (PEP) inhibitors of PfFP3 used in the current work were taken from the reported study by Weldon et al. [221]. This covered a broad range of activities ($60 \text{ nM} \leq IC_{50}^{exp} \leq 47230 \text{ nM}$), suitable for QSAR model construction. The data set has been divided into a training set of eight (08) molecules for generating the subsequent QSAR models and a test set of two (02) molecules for evaluating the predictive quality of the models. Table 2.1 depicts pIC_{50}^{exp} values (nM) of the data set while Fig.2.2 depicts several representative skeletons of the data set.

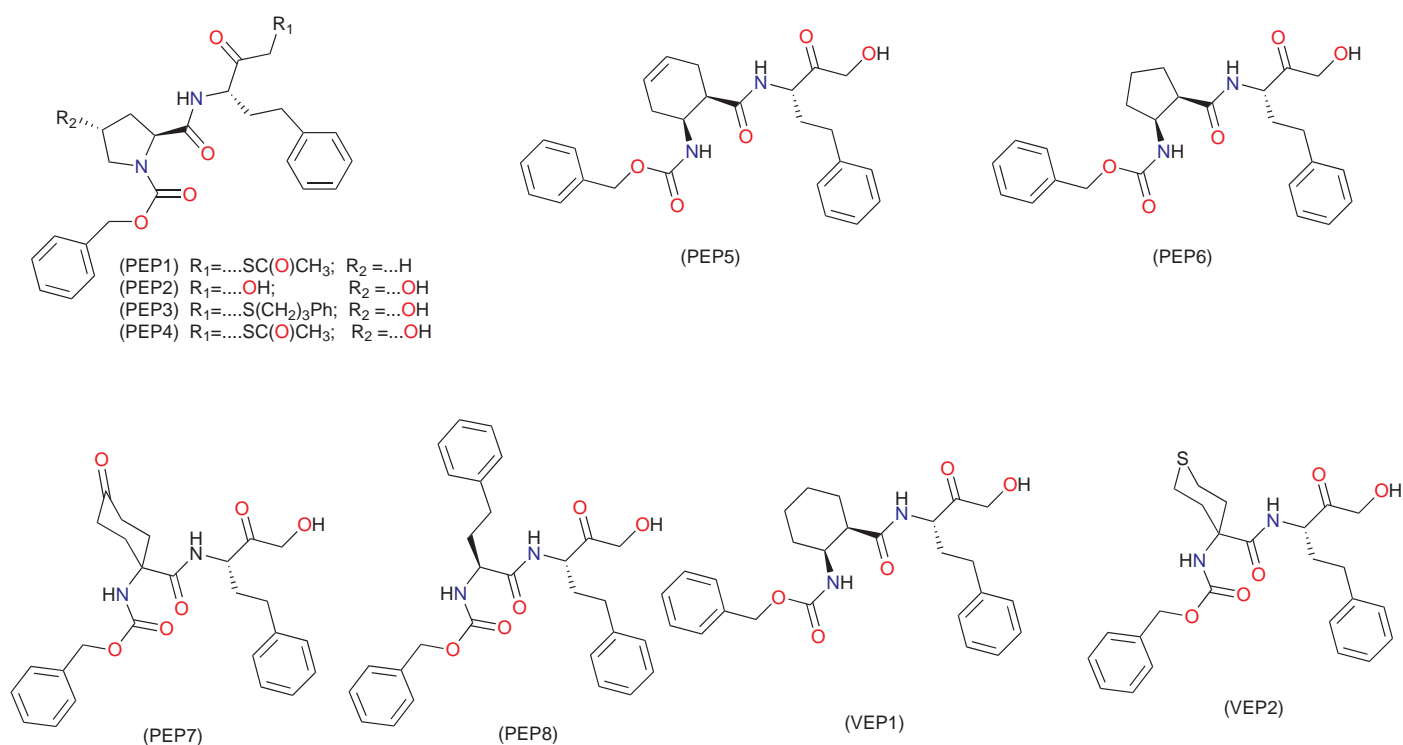


Figure 2.2: The chemical structures of training (PEP x) and validation (VEP x) sets of inhibitors obtained from literature [221].

Table 2.1: Training and validation sets of PEP inhibitors obtained from the literature.

| Training set ^a | M_w ^b (g mol ⁻¹) | $IC_{50}^{exp,c,d}$ (nM) |
|-----------------------------|--|--------------------------|
| PEP1 (ref) | 482 | 36360 |
| PEP 2 | 440 | 60 |
| PEP 3 | 574 | 520 |
| PEP 4 | 498 | 3560 |
| PEP 5 | 450 | 20180 |
| PEP 6 | 438 | 23900 |
| PEP 7 | 466 | 47230 |
| PEP 8 | 488 | 11910 |
| Validation set ^a | M_w ^b (g. mol ⁻¹) | $IC_{50}^{exp,c,d}$ (nM) |
| VEP1 | 452 | 540 |
| VEP2 | 470 | 8220 |

^aSee Fig. 2.2 for the chemical structures of training and validation. ^bMolar weight of PEP analogues.

^cThis is IC_{50}^{exp} expressed in nanomolar concentration. ^dFor potencies of PEP analogues against FP-3, Ref. [221] is referred.

2.3 Model building by *in situ* modification

Three dimensional (3D) molecular models of the free enzyme (E), free inhibitors (I) and the enzyme-inhibitor complexes (E:I), have been constructed from a high-resolution X-rays crystal structure of the native inhibitor K11017 (Mu-Leu-Hph-VSPH where VSPH: phenyl vinyl sulfone; Hph: homophenylalanyl; Mu: morpholino urea:) bound to the plasmodial FP-3 (Protein Data Bank [220] entry code 3BWK) via the graphical user interface available in molecular modelling program Insight-II [222]. Chain A of the enzyme was selected in the computations. The structures of FP-3 and enzyme-inhibitor (E:I) complexes were considered in the computations to be at the pH of 7 with neutral N- and C-terminal residues and all protonated and ionized residues charged. Co-crystallized water molecules, contained in the initial structure were deleted. The inhibitors have been built into the crystal reference structure by *in situ* modifying of derivatives groups in the molecular scaffold of the native inhibitor K11017. All rotatable bonds of the replacing residues have been subject at an exhaustive conformational search coupled with a careful gradual energy minimization of the modified inhibitor and the active site residues of FP-3 located in the vicinity of the inhibitor ($\leq 5\text{\AA}$). This methodology has helped to identify low-energy bound conformations of the modified inhibitor and leads to various low-energy structures of the E:I complexes which have been then carefully refined by minimization of the whole complex. This process has been successfully used in many works on complexation models [223-232].

2.4 Molecular mechanics

Simulations of the models of inhibitors, FP-3 and E:I complexes were carried out by molecular mechanics (MM) using an all-atom representation and the class II consistent force field CFF91 [222]. All MM calculations used a dielectric constant of 4 for representing dielectric shielding effects in the proteins. The optimisation (energy minimisation process) of the free E, free I and E:I complexes were carried out by a gradual relaxation of the structures, beginning by adding H-atoms, then the residue side chain heavy atoms, and ending up with the relaxation of the protein backbone. In all the geometry optimizations, a sufficient number of steepest descent and conjugate gradient iterative cycles have been used with the convergence criterion for the average gradient set to of $0.01 \text{ kcal mol}^{-1} \text{\AA}^{-2}$.

2.5 Conformation search of free inhibitors by molecular dynamics

In the current work, we have used Molecular Dynamics (MD) calculations. These techniques have been used in order to conformational research and to refine structures. MD is a method that studies the movements and evolution of the spatial configuration of molecular systems. In practice, we work

by solving the classical equations of Newton's motion, given a potential energy function and its associated forcefield.

2.5.1 Principles of molecular dynamics

MD is an *in silico* technique which simulates, from the laws of classical mechanics, the trajectories of atoms or molecular systems in the crystalline phase, in solution or in the gas phase [233, 234]. This is done by using the principles of Newton's classical mechanics [235]. It gives the evolution of a system in conformational space over time. A numerical resolution of the classical equations of motion is used in order to calculate the energy of the system along its path over the potential area of the system. In this technique, intramolecular motions are simulated and visualized in real time. These motions correspond to the vibrations around a minimum or the transition from a minimum to another minimum level of energy. MD simulations, therefore, consist of calculating the positions and speeds of a system of atoms or molecular systems [236]. In classical MD, each atom of the molecule is considered as a point mass whose motion is determined by the set of forces exerted on it by other atoms as a function of time [237]. In MD simulations, we need a set of initial conditions (positions and velocities of each particle), a good model to represent the forces acting between the particles (by using forcefields presented in the previous section). Then we need to solve the classical equation of motion: (equation 2.1).

$$\sum \vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i(t)}{dt^2} = -\frac{d \vec{E}_p}{r(t)} \quad (2.1)$$

Where: \vec{F}_i : force vector acting on the atom i ; m_i : mass of the atom i ; \vec{a}_i : acceleration vector of the atom i , and \vec{r}_i : The position vector of the atom i . \vec{E}_p is potential energy of the system

As illustrated by the Newton's equation, the velocity and direction of atomic motion depend on the forces acting on the atoms.

In practice, atoms have their own initial velocity that conforms to the total kinetic energy of the system. This kinetic energy depends on the simulation temperature of the system 2.2. This kinetic energy allows the system to cross potential barriers that other minimizers cannot cross. Velocities are thus assigned to atoms by random draw following a Maxwell distribution:

$$E_c = \frac{1}{2} \sum m_i \langle v_i \rangle^2 = \frac{3}{2} NK_b T \quad (2.2)$$

With:

$\langle v_i \rangle^2$: value of the square average of the speed;

N : number of atoms in the system;

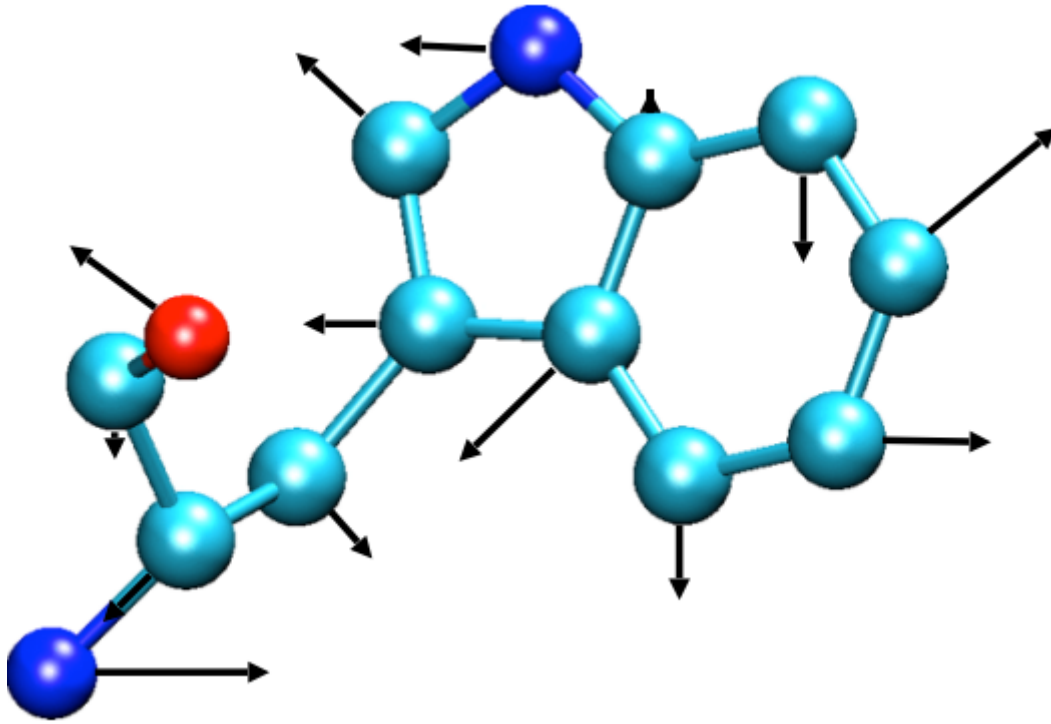


Figure 2.3: Forces acting on atoms of the molecule.

K_b : Boltzmann constant; and

T : absolute temperature

As we already said, the equation (2.1) can only be solved numerically. So we need to discretize the trajectory and use an integrator to advance over small time steps:

$$r_i(t_0) \rightarrow r_i(t_0 + \Delta t) \rightarrow r_i(t_0 + 2\Delta t) \rightarrow \dots r_i(t_0 + n\Delta t) \quad (2.3)$$

The following properties of such an integrator are usually required:

- ▷ Minimal need to compute the forces: This is the most time-consuming step in the simulation, thus any algorithm requiring more than one cycle of evaluation of the forces per time step will not be accurate.
- ▷ Good stability if large time steps Δt are used.
- ▷ Good accuracy.
- ▷ Good conservation of energy and momentum.
- ▷ Time-reversibility.

The immediately apparent solution would be to use a simple Taylor expansion, so:

$$r_i(t_0 + \Delta t) = r_i(t_0) + \frac{dr_i(t_0)}{dt} \Delta t + \frac{1}{2} \frac{d^2 r_i(t_0)}{dt^2} \Delta t^2 + 0(\Delta t^3) \quad (2.4)$$

Unfortunately, this algorithm is unstable and inaccurate. Verlet has proposed a better solution. If we sum the Taylor expansions for $+\Delta t$ and $-\Delta t$, the terms in Δt , Δt^3 , etc. cancel and we obtain:

$$r_i(t_0 + \Delta t) = -r_i(t_0 - \Delta t) + 2r_i(t_0) + a_i(t_0) \Delta t^2 + 0(\Delta t^4) \quad (2.5)$$

The velocities are not used in the algorithm, but they can be obtained as:

$$v_i(t_0) = \frac{1}{2\Delta t} [r_i(t_0 + \Delta t) - r_i(t_0 - \Delta t)] \quad (2.6)$$

And the velocity-Verlet algorithm:

$$r_i(t_0 + \Delta t) = r_i(t_0) + v_i(t_0) \Delta t + \frac{1}{2} a_i(t_0) \Delta t^2 \quad (2.7)$$

$$v_i(t_0 + \Delta t) = v_i(t_0) + \frac{1}{2} [a_i(t_0) + a_i(t_0 + \Delta t)] \Delta t \quad (2.8)$$

A general scheme of the MD procedure is display in Fig. (2.4). The starting point refers to the definition of the forcefield and the molecular topology, which is the essence of the proposed model.

Molecular Dynamics

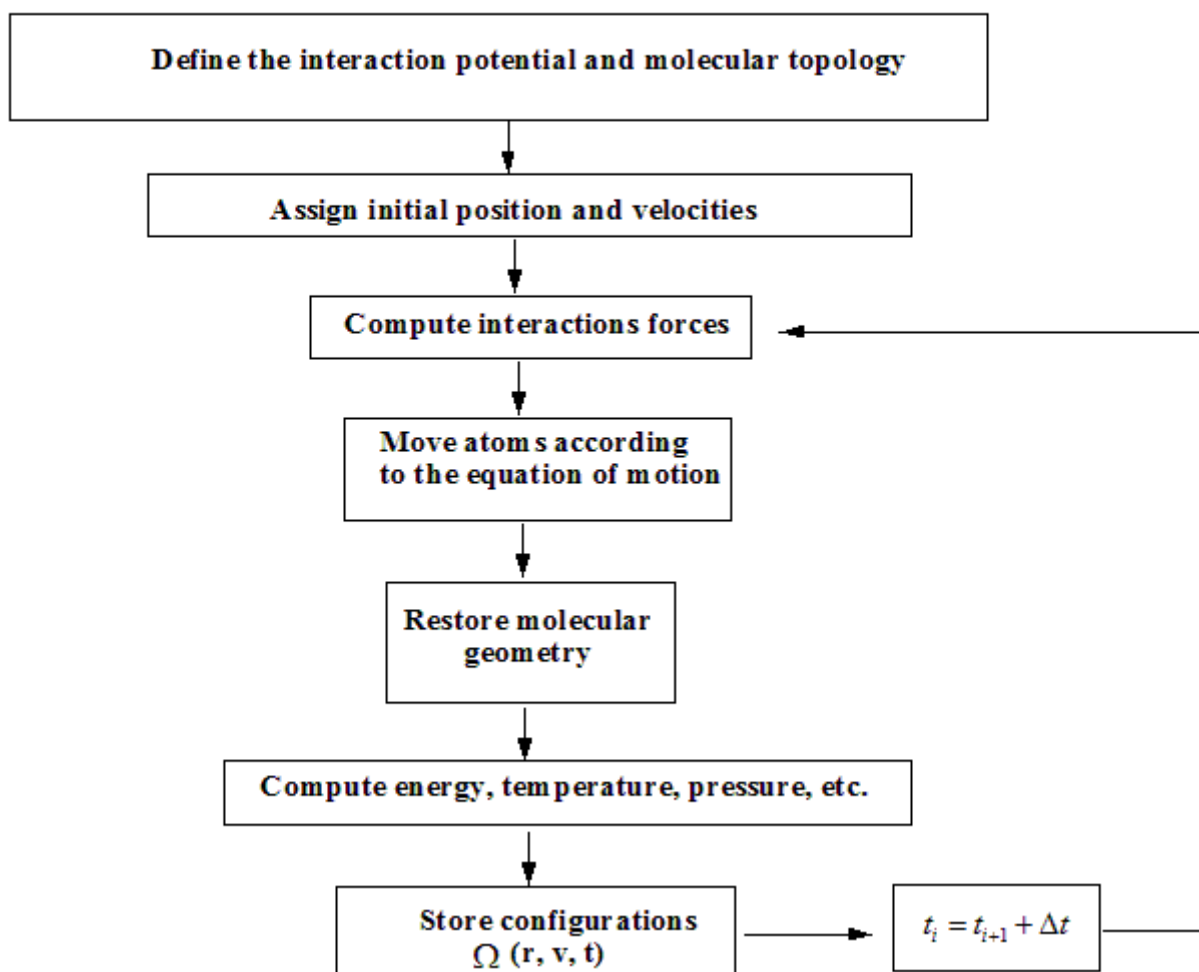


Figure 2.4: Steps of the molecular dynamics simulation procedure.

In the current work, the conformation of each free inhibitor was obtained from its bound conformation in the E:I complex, which had been previously obtained by the gradual relaxation to the nearest local energy minimum. The geometries of the free ligands were then obtained by the molecular dynamics method. These simulations were performed with the Insight II software [222] and the CFF91 forcefield [222]. The temperature T has been gradually increased from 30 to 300 K in 30 K steps. The equations of motion were integrated according to the verlet algorithm with a time increment $\tau=1$ fs. At each value of T , the system is heated for 500 fs. At the end of the calculation, 200 optimized conformers are generated. After the analysis of the results, the lowest conformer energy is selected and then optimized again with a dielectric constant $\epsilon = 4$

2.6 Solvation Gibbs free energy

Drug targets in the body exist in an aqueous environment and the drug needs to cross that environment in order to reach its target. Therefore, both the drug and its target are solvated with water molecules before they meet each other. It is well accepted that the role of solvent in biochemical processes is crucial. The estimation of hydration energy, binding energy of ligands or protein residues heavily depends on a good description of the solvent behavior. Specific water molecules and their placement within protein active sites have been shown to make specific contributions to the energetics of protein-ligand. Thus, it is important to take account the effects of these water molecules in the ligands binding interaction in order to better explain the observed activities compared to experimental ones. Proper theoretical description of the solvation effects of water within a ligand-binding pocket is a significant computational challenge. Several studies have shown that water molecules have an important role in the affinity between the ligand and his target. When a ligand binds, in the active site of a target, solvent molecules within the active site rearrange or become displaced. These rearrangements of water molecules within the binding pocket of the target, affect the binding free energy. For example, water solvation of hydrophobic pockets is very unfavorable energetically because water in this situation cannot form hydrogen bonds. Conversely, the expulsion of water in such regions of the protein leads to an improvement in ligand-protein binding affinity Fig. 2.5.

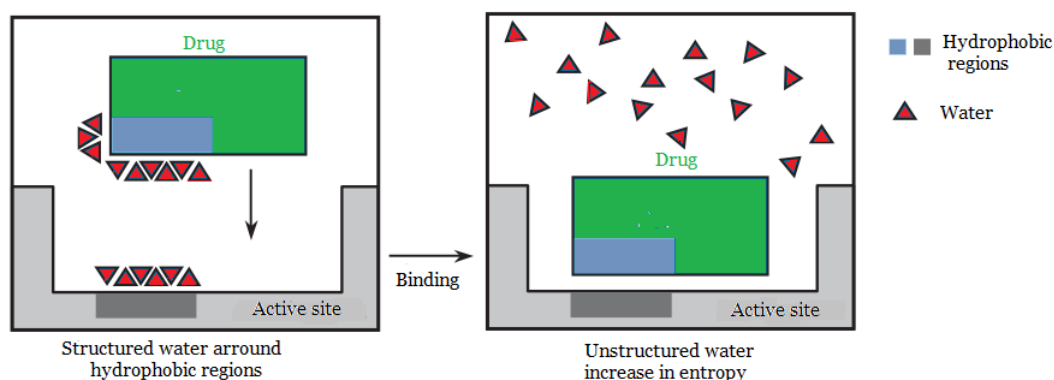


Figure 2.5: Desolvation of a drug and its target active site prior to binding.

It is clear that an accurate description of the thermodynamics of ligand-protein association requires a consideration of all interacting species at every stage of the process. Figure 2.6 shows the thermodynamic cycle of protein-ligand interactions.

In the field of computer-aided drug design, the methods have been developed for describing accurately the effects of solvent. These *in silico* methods describe the physics involved and calculate the solvation energy of the solute (ligand, protein, complex...) [238]. There are two principal types of solvation which are more known: explicit and implicit solvation [239]. We are use implicit solvation method in our calculations.

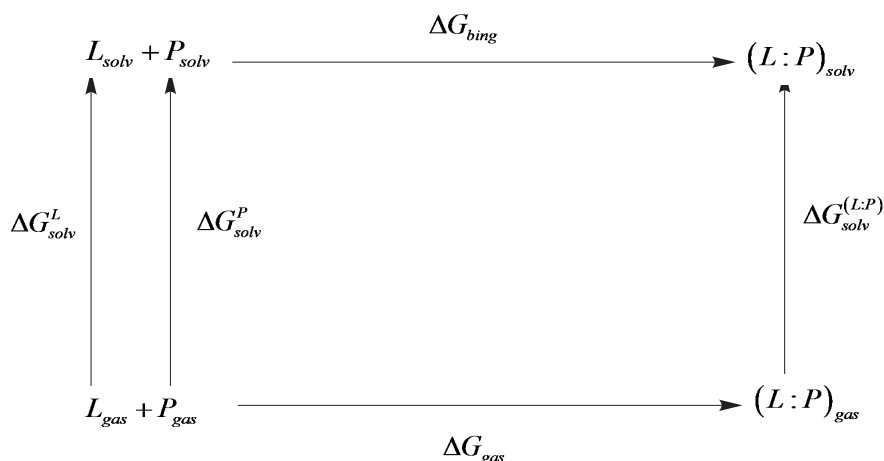


Figure 2.6: The thermodynamic cycle of protein-ligand interactions.

Implicit solvation is the continuum model. Solvent molecules disappear and they are substituted by a continuous dielectric of infinite extent that surrounds a cavity containing the solute molecule (see Figure 2.7). The molecular system (solute) is placed into a cavity described by the surface tension of continuous medium (the solvent) characterized by macroscopic properties. Implicit solvent models are more approximate but incur less computational cost than explicit solvent models [240-242]. Calculations with implicit models can be performed in MM [243] or with QM [244]. The physical phenomenon that is represented (but not exclusively) in implicit solvent approaches is the linear response to the electric field generated by the solute. In this work, the PB Equation has been used to calculate ligand solvation and estimate ligand binding free energies by calculating the total energy of the protein-ligand complex and then subtracting the solvated energy of the protein and ligand separately. This commonly used protocol is known as the Molecular Mechanics-Poisson-Boltzmann (MM-PB) [245, 246]. MM-PB is a scheme currently used for the determination of binding free energies for protein-ligand complexes [247].

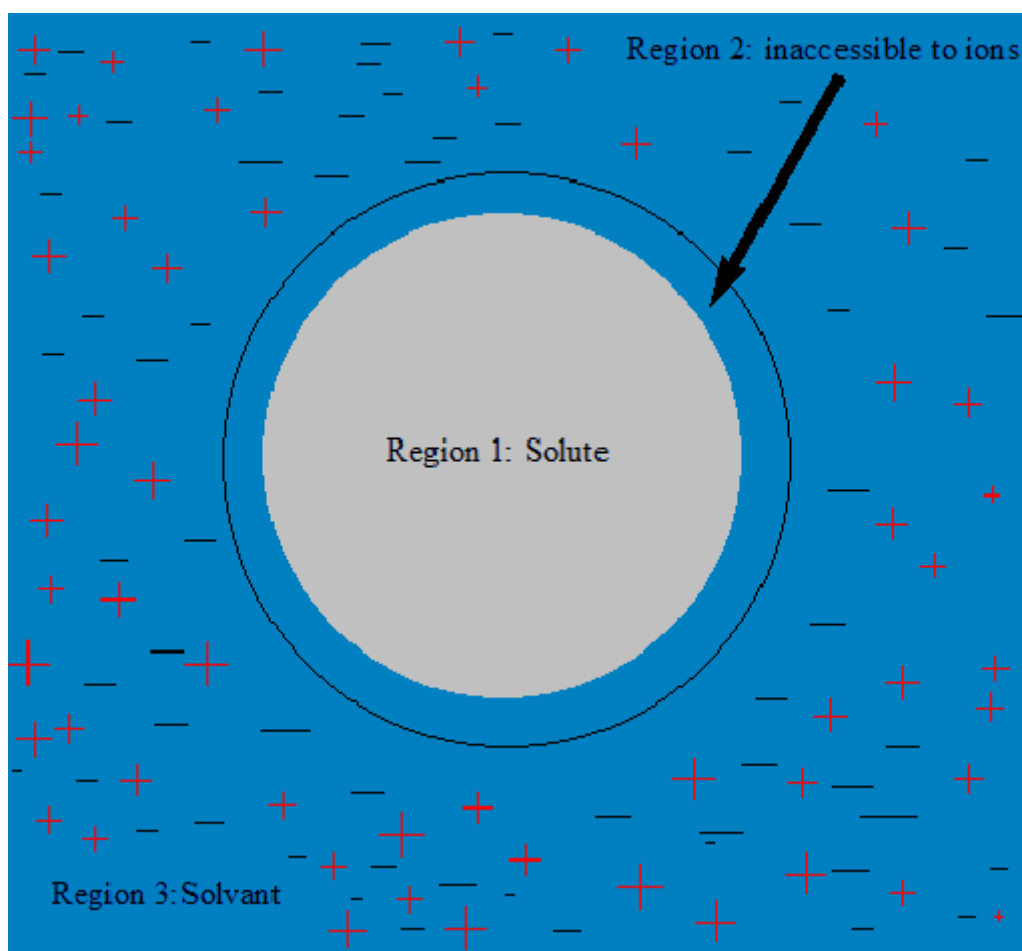


Figure 2.7: Implicit solvation model

In Figure 2.7

- ◆ Region 1 is a cavity within the solvent. In this cavity, the immobile solute (inhibitor, enzyme, or enzyme-inhibitor complex) is placed. This region is characterized by its dielectric permittivity ϵ_1 and its charge volume density ρ_1 . Usually ϵ_1 varies between 2.5 to 4 [247]. Charges of solute are treated by a forcefield.
- ◆ Region 2 is located in the immediate vicinity of the solute. Occupied by the solvent, this region is called the exclusion zone because the ions do not have access to it. Therefore, the density ρ_2 is zero. In this region a dielectric permittivity constant of the solvent is ϵ_2 (ϵ_2 is equal to ϵ_3). For an aqueous solvent, ϵ_2 varies between 78 and 80 [246].
- ◆ Region 3 is also occupied by the solvent. Here the presence of mobile ions that are modeled by a charge density ρ_3 is notable.

The free energy of solvation has three terms each relate to a purely theoretical process: the creation of a cavity (ΔG_{cav}) within the solvent to accommodate the solute molecule and embedding of the solute molecule into the cavity leading to van der Waals (ΔG_{vdW}) and electrostatic (ΔG_{elec}) interactions between solute and solvent (see equation 2.9).

$$\Delta G_{solv} = \Delta G_{cav} + \Delta G_{vdW} + \Delta G_{elec} \quad (2.9)$$

The first term of equation 2.9 represents the entropic cost of reorganization of the solvent molecules. It also represents the cost of the pressure resistance of the solvent to form the cavity and high values of this energy is an indicator that the drug could be ineffective. In some cases, a polar group of a drug can be removed in order to lower its energy of desolvation. The first two terms are generally expressed by a non-polar (np) term (see equation 2.10)

$$\Delta G_{solv,np} = \Delta G_{vac} + \Delta G_{vdW} \quad (2.10)$$

2.6.0.1 The Poisson-Boltzmann Equation with Non-polar Surface Area (PBSA)

In a homogeneous and isotropic dielectric medium, ϵ is constant. When the medium is non-homogeneous, ϵ depends on the position \vec{r} i.e $\epsilon = \epsilon(\vec{r})$. The Poisson equation relates the scalar electric potential ϕ to the charge density ρ is written:

$$\vec{\nabla} \left[\epsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \right] = -4\pi \rho(\vec{r}) \quad (2.11)$$

where $\phi(\vec{r})$ is the electrostatic potential. $\rho(\vec{r}) = \rho_1(\vec{r}) + \rho_3(\vec{r})$; $\epsilon(\vec{r})$ has the interior or exterior value depending on whether r lies inside or outside the molecular surface.

Equation 2.11 becomes:

$$\vec{\nabla} \left[\epsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \right] = -4\pi [\rho_1(\vec{r}) + \rho_3(\vec{r})] \quad (2.12)$$

Suppose that M is the number of ionic species present in region 3, and $N_i(\vec{r})$ the number of ions per unit volume of each ionic species ($i = 1, 2, \dots, M$) of charge $q_i = Z_i e$; Z_i is a valence of atom i and e is the elementary electric charge. The charge distribution $\rho_3(\vec{r})$ is written as follow:

$$\rho_3(\vec{r}) = \sum_{i=1}^M N_i(\vec{r}) q_i = \sum_{i=1}^M N_i(\vec{r}) e z_i \quad (2.13)$$

According to Debye-Hückel theory for ionic concentrations below 0.2 M at thermodynamic equilibrium, $N_i(\vec{r})$ is given by the Boltzmann distribution as follow:

$$N_i(\vec{r}) = N_i^t \exp(-\beta q_i \phi(\vec{r})) \quad (2.14)$$

with N_i^t the total density of ion i , $\beta = (k_B T)^{-1}$; k_B is the Boltzmann constant and T is the absolute temperature of the medium. Equation 2.13 becomes:

$$\rho_3(\vec{r}) = \sum_{i=1}^M q_i N_i^t \exp(-\beta q_i \phi(\vec{r})) \quad (2.15)$$

The Poisson equation (Equation 2.11) designed to describe the electrostatic proteins in the water, was improved in 1923 by Peter Debye and Erich Hückel. They added a Boltzmann term (Equation 2.15) taking into account the mobility of ions [248]. Thus the general form of PB equation is as following:

$$\vec{\nabla} \left[\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \right] + 4\pi \sum_{i=1}^M q_i N_i^t \exp(-\beta q_i \phi(\vec{r})) = -4\pi \rho_1(\vec{r}) \quad (2.16)$$

Equation 2.16 is usually solved numerically. Although the PB calculation is more accurate than other implicit solvent methods deployed in the application, it is computationally expensive. For this, the Debye-Hückel approximation allow to find an approximate solution of this equation. Here the potential electrostatic energy $q_i \phi(\vec{r})$ is low compared to energy due to thermal agitation ($k_B T = \beta^{-1}$). Knowing that $-1 \leq \beta q_i \phi(\vec{r}) \leq 1$, Taylor's development limited to the first order of $\exp(-\beta q_i \phi(\vec{r}))$ give:

$$\exp(-\beta q_i \phi(\vec{r})) \approx 1 - \beta q_i \phi(\vec{r}) \quad (2.17)$$

With this hypothesis, $\rho_3(\vec{r})$ is rewrite as follow:

$$\rho_3(\vec{r}) = \sum_{i=1}^M q_i N_i^t - \left[\sum_{i=1}^M q_i^2 N_i^t \right] \beta \phi(\vec{r}) \quad (2.18)$$

Taking into account the electroneutrality of the solution:

$$\sum_{i=1}^M q_i N_i^t = 0 \quad (2.19)$$

Thus,

$$\rho_3(\vec{r}) = \left[\sum_{i=1}^M q_i^2 N_i^t \right] \beta \phi(\vec{r}) \quad (2.20)$$

By definition, the ionic strength of the solution I_S is:

$$I_s = \frac{1}{2} \sum_{i=1}^M C_i^t Z_i^2 = \frac{1000}{2N_A} \sum_{i=1}^M N_i^t Z_i^2 \quad (2.21)$$

where $C_i^t = \frac{1000}{N_A} N_i^t$ is the molar concentration of each ionic species and, N_A the number of Avogadro. Thus, we have:

$$\sum_{i=1}^M N_i^t Z_i^2 = \frac{2N_A I_s}{1000} \quad (2.22)$$

And finally, the expression of $\rho_3(\vec{r})$ is the following:

$$\rho_3(\vec{r}) = \frac{-2N_A I_s e^2}{1000 k_B T} \phi(\vec{r}) \quad (2.23)$$

So, the Equation 2.16 becomes:

$$\vec{\nabla} \left[\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \right] - \frac{8\pi N_A I_s e^2}{1000 k_B T} \phi(\vec{r}) = -4\pi \rho_1(\vec{r}) \quad (2.24)$$

Using the modified parameter of Debye-Hückel, $\bar{k}(\vec{r})$ is define as follow:

$\bar{k}(\vec{r})=0$ in region 1 or 2; $\bar{k}(\vec{r})=k\sqrt{\varepsilon}$ in the region 3, k is the Debye-Hückel parameter. $k^2 = \frac{8\pi N_A I_s e^2}{1000 \varepsilon_3 k_B T}$

A linearized PB equation can be written in the following form:

$$\vec{\nabla} \left[\varepsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \right] - k^2 \phi(\vec{r}) = -4\pi \rho_1(\vec{r}) \quad (2.25)$$

This linearized PB (Equation 2.25) is solved by the Finite Difference method based on a grid map.

2.6.0.2 The Finite Difference method (FDM)

The numerical solution usually chosen to solve the PB is to map the ligand or a complex onto a three-dimensional grid and apply the finite difference approximation. Here, the solvation takes place in a cube. The volume of this cube is fonction of the volume of the solute. In the calculation with the Delphi module [187] for example, a filling to 70% of the cube volume is recommended. This cube is subdivided into several elementary cubes of side l (l expressed in Å) as display in Fig.2.8.

A charge $q(\bar{k}_0, \phi_0)$ is placed in the center of an elementary cube. A dielectric constant is defined for each face of this cube. (ε_i $i=1,2,\dots,6$), and charges are placed (q_i, \bar{k}_i, ϕ_i). The electrostatic potential

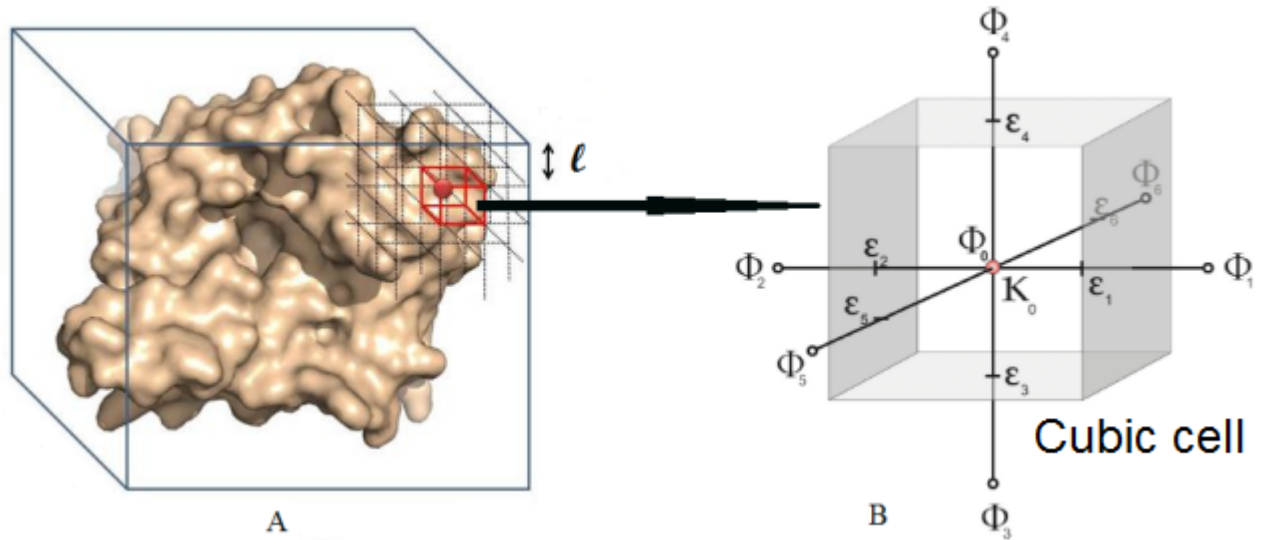


Figure 2.8: Cubic in finite difference method. A: Fragmentation of the medium in identical cubic cells of side l (in red). B: Zoom on one of the cubic cells where the PB equation is discretized.

ϕ_0 at the center of a cell is obtained by integrating the linear equation of PB into the volume V of this cell:

$$\int \int \int_{V_{grid}} \vec{\nabla} [\epsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] dv - \int \int \int_{V_{grid}} k^2 \phi(\vec{r}) dv = -4\pi \int \int \int_{V_{grid}} \rho_1(\vec{r}) \quad (2.26)$$

Knowing that the volume $V = l^3$ and the volume density of charge into the cell $\rho_1(\vec{r}) = q_0 l^{-3}$, Equation 2.26 becomes:

$$\int \int \int_{V_{grid}} \vec{\nabla} [\epsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] dv - k_0^2 l^3 \phi_0 = -4\pi q_0 \quad (2.27)$$

The Green-Ostrogradski theorem allows to write:

$$\int \int \int_{V_{grid}} \vec{\nabla} [\epsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] dv = \oint \epsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \vec{n} ds \quad (2.28)$$

Using the above relation, the equation 2.27 can be expressed as:

$$\oint \epsilon(\vec{r}) \vec{\nabla} \phi(\vec{r}) \vec{n} ds - k_0^2 l^3 \phi_0 = -4\pi q_0 \quad (2.29)$$

Substituting the gradient of the electrostatic potential $\vec{\nabla} \phi(\vec{r})$ by its discrete shape along the three directions, we obtain:

$$\sum_{i=1}^6 \varepsilon_i \frac{(\phi_i - \phi_0)}{l} l^2 - k_0^2 l^3 \phi_0 = -4\pi q_0 \quad (2.30)$$

The expression of the electrostatic potential ϕ_0 in the center of each elementary cube is obtain as follow:

$$\phi_0 = \frac{\sum_{i=1}^6 \varepsilon_i \phi_i + 4\pi \frac{q_0}{l}}{\sum_{i=1}^6 \varepsilon_i + k_0^2 l^2} \quad (2.31)$$

In the current work, the electrostatic component of solvation Gibbs free energy that incorporates also the effects of ionic strength through the solution of nonlinear Poisson–Boltzmann equation [190] was computed by the DelPhi package of Discovery Studio [249]. This DelPhi program treats the solvent as a continuous medium of high dielectric constant ($\varepsilon_0 = 80$) while the solute is treated as a cavity with low dielectric ($\varepsilon_i = 4$). Boundaries are linked to the solute’s molecular surface, which enclose the solute’s atomic charges. The molecular electrostatic potential and reaction field around the solute are solved by a finite difference method on a (235 x 235 x 235) cubic lattice grid for the complexes and free enzyme and (65 x 65 x 65) grid for the free inhibitors, implementing the full Coulombic boundary conditions. In both cases, two (02) subsequent focusing steps led to a similar final resolution of about 0.3 Å per grid unit at 70 % filling of the grid by the solute. Physiological ionic strength of 0.145 mol dm⁻³, atomic partial charges and radii defined in the CFF91 parameter set [190] and a probe sphere radius of 1.4 Å were used. In this implementation, the electrostatic component of the solvation Gibbs free energy was calculated as the reaction field energy [250].

2.7 Calculation of the entropic term

For the drug discovery scientist, the term “thermodynamics” refers to the study of the heat change that occurs when drug targets interact. Thus, entropy variation well knows to be among the physical laws that govern drug-target binding. For example when the hydrophobic region of a drug interacts with a hydrophobic region of a target, water molecules are freed and become less ordered (As display in Fig. 2.5). This leads to an increase in entropy and a gain in binding energy. So the change of entropy is probably a good indicator of the rearrangements undergone by the atoms or solvent (water) molecules when a drug binds to its target. In the gas phase simulation, the vibrational entropy change which occurs as the inhibitor binds to the enzyme was computed by normal mode analysis of the inhibitor vibrations, by using a simplified method previously described by Fischer and co-workers [251, 252]. This approach involves the vibrational analyses of the inhibitor bound at the active site of a ‘frozen’ enzyme, while the low-energy conformers of the free inhibitor were are computed for fully minimised structures. This was done using the Discover package [251] and the formula:

$$T\Delta S_{vib} = T\Delta S_{vib} \{I\}_{FP3} - T\Delta S_{vib} \{I\} \quad (2.32)$$

The method has previously provided a good approximation of the vibrational entropy change of the fully flexible system for small and relatively stiff ligands, i.e. including the degrees of freedom of the protein receptor [29]. In this calculation, the $T\Delta S_{vib} \{I\}$ term can explain vibrational motions of the free inhibitor and the conformational flexibility of the molecule, i.e. low frequency vibrations, which correspond to collective motions of atoms with larger amplitudes (conformational changes contribute mostly to this term). The relative values of $\Delta\Delta T S_{vib}$ with respect to the reference inhibitor were used to compensate partially for the restricted flexibility of the E. In this respect, the entropy term is also recognized as an important factor for drug optimization, even though an enthalpy contribution to binding affinity is known to be more essential [253].

2.8 Binding affinity calculations

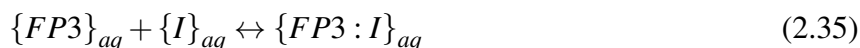
It has been previously proven that the concentration of a competitive tight binding inhibitor that causes 50% reduction of the rate of catalytic substrate conversion (IC_{50}^{exp}) of a reversible inhibitor depends on the enzyme-inhibition constant K_i as follows:

$$IC_{50}^{exp} = K_i + \frac{[S] \cdot K_i}{K_M} + \frac{[E]}{2} \quad (2.33)$$

where [S] and [E] are the substrate and active enzyme concentrations, respectively, while K_M represents the Michaelis constant [254]. The IC_{50}^{exp} value can thus be predicted from the standard Gibbs free energy change during the enzyme: inhibitor complex formation,

$$\Delta G_{comp} = -RT \ln K_i \quad (2.34)$$

assuming the following equilibrium in solution:



Where $\{FP3\}_{aq}$, $\{I\}_{aq}$, and $\{FP3 : I\}_{aq}$ respectively indicate solvated enzyme (E), inhibitor and their complex. The standard Gibbs free energy change of the reaction (1) can be derived by molecular simulations of the complex and the free reactants E and I [255]:

$$\Delta G_{comp} = G \{FP3 : I\} - G \{FP3\} - G \{I\} \quad (2.36)$$

In our the present work, the exact values of standard Gibbs free energies for larger systems such as enzyme: inhibitor complexes have been approximated by the expression [256, 257]:

$$G\{FP3 : I\} \cong [E_{MM}\{FP3 : I\} + RT - TS_{vib}\{FP3 : I\} + G_{sol}\{FP3 : I\}] \quad (2.37)$$

Where $E_{MM}\{FP3 : I\}$ represents the molecular mechanics total energy of the complex (including bonding and non-bonding contributions), $G_{sol}\{FP3 : I\}$ is the solvation Gibbs free energy and $TS_{trv}\{FP3 : I\}$ the entropic term:

$$TS_{trv}\{FP3 : I\} = TS_{trans}\{FP3 : I\} + TS_{rot}\{FP3 : I\} + TS_{vib}\{FP3 : I\} \quad (2.38)$$

i.e., a sum of contributions arising from translational, rotational and vibrational motions. When assuming that the translational and rotational terms for the free enzyme and for the enzyme: inhibitor complex are approximately equal $TS_{trans}\{FP3 : I\} \approx TS_{trans}\{FP3\}$ and $TS_{trv}\{FP3 : I\} \approx TS_{trv}\{FP3\}$, we obtain:

$$\Delta G_{comp} = \Delta H_{MM} + TS_{trans}\{I\} + TS_{rot}\{I\} - T\Delta S_{vib} + \Delta G_{sol} \quad (2.39)$$

with

$$\Delta H_{MM} = [E_{MM}\{FP3 : I\} - E_{MM}\{FP3\} - E_{MM}\{I\}] \quad (2.40)$$

$$T\Delta S_{vib} = [TS_{vib}\{FP3 : I\} - TS_{vib}\{FP3\} - TS_{vib}\{I\}] \quad (2.41)$$

$$\Delta G_{sol} = [G_{sol}\{FP3 : I\} - G_{sol}\{FP3\} - G_{sol}\{I\}] \quad (2.42)$$

$TS_{trans}\{I\}$ and $TS_{rot}\{I\}$ describe the translational and rotational entropy terms of the free inhibitor respectively and ΔTS_{vib} represents the vibrational entropy change upon the complex formation. Comparison between different inhibitors was done via relative changes in the complexation Gibbs free energy with respect to a reference inhibitor, I_{ref} (PEP 1), assuming ideal gas behaviour for the rotational and translational motions of the inhibitors:

$$\Delta\Delta G_{comp} = \Delta G_{comp}(I) - \Delta G_{comp}(I_{ref}) = \Delta\Delta H_{MM} - \Delta\Delta TS_{vib} + \Delta\Delta G_{sol} \quad (2.43)$$

The advantage of such an approach is that the evaluation of relative changes is preferable, since it results in the partial cancellation of errors caused by the approximate nature of the MM method. Additionally, addition solvent and entropic effects are included in the description.

2.9 Quantitative correlation methods

Quantitative structure-activity relationships (QSAR) have been used for decades in the development of relationships between physico-chemical properties of chemical substances and their biological activities in order to obtain a reliable statistical model which is able to predict the activities of new chemical compounds [258]. In terms of drug design, the structure here refers to the properties or descriptors of the molecules, their substituents or interaction energy fields, function corresponds to an experimental biological/biochemical endpoint like binding affinity, activity, toxicity or rate constants. The fundamental principle underlying the formalism is that the difference in structural properties is responsible for the variations in observed biological activities of the molecules. In the classical QSAR studies, for example, affinities of ligands to their binding sites, inhibition constants, and other biological parameters (molecular properties s lipophilicity, polarizability, electronic and steric properties) are often correlated. When some property other than the biological activity is concerned, the term 'quantitative structure-property relationship' (QSPR) is often used. However, the formalism (QSAR) has only a limited utility for designing a new molecule due to the lack of consideration of the 3D structure of the molecules. 3D-QSAR has emerged as a natural extension to the classical Hansch and Free-Wilson approaches, which exploits the three-dimensional properties of the ligands to predict their biological activities. It has served as a valuable predictive tool in drug design and discovery. Although the trial and error factor involved in the development of a new drug cannot be ignored completely, 3D-QSAR certainly decreases the number of compounds to be synthesized by facilitating the selection of the most promising candidates. In general, mostly all the QSAR methods focus on the following goals:

- ▷ To quantitatively correlate and recapitulate the relationships between chemical structure and respective changes in biological activity in order to understand the chemical properties which are responsible for biological activities of the compounds,
- ▷ To optimize the existing leads in order to improve their biological activities, and
- ▷ To predict the biological activities of untested and sometimes yet unavailable compounds.

2.10 Interaction energy calculations

Interaction energy values were computed using Discovery Studio 2.5 [259]. The MM interaction energy (E_{int}) protocol available in this program computes the (non-bonded) van der Waals and electrostatic interactions between enzyme residues and each inhibitor. The CFF force field [190] with a dielectric constant of 4 was used during the calculations. The breakdown of E_{int} into active site residues contributions reveals the significance of individual interactions and permits us to carry out a comparative analysis. The approach leads to the identification of affinity values which would enhance the prediction of favourable and unfavourable PEP substitutions.

2.11 Pharmacophore (PH4) modeling

At the end of the XIXe siècle, Paul Ehrlich has developed the original concept of the pharmacophore (PH4) [259]. At that time, it was known that certain “chemical groups” or functions in a molecule were responsible for a biological effect, and molecules with similar effect had similar functions in common. Schueler has coined much later the word pharmacophore in his 1960 book *Chemobiodynamics and Drug Design*, and was defined as “a molecular framework that carries (phoros) the essential features responsible for a drug’s (pharmacon) biological activity [260]. Thus, the definition of a PH4 was therefore no longer concerned with “chemical groups” but “patterns of abstract features.” Since 1997, a PH4 has been defined by the International Union of Pure and Applied Chemistry as following: A PH4 is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) it’s biological response [261]. The PH4 must be considered as the largest common denominator of the molecular interaction features shared by a set of active molecules. Thus a pharmacophore does not represent a real molecule or a set of chemical groups but is an abstract concept. Often the long definition is simplified as follow “A pharmacophore is the pattern of features of a molecule that is responsible for a biological effect.” This means that a pharmacophore is built from features rather than defined chemical groups.

A pharmacophore is often regarded as a set of features arranged in 3D space which are essential for a molecule to have in order to exert a certain biological activity. The perception of a pharmacophore is often an essential step towards understanding the interaction between a ligand and its receptor. The PH4 concept is based on the assumption that a set of structural features in a molecule is recognised at the receptor site and is responsible for the molecule’s biological activity. Bound conformations of inhibitors taken from E:I complexes were used to construct 3D-QSAR pharmacophore models using Catalyst HypoGen algorithm implemented in DS 2.5 [259]. This consisted in building a top scoring pharmacophore hypothesis from the most active inhibitor. Three stages (construction, subtraction and optimisation) are involved, meanwhile the inactive ones were used to define the exclusion volume. A maximum number of five features were selected according to the PEP scaffold and substituents, i.e. hydrophobic aliphatic (HYd), hydrophobic aromatic (HYdAr), hydrogen-bond acceptor (HBA) hydrogen-bond donor (HBD) and ring aromatic (Ar). As per the adjustable parameters in the HypoGen protocol, all were kept by default excepting for the uncertainty on the activity and the minimum inter-feature distance, which were set to 1.6 and 1.25 Å, respectively. These parameters were carefully chosen in order to bring the uncertainty interval of experimental activity from a wide span $\left[\frac{IC_{50}}{3}, 3 \times IC_{50}\right]$ to a relatively narrow one $\left[\frac{4 \times IC_{50}}{5}, 5 \times \frac{IC_{50}}{4}\right]$. This is important because the accuracy and homogeneity of the measured inhibitory activities based on the fact that the experimental biological activities were derived from the same laboratory must be taken into account [221]. During the generation of 10 pharmacophores, 0 was set as the number of missing features and the best pharmacophore models were selected.

2.12 ADME-related properties

In the past, many clinical candidates failed during development. The reasons for failure are now much better understood. The author of this contribution, Chris Lipinski, was among the first to point out that drugs typically have physicochemical and structural properties within certain ranges. Drug metabolism and pharmacokinetic (DMPK) assessment has come to occupy today an important place during the early stages of drug discovery. Computer-based methods are gradually gaining ground in this area and are often used as an initial tool to eliminate compounds likely to present uninteresting pharmacokinetic profiles and unacceptable levels of toxicity from the list of potential drug candidates, hence cutting down the cost of discovery of a drug. A typical drug discovery workflow, which involves CADD, combined with early stage *in silico* DMPK assessment methods is. Note that designing “drug-like” molecules having interesting pharmacokinetic properties is an important paradigm in drug discovery programs [262]. This involved the search for lead compounds which can be easily orally absorbed, easily transported to their desired receptor, but not easily attacked by metabolising enzymes so as to form toxic metabolic products before reaching the targeted site of action and easily eliminated from the body before accumulating in sufficient amounts that may produce adverse side effects. The ensemble of the above properties is often referred to as ADME (absorption, distribution, metabolism and elimination) properties, or better still ADMET, ADME/T or ADMETox (i.e., if toxicity criteria are also taken into consideration). Computer-based *in silico* methods for the prediction of ADMET profiles of drug leads at early stages of drug discovery are increasingly gaining ground [263-265]. This could be explained by the relative cost advantage added to the time factor when compared to standard experimental approaches for ADMET profiling [264, 265]. As an example, it takes a minute in an *in silico* model to screen 20,000 molecules but takes 20 weeks in the “wet” laboratory to do the same exercise [262]. Due to the accumulated ADMET data in the late 1990s, many pharmaceutical companies are now using computational models that, in some cases, are replacing the “wet” screens [262]. This paradigm shift has, therefore, stimulated the development of several theoretical methods for the prediction of ADMET parameters. Several theoretical methods for the determination of ADMET parameters have been developed and implemented in a number of currently available softwares for drug discovery protocols [266-268], even though predictions are sometimes disappointing [268]. Such software often make use of quantitative structure-activity relationships (QSAR) [267, 268] or knowledge-base methods [269]. The goal has been to considerably cut down on the currently very high cost of discovery of a drug [270]. A promising lead is often defined as a compound which combines potency with an admirable ADMET profile. So, compounds with unfavourably predicted pharmacokinetic profiles are either completely dismissed from the list of potential drug candidates (even if they prove to be highly potent) or the DMPK properties are “fine-tuned” in order to improve their chances of making it to clinical trials [270]. This explains why the “graveyard” of very highly potent compounds which do not make it to clinical trials keeps filling up, to the extent that the process of drug discovery often presents the challenge of either resorting to new leads or “resurrecting” some buried leads with the view of “fine-tuning” their ADMET profiles.

Properties that determine the pharmacokinetics profile of a compound, besides octanol/water par-

titiation coefficient, aqueous solubility, blood/brain partition coefficient, Caco-2 cell permeability, serum protein binding, number of likely metabolic reactions and eighteen more descriptors related to adsorption, distribution, metabolism and excretion (ADME properties) of the inhibitors were computed by the QikProp program [271] based on the methods of Jorgensen [272, 273]. According to those methods, experimental results of more than 710 compounds among which about 500 drugs and related heterocycles were correlated with computed physico-chemical descriptors resulting in an accurate prediction of molecule's pharmacokinetic profile. Drug likeness (#stars) the number of property descriptors that fall outside the range of values determined for 95 % of known drugs out of 24 selected descriptors computed by the QikProp [271], was used as an additional ADME-related compound selection criterion.

2.13 Generation of the virtual library

Molecular models of new analogues were generated using the Molecular Operating Environment (MOE) program [274]. This was carried out by attaching the R-groups (fragments, building blocks) onto the PEP scaffolds using the Quasar CombiDesign module of MOE program. Chemical reagents considered in this study were taken from the directories of chemicals available from the commercial suppliers of chemicals [275]. Each analogue was built as a neutral molecule in MOE program [274] and its molecular geometry was refined by MM optimization, implemented in the Discovery Studio 2.5 [259] smart minimizer. Convergence criteria (energy difference of 10^{-4} kcal mol⁻¹, R.M. S. displacement of 10^{-5} Å) and a dielectric constant of 4 using class II consistent force field CFF [259] as described in Molecular Mechanics section.

2.14 ADME-based library focusing

The library focusing strategy based on ADME-related properties and the pharmacophore-based library focusing procedure was previously described [231]. Twenty four pharmacokinetic molecular descriptors available in QikProp [271], which characterize a wide spectrum of molecular properties as described in section 2.12. such as molecular mass, total solvent-accessible molecular surface, hydrophobic portion of the solvent-accessible molecular surface, total volume of molecule enclosed by solvent-accessible molecular surface, number of non-trivial non-hindered rotatable bonds, estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution, estimated number of hydrogen bonds that would be accepted by the solute from water molecules, logarithm of partitioning coefficient between n-octanol and water phases, logarithm of predicted aqueous solubility, logarithm of predicted binding constant to human serum albumin, logarithm of predicted brain/blood partition coefficient, apparent Caco-2 cell membrane permeability in Boehringer-Ingelheim scale, number of likely metabolic reactions, percentage of human oral absorption in gastrointestinal tract, etc. Optimum ranges of the 24 descriptors were defined in terms of

upper and lower bounds, and average values according to QikProp [272]. Drug likeness was used as ADME-related compound selection criterion.

2.15 Pharmacophore-based library focusing

The pharmacophore models (PH4) previously describe was derived from the bound conformations of PEPs at the active site of FP3. The enumerated and ADME-focused virtual library was further focused by using the ligand pharmacophore mapping protocol available of Discovery Studio [259]. Within this protocol, each generated conformer of the analogues was geometry optimized by means of the CFF forcefield for a maximum of 500 energy minimization steps and subsequently aligned and mapped to the PH4 model in order to select the top ranking overlaps. 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 [276]. The best representative of each cluster was considered in the virtual screening of analogues.

2.16 *In silico* screening

The conformer with the best mapping on PH4 pharmacophore in each cluster was selected for virtual screening using the complexation QSAR model. For each E:I complex, the relative complexation Gibbs free energy $\Delta\Delta G_{comp}$, was calculated. This was then used to compute the predicted activities pIC_{50}^{pre} of each of the newly designed analogues against FP-3. The predicted IC_{50}^{pre} values were then calculated using the formular:

$$IC_{50}^{pre} = 10^{-pIC_{50}^{pre}(FP3)} \quad (2.44)$$

RESULTS AND DISCUSSION

Our aim was to design low nanomolar antimalarial agents, peptidomimetics (PEP) inhibitors, which inhibit cysteine protease falcipain-3 protein of *Plasmodium falciparum* (*Pf*). Three-dimensional (3D) models of FP3-PEP complexes for training sets of published PEP analogues with known *in vitro* biological activities (IC_{50}^{exp}) have been prepared by *in situ* modification of the crystal structure of *Pf*FP3 inhibited by K11017 (Protein Data Bank entry 3BWK [30]). Molecular mechanics, conformational searching and implicit solvation model have been used to compute Gibbs free energies of inhibitor-FP3 receptor complex formation. A Quantitative structure-activity relationships (QSAR) model has been built leading to a linear correlation between the calculated free energies of complexation ($\Delta\Delta G_{comp}$) and *in vitro* biological activities: $pIC_{50}^{exp} = \log_{10} IC_{50}^{exp} = -0.8037 \times \Delta\Delta G_{comp} + 4.0009$; $R^2 = 0.95$). The obtained complexation model could explain about 95% of the FP-3 inhibition. In addition, ligand-based quantitative pharmacophore model (PH4) was built from bound conformations of the training set compounds and confirmed the correlation between molecular models and observed *in vitro* biological activities $pIC_{50}^{exp} = -\log_{10} IC_{50}^{exp} = 0.9958 \times pIC_{50}^{pre} + 0.0219$; $R^2 = 0.95$). Structural information obtained from the both model guided us in designing virtual combinatorial libraries of novel PEP analogues. Comparative analysis of the active site interactions directed us in the selection of building blocks used in the libraries. The initial virtual library was focused by means of computationally predicted oral bioavailability (ADME profile) and subsequently *in silico* screened with the PH4 pharmacophore models to identify new PEP inhibitor candidates. Their inhibitory activities predicted by the QSAR model fall into the low nanomolar concentration range. The best inhibitor reached a predicted of 0.003M.

3.1 Selection of training and validation (or test) data sets

A data set of ten (10) FP-3 inhibitors with a broad range of *in vitro* activities, obtained from (IC_{50}^{exp}) the same laboratory, with a sufficiently broad range of activities (60 – 47230 nM) [221] were used to generate a 3D-QSAR model. This was divided into a training set of eight (08) analogues (PEP_{1–8}) used to build the QSAR model and a validation (or test) set of two (02) analogues for evaluating the model. Table 3.1 give a summary of the *in vitro* activities (in IC_{50}^{exp} values against FP-3) of the

selected PEP analogues used in this study. Their chemical structures are provided in the Materials and Methods chapter. The IC_{50}^{exp} is the inhibitory concentration that causes 50% decrease in the rate of substrate conversion by the enzyme, measured in the enzyme inhibition assay [221]:

$$IC_{50}^{exp} = K_i + \frac{[S] \cdot K_i}{K_M} + \frac{[E]}{2} \quad (3.1)$$

where [S] and [E] are the substrate and active enzyme concentrations and K_M the Michaelis constant [254].

Comparison of IC_{50}^{exp} s let observe that PEP2 and related compounds (PEP3 and PEP4) may serve as a promising direction forward to the design of potent inhibitors of FP-3 to be used as leads for the effective treatment of resistant forms of malaria.

Table 3.1: *In vitro Biological Activities of data set of PEP inhibitors obtained from the literature.*

| Inhibitor ^a | Mw ^b (g. mol ⁻¹) | $IC_{50}^{exp,c,d}$ (nM) |
|------------------------|---|--------------------------|
| PEP 1 (ref) | 482 | 36360 |
| PEP 2 | 440 | 60 |
| PEP 3 | 574 | 520 |
| PEP 4 | 498 | 3560 |
| PEP 5 | 450 | 20180 |
| PEP 6 | 438 | 23900 |
| PEP 7 | 466 | 47230 |
| PEP 8 | 488 | 11910 |
| PEP 9 | 452 | 540 |
| PEP 10 | 470 | 8220 |

^aSee Fig 3.2. chapter 3 for chemical structures of inhibitors, ^bMolar weight of PEP analogues, ^cThis is IC_{50}^{exp} expressed in nanomolar concentration, ^dFor potencies of PEP analogues against FP-3, Ref. [221] is referred.

3.2 Data set and complexation model

3.2.1 Energies of the protein-ligand complexes

The lowest energy structures for the protein-ligand complexes obtained using the procedure earlier described in chapter 3 by torsion drives were retained. The same physical quantities were calculated as previously. These are presented in Table 3.2.

Table 3.2: Results of Interaction Energies, Solvation Enthalpies and Entropic Terms of the Ligands in Complex with FP-3.

| Inhibitor | $\Delta H_{MM}\{FP3 : I\}$ (kcal.mol ⁻¹) | $G_{sol}\{FP3 : I\} \times kT$ | $G_{sol}\{FP3 : I\}$ (kcal.mol ⁻¹) | $TS_{vib}\{FP3 : I\}$ (kcal.mol ⁻¹) | $TS_{vib}\{FP3 : I\}$ (kcal.mol ⁻¹) |
|------------------|--|--------------------------------|--|---|---|
| PEP 1 | -3 481.59 | -1 011.52 | -596.79 | 48.47 | 48.47 |
| PEP 2 | -3 472.47 | -1 013.39 | -597.90 | 43.09 | 43.09 |
| PEP 3 | -3 493.17 | -1 014.81 | -598.73 | 58.25 | 58.25 |
| PEP 4 | -3 485.21 | -1 013.30 | -597.85 | 49.36 | 49.36 |
| PEP 5 | -3 500.58 | -1 014.42 | -598.51 | 45.27 | 45.27 |
| PEP 6 | -3 497.75 | -1 017.70 | -600.44 | 46.02 | 46.02 |
| PEP 7 | -3 478.92 | -1 014.40 | -598.50 | 47.83 | 47.83 |
| PEP 8 | -3 482.62 | -1 008.20 | -594.84 | 47.78 | 47.78 |
| PEP 9 | -3 509.15 | -1 013.73 | -598.10 | 45.87 | 45.87 |
| PEP 10 | -3 489.91 | -1 013.55 | -598.00 | 45.27 | 45.27 |

3.2.2 Energy of free ligands

In the complexation model proposed within this study a conformational analysis of the free ligands was carried out. The enthalpy (ΔH_{MM}), the entropy factor (TS_{vib}) and the solvation energy (G_{sol}) were each calculated for the series of 10 ligands studied [221]. The search for stable conformations was carried out using a Molecular dynamic and Monte Carlo search at 5000 K (see detailed description in chapter 2). This generated 200 different conformations, corresponding to different energies. Figure 3.1, for example, displays 200 conformations obtained by MD simulation of the most active ligand.

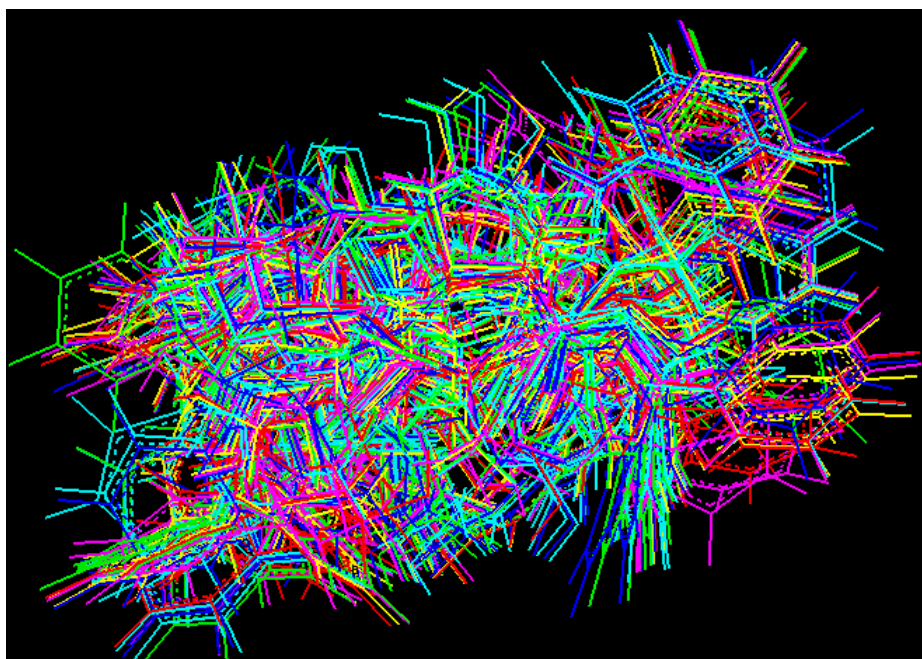


Figure 3.1: 200 snapshots obtained during the molecular dynamics simulation of the most active ligand.

The lowest energy (most stable) conformations were selected and further minimized by molecular mechanics method using the suitable forcefield. The results from energy calculations of the free ligands were then used in the complexation model for FP-3. These values are presented in Table 3.3

Table 3.3: Results of Interaction Energies, Solvation Enthalpies and Entropic Terms of the Free Ligands.

| Inhibitor | $\Delta H_{MM}\{FP3:I\}$ (kcal.mol ⁻¹) | $G_{sol}\{FP3:I\} \times kT$ | $G_{sol}\{FP3:I\}$ (kcal.mol ⁻¹) | $TS_{vib}\{FP3:I\}$ (kcal.mol ⁻¹) |
|------------------|--|------------------------------|--|---|
| PEP 1 | 3.00 | -7.00 | -4.13 | 40.66 |
| PEP 2 | 14.04 | -8.76 | -5.17 | 37.06 |
| PEP 3 | -7.43 | -7.68 | -4.53 | 50.51 |
| PEP 4 | 0.63 | -8.69 | -5.13 | 42.30 |
| PEP 5 | -14.98 | -9.54 | -5.63 | 37.34 |
| PEP 6 | -14.11 | -8.52 | -5.03 | 37.40 |
| PEP 7 | 7.12 | -9.02 | -5.32 | 38.69 |
| PEP 8 | 2.52 | -9.20 | -5.43 | 43.69 |
| PEP 9 | -22.47 | -8.81 | -5.20 | 38.24 |
| PEP 10 | -5.04 | -10.08 | -5.95 | 39.06 |

3.2.3 Complexation energies of the protein-ligand complexes

The relative Gibbs free energy change related to the enzyme-inhibitor complex formation was determined using the following relation (equation 3.2):

$$\Delta\Delta G_{comp} = \Delta\Delta H_{MM} + \Delta\Delta G_{sol} - \Delta\Delta TS_{vib} \quad (3.2)$$

where $\Delta\Delta H_{MM}$ ($\Delta\Delta E_{MM}$ or) is the relative enthalpic contribution to the Gibbs free energy change related to the intermolecular interactions in the enzyme-inhibitor complex derived by molecular mechanics, i.e., the difference between the interaction energy (E_{MM}) of each complex within the series and that of the reference complex $\{FP3 : I_{ref}\}$, expressed in equation 3.3. I_{ref} is the reference inhibitor PEP 1.

$$\Delta\Delta H_{MM} = [H_{MM} \{FP3 : I_{ref}\} - H_{MM} \{I_x\}] - [H_{MM} \{FP3 : I_{ref}\} - H_{MM} \{I_{ref}\}] \quad (3.3)$$

$\Delta\Delta G_{sol}$ is the relative solvation Gibbs free energy contribution to the Gibbs free energy of the enzyme-inhibitor complex formation, i.e., the difference between the solvation energy (G_{sol}) of each complex within the series and that of the reference complex (equation 3.4):

$$\Delta\Delta G_{sol} = [G_{sol} \{FP3 : I_{ref}\} - G_{sol} \{I_x\}] - [G_{sol} \{FP3 : I_{ref}\} - G_{sol} \{I_{ref}\}] \quad (3.4)$$

and $-\Delta\Delta TS_{vib}$ is the relative entropic contribution of the inhibitor to the Gibbs free energy related to the enzyme-inhibitor complex, i.e., the difference between the entropy term (ΔTS_{vib}) of each ligand within the series and that of the reference ligand (equation 3.5):

$$-\Delta\Delta TS_{vib} = [TS_{vib} \{FP3 : I_{ref}\} - TS_{vib} \{I_x\}] - [TS_{vib} \{FP3 : I_{ref}\} - TS_{vib} \{I_{ref}\}] \quad (3.5)$$

The results presented in Table 3.4:

Table 3.4: Results of Binding Free Energies of Complexation of the Ligands in Complex with FP-3.

| Inhibitor | ΔH_{MM} (kcal.mol ⁻¹) | $\Delta\Delta G_{sol}$ (kcal.mol ⁻¹) | $\Delta\Delta T S_{vib}$ (kcal.mol ⁻¹) | $\Delta\Delta G_{comp}$ (kcal.mol ⁻¹) | pIC_{50}^{exp} |
|-----------|---|--|--|---|------------------|
| PEP 1 | 0 | 0 | 0 | 0 | 4.44 |
| PEP 2 | -1.92 | -0.08 | 1.78 | -3.78 | 7.22 |
| PEP 3 | -1.16 | -1.54 | 0.07 | -2.77 | 6.28 |
| PEP 4 | -1.26 | -0.07 | 0.74 | -2.07 | 5.45 |
| PEP 5 | -1.01 | -0.22 | -0.13 | -1.10 | 4.70 |
| PEP 6 | 0.94 | -2.75 | -0.81 | -1.00 | 4.62 |
| PEP 7 | -1.45 | -0.51 | -1.33 | -0.63 | 4.33 |
| PEP 8 | -0.56 | 3.25 | 3.72 | -1.03 | 4.92 |
| PEP 9 | -2.10 | -0.23 | 0.18 | -2.51 | 6.27 |
| PEP 10 | -0.28 | 0.61 | 1.61 | -1.28 | 5.09 |

3.2.4 Selection of training and validation (or test) data sets

A data set of FP-3 inhibitors (ten inhibitors 10) with measured *in vitro* biological activities (IC_{50}^{exp} and $pIC_{50}^{exp} = -\log_{10}IC_{50}^{exp}$) values obtained from the same laboratory [221] was used to generate a 3D-QSAR model. The dataset was divided into a training set of eight (08) molecules for generating the subsequent QSAR models and a test set of two (02) molecules for evaluating the predictive quality of the models. Table 3.1 depicts IC_{50}^{exp} values of the data set while Fig 3.2 depicts several representative skeletons of the data set. By comparison of IC_{50}^{exp} of PEP derivatives against FP-3, it is observed that PEP2 and related compounds (PEP3 and PEP4) may serve as a promising direction forward in the design of potent inhibitors of FP-3 that will be used as drugs for the effective treatment of resistant forms of malaria.

3.3 Correlation curve including 8 ligands of training set

A plot of pIC_{50}^{exp} against $\Delta\Delta G_{comp}$ (Figure 3.2) reveals that 94.71% of the activities of all the 8 PEP derivatives against FP-3 could be explained by the complexation model (a squared correlation coefficient, $R^2 = 0.95$). The most active inhibitor (PEP 2, $pIC_{50}^{exp} = 7.22$) and the least active inhibitor (PEP 1, $pIC_{50}^{exp} = 4.33$) against FP-3 at the two extremes of the best fit straight line. The following statistically significant correlation equation was obtained (equation 3.6):

$$pIC_{50}^{exp} = \log_{10}IC_{50}^{exp} = -0.8037 \times \Delta\Delta G_{comp} + 4.0009 \quad (3.6)$$

3.4 Validation of QSAR model

The reason behind any QSAR study is to obtain a model with the highest predictive and generalization abilities. The QSAR model obtain was validated both internally and externally. Internal validation was used to verify the robustness of the model, while external validation was used to determine its predictivity.

3.4.1 Internal validation by leave-one-out (LOO) method

One compound in the training set was left out at a time and its biological activity was then calculated (or predicted) using the QSAR model derived from the biological activity of the 7 other molecules. The theoretically predicted activities (pIC_{50}^{pre}) of each of the 8 molecules obtained by the LOO method were then compared with the experimental activities (Table 3.5). A correlation plot of the experimentally measured activities, pIC_{50}^{exp} (on the y-axis) against pIC_{50}^{pre} (on the x-axis), was then obtained

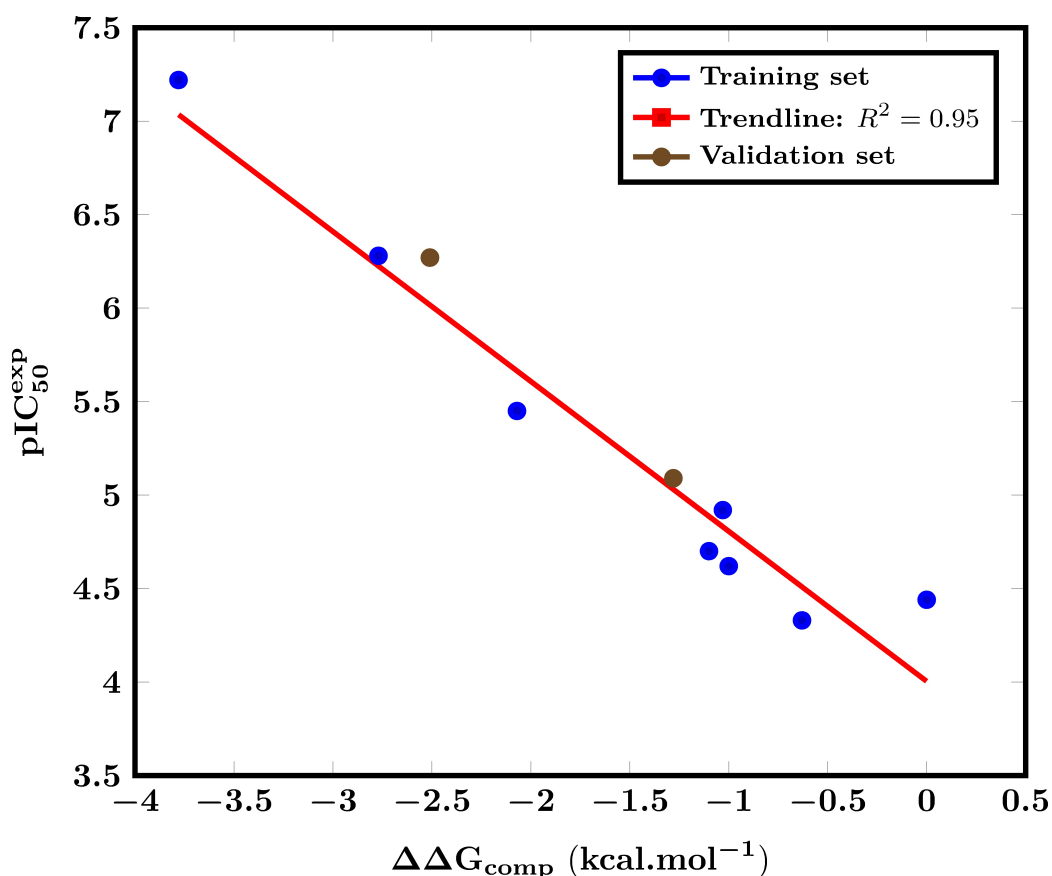


Figure 3.2: A plot of the correlation equation between pIC_{50}^{exp} and relative complexation Gibbs free energies of the 8 training set of FP-3 inhibitors $\Delta\Delta G_{comp}$

(Fig 3.3). The cross validated correlation coefficient R_{xv}^2 value of 0.90, which indicates that a major portion (about 90 %) of the variance of the training set data was well described by this QSAR model.

The SDEP (standard deviation of error prediction, s) was also calculated from the experimentally observed and theoretically predicted potencies of each compound.

where n is the number of compounds in the dataset ($n = 8$). This gave us the following linear relationship between the experimentally observed and theoretically predicted potencies as (shown in Figure 3.3):

$$pIC_{50}^{exp} = -\log_{10} IC_{50}^{exp} = -0.8037 \times \Delta\Delta G_{comp} + 4.0009, n = 8, R_{xv}^2 = 0.90, s = 0.24$$

The cross validation correlation coefficient obtained (0.90) is high. This reflected the high positive gradient of this curve. Moreover low SDEP (0.24) indicate that there is a true correlation between the experimental and predicted potencies.

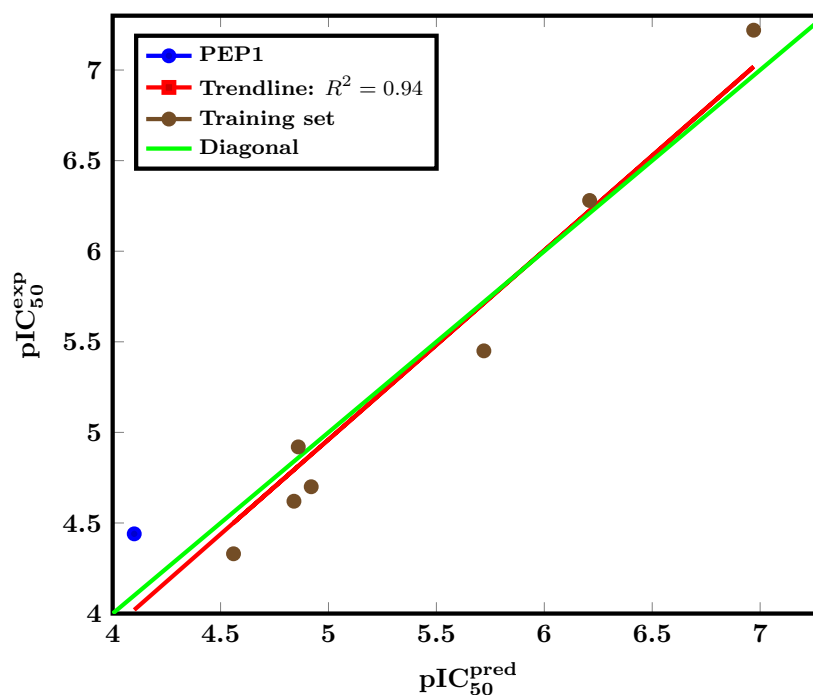


Figure 3.3: Internal validation of QSAR model and calculation of cross validation coefficient (). Outlier (PEP 1) is highlighted in blue.

3.4.1.1 Comparison between theoretical and experimental activities

The ratio of $\frac{pIC_{50}^{pred}}{pIC_{50}^{exp}}$ is close to 1 for all the molecules in the training set (Table 3.5). Concerning the ratio of $\frac{IC_{50}^{pred}}{IC_{50}^{exp}}$, only the values corresponding to PEP 1 and PEP 4 are quite far from 1. The rest of the values are close to 1 (practically along the regression line). However, we can note that PEP 1 is the almost least active compound within this dataset, which could explain the abnormality. However the residual values ($pIC_{50}^{pred} - pIC_{50}^{exp}$) of -0.34 corresponding to PEP 1 is < 0.40 in absolute value, which is a good indicator that this model is acceptable.

Table 3.5: Comparison of Experimentally Measured and Theoretically Predicted Activities of the Training set.

| Inhibitor (I) | pIC_{50}^{exp} | pIC_{50}^{pred} | IC_{50}^{exp} (μM) | IC_{50}^{pred} (μM) | Ration of pIC_{50}^{pred} $\left(\frac{pIC_{50}^{pred}}{pIC_{50}^{exp}}\right)$ | Residual of pIC_{50}^{pred} $\left(pIC_{50}^{pred} - pIC_{50}^{exp}\right)$ | Ration of IC_{50}^{pred} $\left(\frac{IC_{50}^{pred}}{IC_{50}^{exp}}\right)$ |
|---------------|------------------|-------------------|--------------------------------|---------------------------------|--|--|---|
| PEP 1 | 4.44 | 4.10 | 36.36 | 79.43 | 0.92 | -0.34 | *2.19 |
| PEP 2 | 7.22 | 6.97 | 0.06 | 0.11 | 0.97 | -0.25 | 1.78 |
| PEP 3 | 6.28 | 6.21 | 0.52 | 0.61 | 0.99 | -0.07 | 1.17 |
| PEP 4 | 5.45 | 5.72 | 3.56 | 1.89 | 1.05 | 0.27 | *0.53 |
| PEP 5 | 4.70 | 4.92 | 20.18 | 12.13 | 1.05 | 0.22 | 0.60 |
| PEP 6 | 4.62 | 4.84 | 23.90 | 14.45 | 1.05 | 0.22 | 0.61 |
| PEP 7 | 4.33 | 4.56 | 47.23 | 27.76 | 1.05 | 0.23 | 0.60 |
| PEP 8 | 4.92 | 4.86 | 11.91 | 13.81 | 0.99 | -0.06 | 1.16 |

*With the exception of these values, the rest of the points are close to the regression line.

Noted that the correlation coefficient R is a measure of quality of fit of the model. It constitutes the variance in the data. The ideal situation correspond the correlation coefficient is equal to or approach 1. But in reality because of the complexity of biological data, any value above 0.90 (corresponding to $R^2 = 0.81$) is adequate. The standard deviation is an absolute measure of the quality of fit. In an ideal situation, s should approach zero, but in experimental situations, this is not so. It should be small but it cannot have a value lower than the standard deviation of the experimental data. The magnitude of s may be attributed to some experimental error in the data as well as imperfections in the biological model. Generally, a larger dataset and a smaller number of variables lead to lower values of s . The F value is usually used as a measure of the level of statistical significance of the regression model. Thereby, the above model was further validated by calculation of the F -statistic. This was calculated from the value of R^2 , the number of data points (or degrees of freedom) in the training set and the number of variables used to build the QSAR model.

F -statistic of 120.7 for a level of the statistical significance $\alpha > 95\%$, signifying that the probability of having a chance correlation¹ is < 0.128 . When one has a larger value of F , this implies that a more significant correlation has been reached. Table 3.6 gives a summary of the statistical data of the model.

Table 3.6: Statistical information of correlation between computed $\Delta\Delta G_{comp}$ of training set inhibitors and their experimental activities IC_{50}^{exp} against FP-3.

| | |
|---|----------|
| $pIC_{50}^{exp} = \log_{10}IC_{50}^{exp} = -0.8037 \times \Delta\Delta G_{comp} + 4.0009$ | |
| Statistical data of linear regression: | |
| Number of compounds n | 8 |
| Squared correlation coefficient of regression R^2 | 0.95 |
| Leave-one-out cross-validated squared correlation Coefficient R_{xv}^2 | 0.88 |
| Standard error of the regression σ | 0.24 |
| Statistical significance of regression, Fisher F-test | 120.7 |
| Level of statistical significance α | >95% |
| Range of activities of IC_{50}^{exp} (nM) | 60–47230 |

3.4.1.2 External validation by using a test set

In addition to internal validation, the quality of the complexation model was confirmed by predicting the FP-3 inhibitory activities for a validation (or test) set of 2 randomly selected PEP analogues PEP9 and PEP10 (not included into the training set) with known IC_{50}^{exp} values (Table 3.7). The ratio of the predicted activities derived from equation 3.6 obtained from the QSAR model of a training set of 8 analogues and observed activities (both $\frac{pIC_{50}^{pred}}{pIC_{50}^{exp}}$ and $\frac{IC_{50}^{pred}}{IC_{50}^{exp}}$, see Table 3.5), which yielded values close to 1, confirmed the predictive power of the QSAR model. The training and validation sets used [221] display some change at the position P1' and P2. However, the validated model could be used to

¹This was verified from the Student-Fischer statistics distribution table

predict the activities of PEP analogues against FP-3, provided that these ones bind to the FP-3 active site in the same way as PEP2.

Table 3.7: Training and Validation Sets of PEP derivatives used in the QSAR Model of FP-3 Inhibition and in the Validated Complexation Model[221].

| <i>Training Set</i> ^a | pIC_{50}^{exp} | IC_{50}^{exp} [μM] |
|------------------------------------|------------------|-----------------------------|
| PEP 1 | 4.44 | 36.36 |
| PEP 2 | 7.22 | 0.06 |
| PEP 3 | 6.28 | 0.52 |
| PEP 4 | 5.45 | 3.56 |
| PEP 5 | 4.70 | 20.18 |
| PEP 6 | 4.62 | 23.90 |
| PEP 7 | 4.33 | 47.23 |
| PEP 8 | 4.92 | 11.91 |
| <i>Validation Set</i> ^a | pIC_{50}^{exp} | IC_{50}^{exp} |
| PEP 9 | 6.27 | 0.54 |
| PEP 10 | 5.09 | 8.22 |

^aSee Fig 3.2. chapter 2 for chemical structures of training and validation, pIC_{50}^{pred} and experimental IC_{50}^{exp} . For experimental data (IC_{50}^{exp}), see Ref [221] and Table 3.1.

3.5 Binding mode of inhibitors of training set

The binding mode and the Connolly surface of the best active PEP2 coming from the complexation model are illustrated in 3D and 2D depiction in Figure 3.4 and 3.6. The main interactions with the active site residues namely the H-Bond with the catalytic residue Cys51 are in line with docking study and WaterMap calculations [221] which, unfortunately, did not provide any statistical correlation between binding affinity and activity (results not shown). The bound conformation of PEP shed light on the structural features for binding affinity, which are vital for the design of novel potent non-peptidic FP-3 inhibitors by exploiting the S1' to S2 pockets.

In Figure 3.6, the binding site surface is colored according to residue hydrophobicity: red -hydrophobic, blue -hydrophilic and white – intermediate.

3.6 Interaction energy calculations

In order to check if other interesting interactions not displayed have to be taken account in the description of PEP binding mode at FP-3 active site for the rational design of new analogues, interaction energy (IE) between some active site residue of FP-3 (S2) and PEPx (P2) was computed.

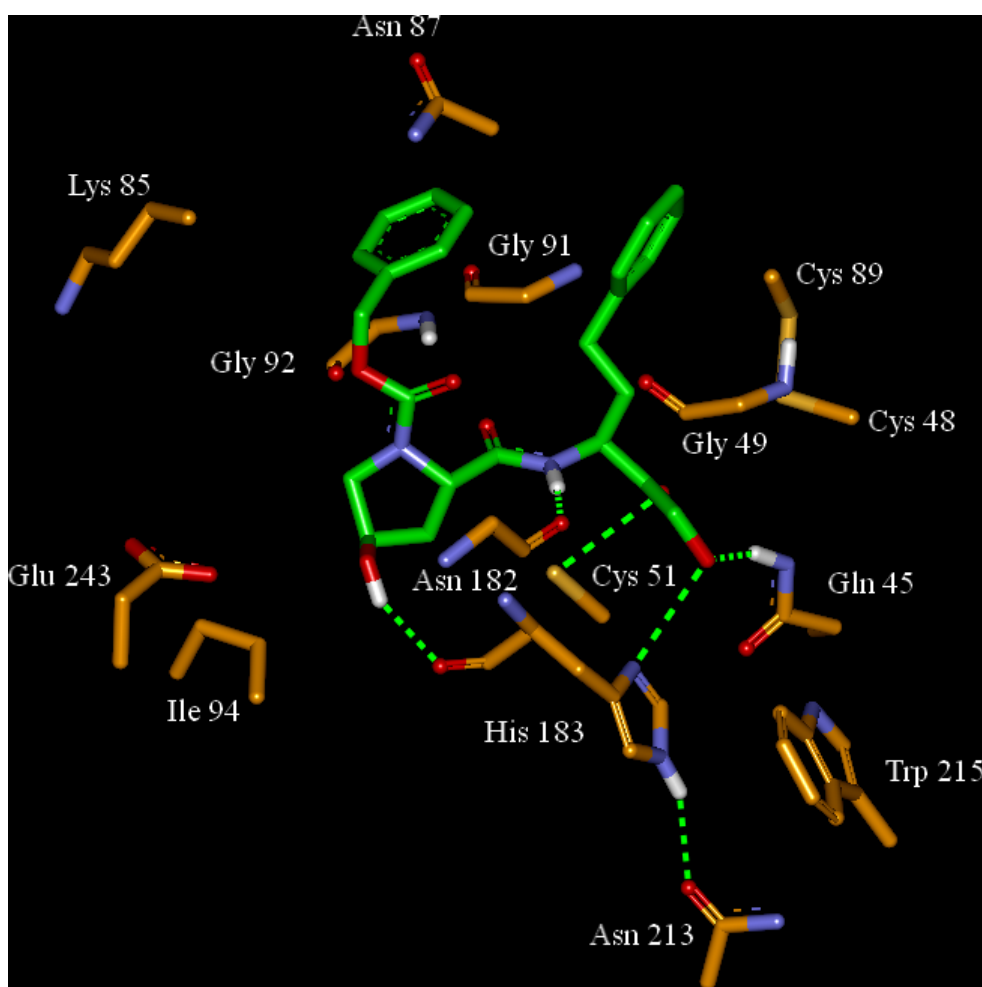


Figure 3.4: Enzyme-ligand interactions at the active site of FP-3 depicted in 3D for the most active PEP2.

Table 3.8: Interaction energy between some active site residue of FP-3 (S2) and PEPx (P2).

| inhibitor | ser 158 | pro 181 | asn 182 | his 183 | ala 184 | glu- 245 | total ie (kcal.mol ⁻¹) | pIC_{50}^{exp} |
|-----------|---------|---------|---------|---------|---------|----------|------------------------------------|------------------|
| pep 1 | -1.00 | -1.01 | -1.50 | -1.56 | -0.87 | -0.99 | -6.94 | 4.44 |
| pep 2 | -0.94 | -1.18 | -1.75 | -2.97 | -1.34 | -1.31 | -9.50 | 7.22 |
| pep 3 | -0.99 | -1.06 | -2.06 | -2.76 | -1.45 | -1.00 | -9.32 | 6.28 |
| pep 4 | -0.95 | -1.60 | -1.77 | -2.13 | -1.37 | -1.02 | -8.83 | 5.45 |
| pep 5 | -1.68 | -1.24 | -0.88 | -0.13 | -0.56 | -0.42 | -4.92 | 4.70 |
| pep 6 | -1.19 | -1.18 | -1.60 | -1.37 | -1.32 | -0.81 | -7.48 | 4.62 |
| pep 7 | -1.20 | -1.23 | -1.68 | -1.10 | -1.24 | -0.96 | -7.41 | 4.33 |
| pep 8 | -1.33 | -2.03 | -1.08 | -1.52 | -1.01 | -1.02 | -7.99 | 4.92 |

The breakdown of IE diagram into some active sub-site residues S2 contribution of FP-3 for PEPs, displayed in Fig 3.7, indicate the particular behavior of P2. Also, Table 3.8 show that the most active inhibitor PEP2 interacts more with residues from sub-site S2 (IE= -9.50 kcal.mol⁻¹) than other inhibitors. The particular contribution of His 183 to IE for the most active compound PEP2 is observe. This contribution of His 183 to IE at the P2 position of PEPs listen in Table3.9 can be correlated with the observed activity of the training set inhibitors as shown in Fig 3.8.

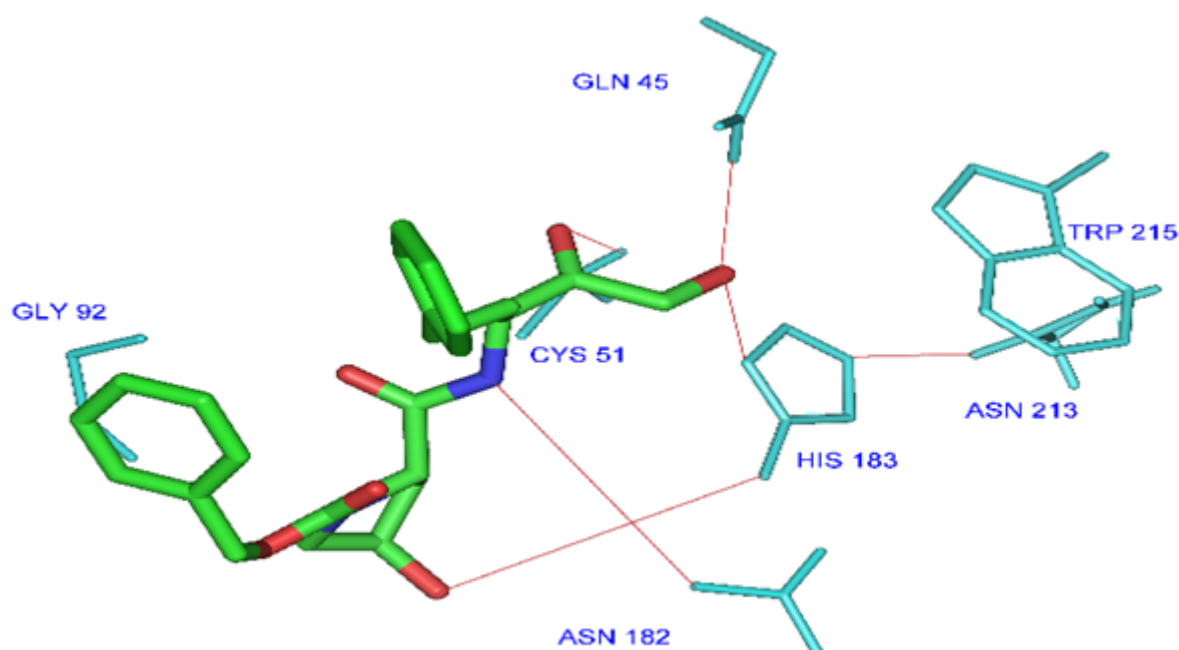


Figure 3.5: 2D schematic interaction diagram of the most potent inhibitor PEP2 at the active site of PfFP3.

Table 3.9: Interaction energy between P1 chains and HIS183 residues.

| Inhibitor | HIS 183 (E_{int}) | ΔE_{int} | pIC_{50}^{exp} |
|--------------|-----------------------|------------------|------------------|
| PEP 1 | -1.56 | 0.00 | 4.44 |
| PEP 2 | -2.97 | 1.43 | 7.22 |
| PEP 3 | -2.76 | 0.19 | 6.28 |
| PEP 4 | -2.13 | 0.46 | 5.45 |
| PEP 5 | -0.13 | 0.04 | 4.70 |
| PEP 6 | -1.37 | -1.41 | 4.62 |
| PEP 7 | -1.10 | -1.20 | 4.33 |
| PEP 8 | -1.52 | -0.57 | 4.92 |

Moreover, the breakdown of IE diagram into each S1' pocket residues contribution of FP-3 for PEPs, as displayed in Fig 3.9 shows that TPR 215 is mostly involved in the contribution of IE of the two most active compounds PEP2 and PEP3, and that ASN 213 residue is most involved in IE with PEP 2 the most active compound than the others residues in S1' pocket. We have focused on the S1' and S2 subsites because substitutions were made in P1' and P2 during the original SAR performed by Weldon et al [221] (see Fig 3.2).

The above results suppose that it is important to design molecules capable to interact with His 183 at S2 and interact with TPR215 in S1'.

It was observed that the IE diagrams analysis could not significantly guide the choice of the R-groups in S1' and S2 subsites, when compared with the case for the design of thymine-like inhibitors of

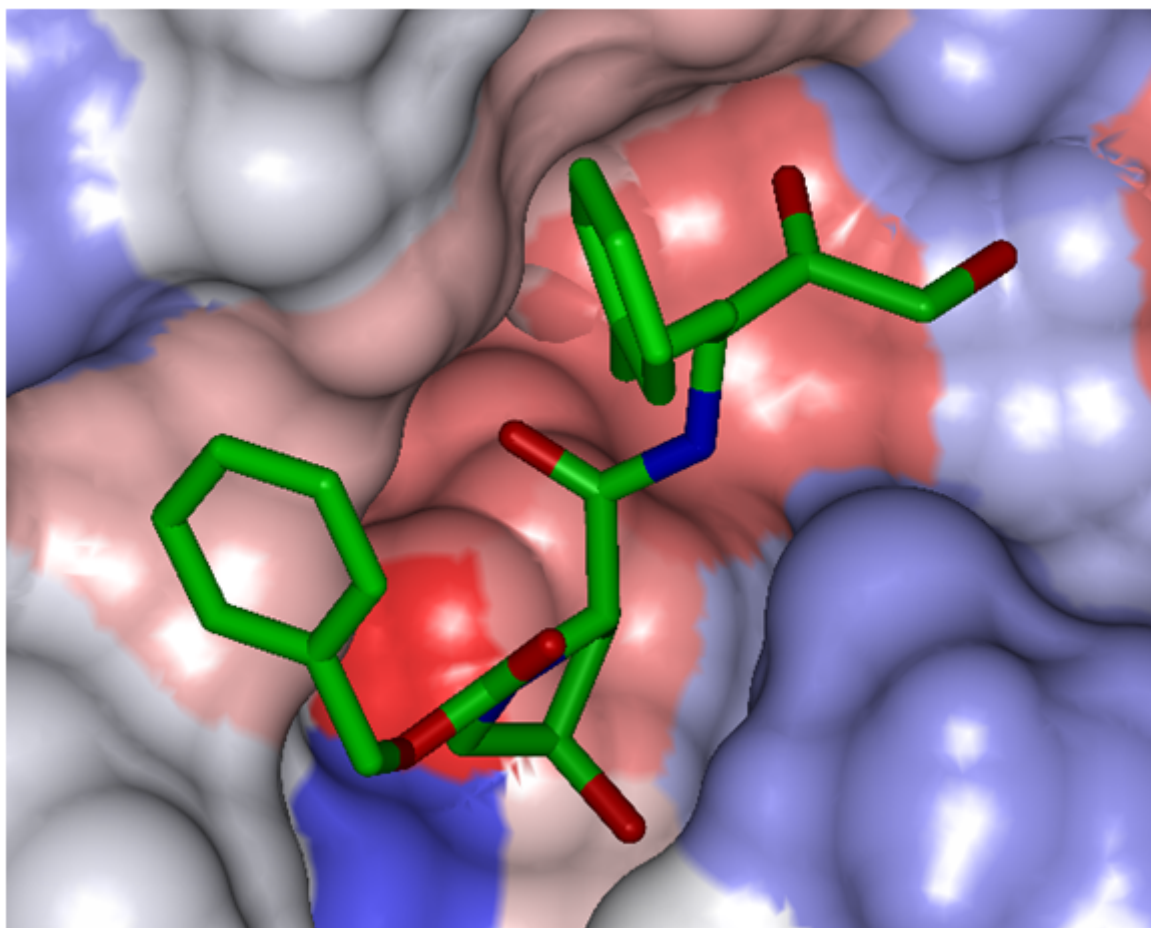


Figure 3.6: Connolly surface of the active site of PfFP-3 with the most potent inhibitor PEP2. The binding site surface is colored according to residue hydrophobicity: red -hydrophobic, blue - hydrophilic and white – intermediate.

thymidine monophosphate kinase [275]. It would rather be suggested that a large and diverse combinatorial virtual library (VL) of PEPs be built and screened with our FP-3 inhibition 3D PH4, based on the complexation one descriptor QSAR model. A successful case study was in the design pyrrolidine carboxamide inhibitors of *Mycobacterium tuberculosis* InhA [231].

3.7 Ligand-based 3D-QSAR PH4 model of FP3 inhibition

In order to build 3D-QSAR pharmacophore based on Catalyst HypoGen algorithm implemented in DS 2.5 [259], the active conformations of inhibitors taken from E:I complexes have been used. Fig: 3.10 shows the active conformations of these inhibitors in the active site of FP-3.

The 3D-QSAR PH4 pharmacophore generation process follows three main steps namely: the constructive, the subtractive and the optimization steps [259]. The constructive phase of HypoGen has automatically selected the most active compounds for which $IC_{50}^{exp} \leq 1.6 \times 60$ nM as leads. Thus,

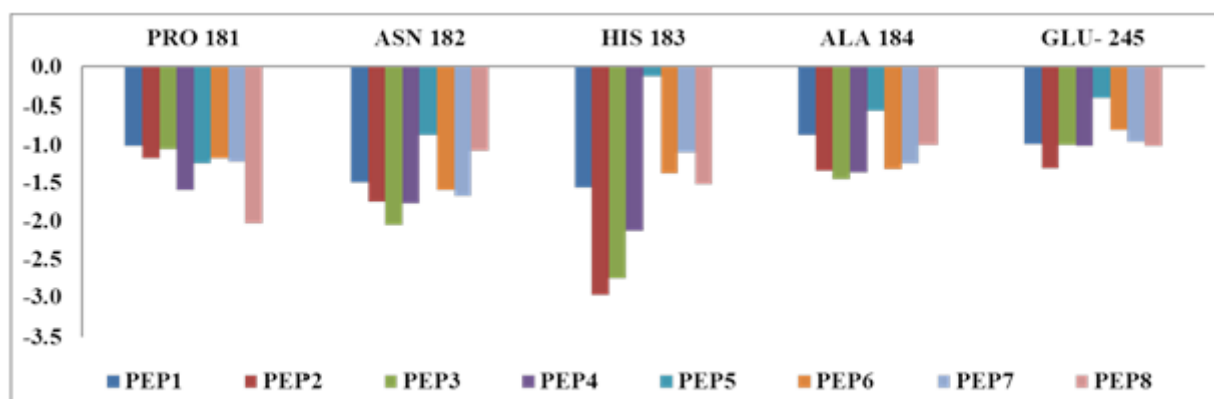


Figure 3.7: Breakdown of FP3-PEP interaction energy into each active site residue contribution at S2 pocket.

only the most active compound PEP2 ($IC_{50}^{exp} = 60$ nM) was used to generate the starting PH4 features. Only those features were retained which matched this lead. In the subtractive phase, which is normally used to remove pharmacophoric features present in poorly active molecules, none of the training set compounds were found inactive ($IC_{50}^{exp} > 60 \times 10^{3.5} = 189736.7$). During the optimization phase, the score of the pharmacophoric hypothesis is improved. Hypotheses are 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 pharmacophoric hypotheses (see Table 3.10) were kept, all displaying four features. The generated pharmacophore models were then assessed for their reliability based on the calculated cost parameters. The overall costs ranged from 37.97 (Hypo1) to 38.94 (Hypo10). The relatively small gap between the highest and lowest cost parameter corresponds well with the homogeneity of the generated hypotheses and the consistency of the training set.

Table 3.10: Output parameters of the 10 generated PH4 pharmacophoric hypotheses for FP-3 inhibitors after CatScramble validation procedure.

| Hypothesis | RMSD ^a | R ^{2b} | Total costs ^c |
|------------|-------------------|-----------------|--------------------------|
| Hypo 1 | 1.024 | 0.978 | 37.97 |
| Hypo 2 | 1.050 | 0.977 | 38.15 |
| Hypo 3 | 1.033 | 0.975 | 38.21 |
| Hypo 4 | 0.994 | 0.977 | 38.32 |
| Hypo 5 | 1.053 | 0.974 | 38.54 |
| Hypo 6 | 1.065 | 0.973 | 38.56 |
| Hypo 7 | 1.075 | 0.973 | 38.63 |
| Hypo 8 | 1.013 | 0.976 | 38.67 |
| Hypo 9 | 1.102 | 0.972 | 38.92 |
| Hypo 10 | 1.098 | 0.971 | 38.94 |
| Fixed Cost | 0.0 | 1 | 33.76 |
| Null Cost | 4.685 | 0.0 | 107.93 |

^aroot mean square deviation; ^bcorrelation coefficient; ^coverallcost parameter of the PH4 pharmacophore.

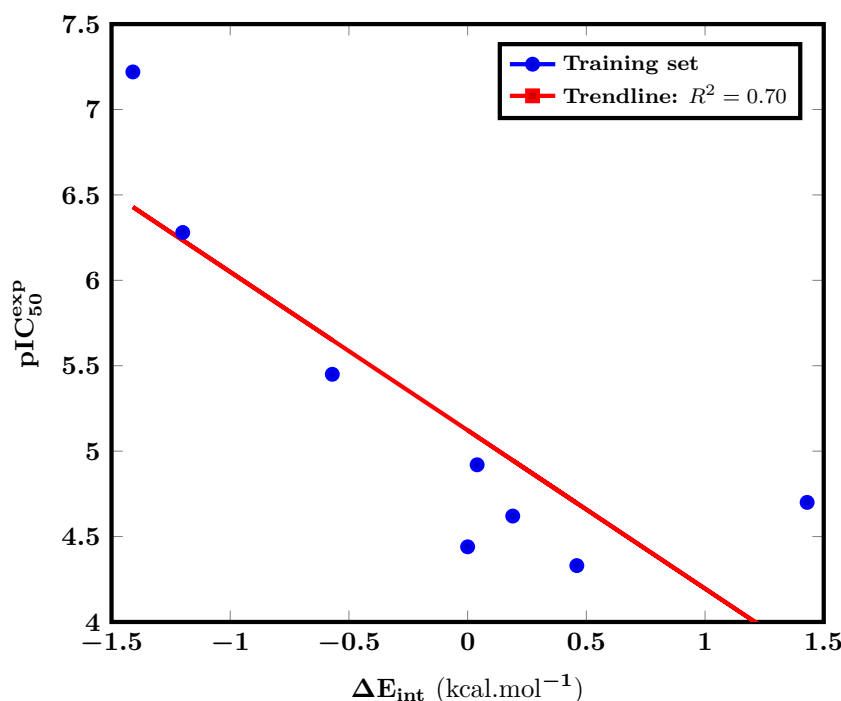


Figure 3.8: Plot of correlation between pIC_{50}^{exp} and $P2$, His 183 contribution to the intermolecular interaction energy of PffFP3-PEPs.

For this PH4 model, the fixed cost (33.76) is lower than the null cost (107.92) by a difference $\Delta = 74.17$. This difference is a major quality indicator of the PH4 predictability ($\Delta > 70$ correspond to an excellent chance or a probability higher than 90% that the model represents a true correlation [259]). To be statistically significant the hypotheses have to be as close as possible to the fixed cost and as far as possible from the null cost. The difference $\Delta \geq 70.03$ for the set of 10 hypotheses confirm the high quality of the pharmacophore model.

The standard indicators such as the root-mean-square deviations (RMSD) between the hypotheses ranged from 1.024 to 1.098 and the squared correlation coefficient (R^2) falls to an interval from 0.978 to 0.971. The first PH4 hypothesis with the 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 3.10. The geometry of the Hypo1 pharmacophore of FP-3 inhibition is displayed in Fig. 3.11. Table 3.12 list the regression equation for pIC_{50}^{exp} vs pIC_{50}^{pred} estimated from Hypo1 $pIC_{50}^{\text{exp}} = -\log_{10}(IC_{50}^{\text{exp}}) = 0.9958 \times pIC_{50}^{\text{pred}} + 0.0219$. The regression equation coefficient is close to one indicating that the ratio of predicted and observed activities $\left(\frac{pIC_{50}^{\text{pred}}}{pIC_{50}^{\text{exp}}}\right)$ for the training set is relatively close to one (Table 3.11). Moreover, the statistical data in Table 3.12 such as R^2 , R_{xy}^2 , greater than 0.9 and a significance Fisher F-test 108.41 document substantial predictive power of this regression for the best PH4 model. Fig. 3.12 display a plot of estimated and experimental activity for PH4 (pIC_{50}^{exp} vs pIC_{50}^{est}). Also, to check the consistency of the generated pharmacophore model we have computed the ratio of predicted and observed activities $\left(\frac{pIC_{50}^{\text{pred}}}{pIC_{50}^{\text{exp}}}\right)$ for the validation set. The computed ratios are as follows: PEP9: 0.95; PEP10: 1.10 all of them relatively close to one, which documents substantial predictive power of the

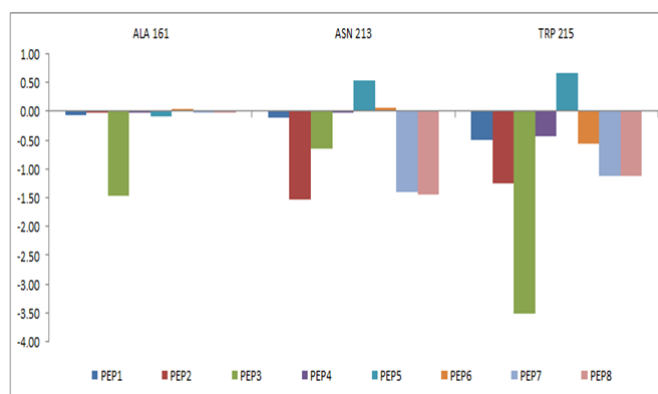


Figure 3.9: Breakdown of FP3-PEP interaction energy into each active site residue contribution at SI' pocket.

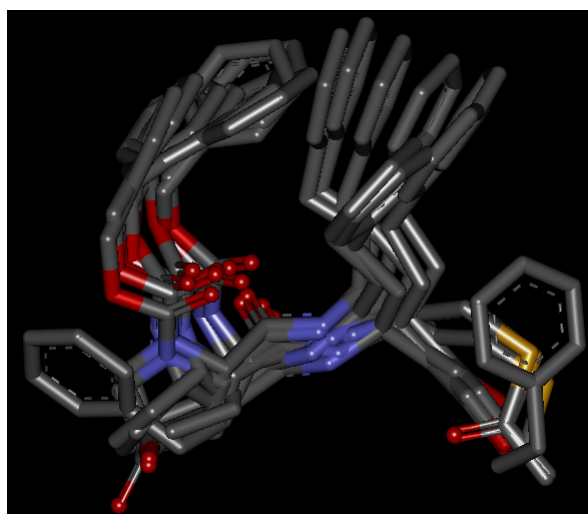


Figure 3.10: Bound conformations of inhibitors taken from E:I complexes

regression for the best PH4 model.

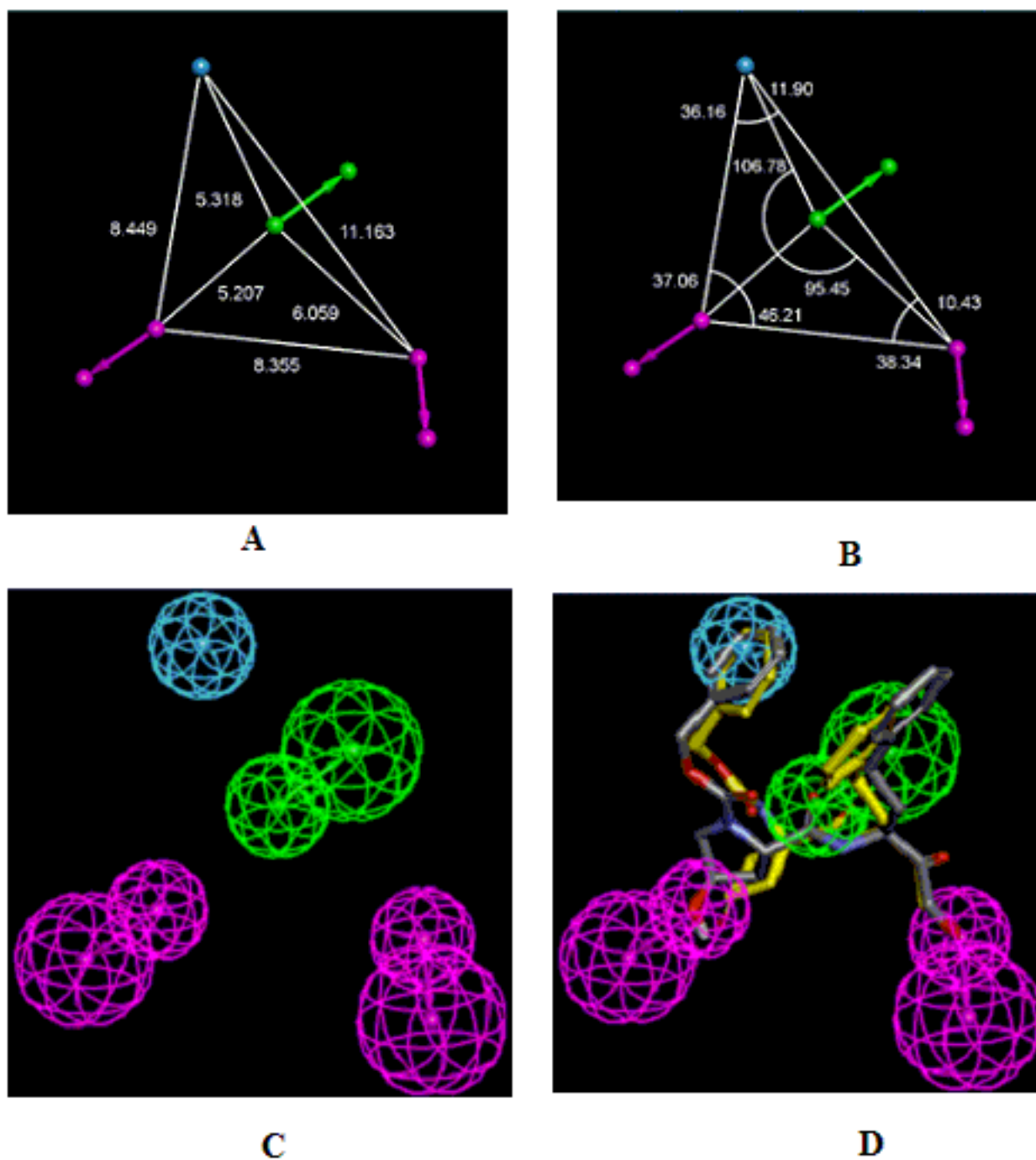


Figure 3.11: Coordinates (A,B) of the Hypo1 pharmacophore of FP-3 inhibition, (C) features of the pharmacophore of FP-3 inhibition, (D) pharmacophore mapping with PEP2 (with blue carbons) and PEP7 (with yellow carbons). Color legend of features: hydrophobic (blue), acceptor (green), donor (purple).

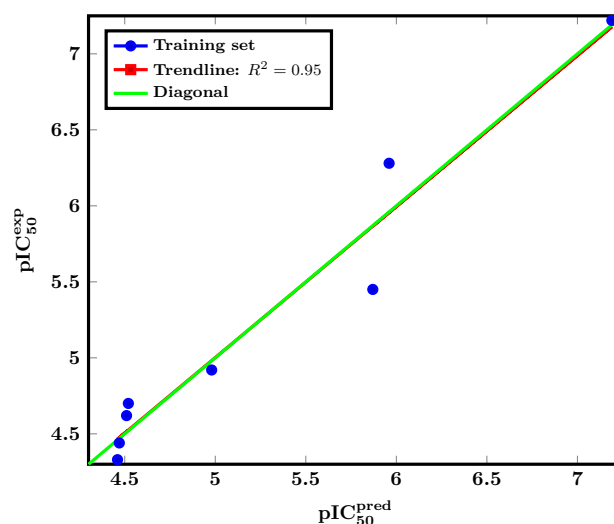


Figure 3.12: Plot of estimated and experimental activity for PH4.

Table 3.11: Computed the ratio of predicted and observed activities ($\frac{pIC_{50}^{pred}}{pIC_{50}^{exp}}$) for the training set.

| Inhibitor (I) | pIC_{50}^{exp} | pIC_{50}^{pred} | $\frac{pIC_{50}^{pred}}{pIC_{50}^{exp}}$ |
|---------------|------------------|-------------------|--|
| PEP 1 | 4.44 | 4.47 | 1.01 |
| PEP 2 | 7.22 | 7.19 | 0.99 |
| PEP 3 | 6.28 | 5.96 | 0.95 |
| PEP 4 | 5.45 | 5.87 | 1.08 |
| PEP 5 | 4.70 | 4.52 | 0.96 |
| PEP 6 | 4.62 | 4.51 | 0.98 |
| PEP 7 | 4.33 | 4.46 | 1.03 |
| PEP 8 | 4.92 | 4.98 | 1.01 |

Table 3.12: Statistical data on regression analysis of correlation for the training set between PH4 predicted activity (pIC_{50}^{exp} vs pIC_{50}^{pred}) and experimental one (pIC_{50}^{exp}) against FP-3.

| | |
|--|----------|
| $pIC_{50}^{exp} = \log_{10}IC_{50}^{exp} = 0.9958 \times pIC_{50}^{pred} + 0.0219$ | |
| Statistical data of linear regression: | |
| Number of compounds n | 8 |
| Squared correlation coefficient of regression R^2 | 0.95 |
| Leave-one-out cross-validated squared correlation Coefficient R_{xv}^2 | 0.91 |
| Standard error of the regression σ | 0.25 |
| Statistical significance of regression, Fisher F-test | 108.41 |
| Level of statistical significance α | >95% |
| Range of activities of IC_{50}^{exp} (nM) | 60–47230 |

3.8 Library design and ADME focusing

In order to identify more potent orally bioavailable *PfFP-3* inhibitors, We have built a virtual library of new analogues inhibitors of *PfFP-3* based on substitutions at two positions (P1', and P2) of a scaffold of a lead compound PEP2 in order to better fit S1' and S1 pockets. This virtual library was built to identifying more potent orally bioavailable *PfFP-3* inhibitors. The 35 R-groups listed in Fig. 3.13 have been attached in positions R1 to R5 of the appropriate scaffold to form a combinatorial library of the size: $R1 \times R2 \times R3 \times R4 \times R5 = 35 \times 35 \times 35 \times 35 \times 35 = 52,521,875$ PEPA. It should be noted that one of the important criteria for the design of new anti-malarials, is their oral bioavailability. Thus in the goal to design a more focused library of a reduced size and increased content of drug-like and orally bioavailable molecules, we have introduced a set of filters and penalties, which can help to select smaller number of suitable PEPs which can undergo *in silico* screening. The initial virtual library has been then filtered in an ADME-based focusing step to remove compounds with expected poor oral bioavailability and low drug likeness. Only compounds not violating any rule of Lipinski's [277] computed for the entire virtual library using QikProp software [271], were kept. This focusing has reduced the size of the initial library to 125,887 PEPA less than 5% its original number size.

3.9 Screening PEPs virtual library using the obtained in silico model

The library of PEP analogues has been further screened for molecular structures matching to the 3D-QSAR PH4 pharmacophore model Hypo1 of FP-3 inhibition. From the set of 125,887 analogues, few thousands of PEPAs mapped to at least 2 features, 592 of which mapped to 4 features of the pharmacophore. Out of then, only 80 best fitting analogues (PH4 hits) have been retained and submitted to screening with help of the complexation QSAR model. Their Gibbs free energy (GFE) upon complex formation with *PfFP-3* has been computed along with its component and their predicted half-maximal inhibitory concentration IC_{50}^{pred} has been estimated with the correlation equation ($pIC_{50}^{exp} = \log_{10} IC_{50}^{exp} = -0.8037 \times \Delta\Delta G_{comp} + 4.0009$). The results obtain are given in Table 3.13. Of the 80 analogues whose inhibitory activities were predicted in Table 7, 30 showed better activities than the most active compound of the training set among them three showed even more activity PEP-2-1-1-16-2 $IC_{50}^{pred} = 0.004$ nM, PEP-2-1-12-21-30 $IC_{50}^{pred} = 0.04$ nM and PEP-2-1-30-7-2 $IC_{50}^{pred} = 0.003$ nM.

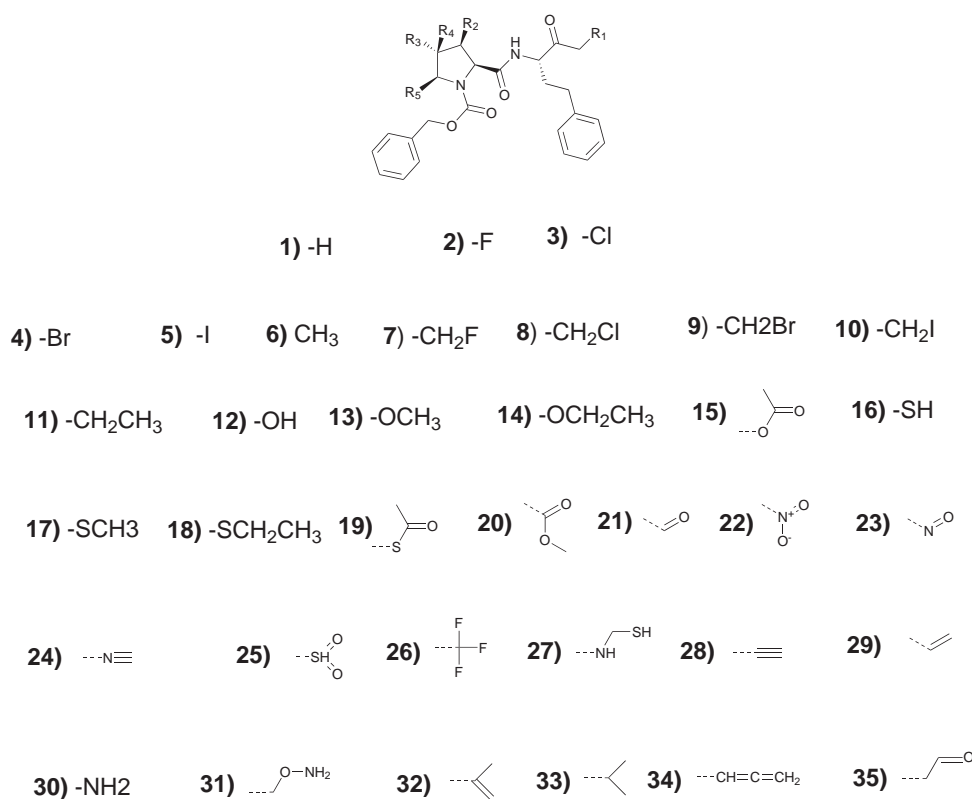


Figure 3.13: R-groups (fragments, building blocks, substituents) used in the design of the initial diversity library of PEP analogues. Dashed bonds indicate the attachment points of the fragments.

Table 3.13: Molar masses, complexation Gibbs free energies and their components for the top 80 scoring virtually designed analogues.

| Analogue ^a | M _w ^b (g.mol ⁻¹) | ΔΔH _{MM} ^c (kcal.mol ⁻¹) | ΔΔG _{sol} ^d (kcal.mol ⁻¹) | ΔΔTS _{vib} ^e (kcal.mol ⁻¹) | ΔΔG _{comp} ^f (kcal.mol ⁻¹) | pIC ₅₀ ^g (nM) |
|-----------------------|---|---|--|---|---|--|
| PEP 1 | 482 | 0 | 0 | 0 | 0 | 47230 |
| 12-1-6-16-30 | 485 | -2.52 | 2.20 | 2.93 | 3.25 | 242.01 |
| 12-1-1-16-30 | 471 | -2.01 | 0.89 | 2.93 | -4.05 | 55.76 |
| 30-6-1-16-29 | 495 | 0.58 | 0.07 | 1.30 | -0.66 | 29515.86 |
| 27-1-1-30-6 | 498 | -4.90 | -0.92 | 0.90 | -6.71 | 0.40 |
| 30-11-1-16-30 | 498 | -0.83 | 0.74 | 0.26 | -0.35 | 52240.29 |
| 31-1-12-13-1 | 485 | -5.72 | -0.27 | 1.31 | -7.30 | 0.14 |
| 12-1-12-13-1 | 470 | -5.91 | 0.36 | 0.29 | -5.84 | 2.02 |
| 30-30-6-12-30 | 483 | -0.13 | -0.05 | 0.46 | 0.28 | 168136.11 |
| 16-21-30-30-1 | 498 | 3.33 | 2.72 | -4.98 | 1.07 | 725268.72 |
| 16-21-24-1-1 | 493 | 2.24 | 2.50 | 6.56 | -1.82 | 3429.37 |
| 16-1-21-12-30 | 499 | -5.11 | 1.90 | 1.62 | -4.83 | 13.20 |
| 16-1-12-1-13 | 486 | -2.48 | 2.20 | 1.90 | -2.19 | 1749.69 |
| 30-6-12-1-12 | 469 | -2.43 | 1.59 | 3.34 | -4.18 | 43.98 |
| 16-1-12-1-1 | 456 | -0.29 | -0.31 | 3.82 | -4.42 | 28.06 |

^aAnalogue; ^bM_w is molecular mass of the inhibitor; ^cΔΔH_{MM} is the relative enthalpic contribution to the Gibbs free energy change related to the FP3-PEP complex formation ΔΔG_{comp}; ^dΔΔG_{sol} is the relative solvation Gibbs free energy contribution to ΔΔG_{comp}; ^eΔΔTS_{vib} is the relative entropic (vibrational) contribution to ΔΔG_{comp}; ^fΔΔG_{comp} is the relative Gibbs free energy change related to the enzyme-inhibitor FP3-PEP complex formation ΔΔG_{comp} = ΔΔH_{MM} + ΔΔG_{sol} - ΔΔTS_{vib}; ^gIC₅₀^{pre} is the predicted inhibition constant towards PfFP-3 calculated from ΔΔG_{comp}, using correlation equation (pIC₅₀^{exp} = log₁₀ IC₅₀^{exp} = -0.8037 × ΔΔG_{comp} + 4.0009).

Table 3.14: Table 3.13 continued.

| Analog ^a | M_w ^b ($g.mol^{-1}$) | $\Delta\Delta H_{MM}$ ^c ($kcal.mol^{-1}$) | $\Delta\Delta G_{sol}$ ^d ($kcal.mol^{-1}$) | $\Delta\Delta T S_{vib}$ ^e ($kcal.mol^{-1}$) | $\Delta\Delta G_{comp}$ ^f ($kcal.mol^{-1}$) | pIC_{50} ^g (nM) |
|---------------------|--|---|--|--|---|---------------------------------|
| PEP 1 | 482 | 0 | 0 | 0 | 0 - | 47230 |
| 30-1-12-6-12 | 469 | -5.79 | 1.52 | -0.36 | -3.91 | 72.03 |
| 30-1-12-24-13 | 494 | -3.05 | 1.79 | 1.20 | -2.46 | 1045.39 |
| 30-1-12-7-29 | 497 | -5.14 | 1.48 | 1.02 | -4.68 | 17.43 |
| 30-1-3-12-29 | 500 | -4.53 | 1.04 | 1.87 | -5.36 | 4.94 |
| 30-1-1-16-3 | 490 | -4.14 | 0.14 | 0.85 | -4.85 | 12.58 |
| 30-1-1-16-29 | 481 | -3.67 | 1.14 | 1.63 | -4.16 | 45.26 |
| 30-1-1-30-7 | 452 | 0.38 | 0.22 | 1.20 | -0.61 | 32506.46 |
| 30-1-2-24-29 | 492 | -5.05 | 0.53 | -0.19 | -4.32 | 33.52 |
| 16-29-30-1-30 | 496 | 2.53 | 0.73 | 3.39 | -0.13 | 77834.04 |
| 17-21-12-1-1 | 498 | 0.87 | 2.62 | 4.30 | -0.81 | 22383.93 |
| 16-31-1-6-1 | 499 | 1.52 | 2.23 | 3.36 | 0.39 | 204327.86 |
| 16-29-12-1-12 | 498 | 1.03 | 1.46 | 5.53 | -3.03 | 363.75 |
| 16-30-30-1-30 | 485 | 1.30 | 1.18 | 3.63 | -1.15 | 11876.59 |
| 16-30-30-6-30 | 499 | 0.17 | 0.81 | 3.10 | -2.12 | 1979.706 |
| 30-1-30-30-35 | 495 | -5.08 | 2.34 | -0.33 | -2.41 | 1149.01 |
| 30-1-30-6-16 | 484 | -2.11 | 0.25 | -0.49 | -1.38 | 7806.17 |
| 30-1-30-7-29 | 478 | -3.77 | 1.59 | -0.66 | -1.53 | 5853.75 |
| 30-1-35-12-30 | 496 | -3.32 | 4.25 | -2.38 | -1.45 | 6853.57 |
| 30-1-34-1-29 | 475 | -3.14 | -1.31 | -0.77 | -3.69 | 108.35 |
| 30-2-12-6-29 | 514 | -2.99 | 1.40 | -1.66 | 0.07 | 113594.2 |
| 30-2-28-1-12 | 497 | 0.44 | 0.85 | 1.63 | -0.34 | 53132.96 |
| 30-6-12-24-30 | 493 | -1.52 | 0.48 | 2.85 | -3.90 | 73.23 |
| 24-1-12-1-21 | 477 | -4.04 | 0.55 | 3.42 | -6.91 | 0.28 |
| 24-1-6-6-21 | 489 | -3.08 | -0.45 | 0.36 | -3.89 | 75.14 |
| 24-29-30-1-12 | 490 | 0.49 | -0.08 | 1.02 | -0.62 | 31858.85 |

^aAnalogue; ^b M_w is molecular mass of the inhibitor; ^c $\Delta\Delta H_{MM}$ is the relative enthalpic contribution to the Gibbs free energy change related to the FP3-PEP complex formation $\Delta\Delta G_{comp}$; ^d $\Delta\Delta G_{sol}$ is the relative solvation Gibbs free energy contribution to $\Delta\Delta G_{comp}$; ^e $\Delta\Delta T S_{vib}$ is the relative entropic (vibrational) contribution to $\Delta\Delta G_{comp}$; ^f $\Delta\Delta G_{comp}$ is the relative Gibbs free energy change related to the enzyme-inhibitor FP3-PEP complex formation $\Delta\Delta G_{comp} = \Delta\Delta H_{MM} + \Delta\Delta G_{sol} - \Delta\Delta T S_{vib}$; ^g IC_{50}^{pre} is the predicted inhibition constant towards PfFP-3 calculated from $\Delta\Delta G_{comp}$, using correlation equation ($pIC_{50}^{exp} = \log_{10} IC_{50}^{exp} = -0.8037 \times \Delta\Delta G_{comp} + 4.0009$).

Table 3.15: Table 3.13 continued.

| Analog ^a | M_w ^b (g.mol ⁻¹) | $\Delta\Delta H_{MM}$ ^c (kcal.mol ⁻¹) | $\Delta\Delta G_{sol}$ ^d (kcal.mol ⁻¹) | $\Delta\Delta T S_{vib}$ ^e (kcal.mol ⁻¹) | $\Delta\Delta G_{comp}$ ^f (kcal.mol ⁻¹) | pIC_{50} ^g (nM) |
|---------------------|--|---|--|--|---|---------------------------------|
| PEP 1 | 482 | 0 | 0 | 0 | 0 - | 47230 |
| 24-21-12-1-12 | 493 | -1.57 | 3.38 | 6.88 | -5.07 | 8.39 |
| 24-21-34-1-1 | 487 | 2.98 | -0.05 | 4.76 | -1.84 | 3337.61 |
| 28-21-12-1-2 | 494 | -3.78 | 5.35 | 6.37 | -4.80 | 13.97 |
| 29-21-6-12-1 | 492 | -6.27 | 4.79 | 3.19 | -4.68 | 17.34 |
| 28-1-31-30-1 | 492 | 5.00 | 1.19 | 5.05 | 1.15 | 833592.8 |
| 28-30-30-30-30 | 492 | -0.37 | 1.97 | 4.39 | -2.79 | 3 567.71 |
| 28-12-31-1-1 | 493 | 2.62 | 2.75 | 6.84 | -1.47 | 6608.46 |
| 16-1-1-35-12 | 498 | -2.58 | 3.49 | 5.05 | -4.14 | 47.07 |
| 16-1-21-12-30 | 499 | -5.23 | 2.54 | 1.01 | -3.69 | 107.40 |
| 30-6-30-29 | 494 | -0.18 | -0.51 | -2.91 | 2.22 | 6070563.65 |
| 12-32-1-12-12 | 496 | 0.78 | 1.18 | -0.23 | 2.19 | 5721021.85 |
| 12-30-12-6-29 | 495 | -1.72 | 3.76 | 0.25 | 1.79 | 2729505.81 |
| 12-30-6-30-13 | 498 | -2.11 | 0.57 | -2.93 | 1.39 | 1295249.99 |
| 12-21-1-30-13 | 497 | -0.41 | 1.05 | 1.13 | -0.50 | 39472.32 |
| 12-21-12-1-13 | 498 | -4.94 | 3.24 | 1.87 | -3.56 | 136.55 |
| 12-21-12-6-12 | 498 | -1.21 | 1.34 | 1.17 | -1.05 | 14416.23 |
| 12-21-24-6-1 | 491 | 2.64 | -0.85 | 3.51 | -1.72 | 4113.86 |
| 12-1-1-12-13 | 470 | -7.81 | 2.59 | 0.77 | -5.98 | 1.55 |
| 12-1-1-35-2 | 484 | -1.26 | 1.65 | 1.02 | -0.63 | 31044.62 |
| 12-11-1-16-30 | 499 | 0.89 | 1.65 | 1.94 | 0.60 | 304896.76 |

^aAnalogue; ^b M_w is molecular mass of the inhibitor; ^c $\Delta\Delta H_{MM}$ is the relative enthalpic contribution to the Gibbs free energy change related to the FP3-PEP complex formation $\Delta\Delta G_{comp}$; ^d $\Delta\Delta G_{sol}$ is the relative solvation Gibbs free energy contribution to $\Delta\Delta G_{comp}$; ^e $\Delta\Delta T S_{vib}$ is the relative entropic (vibrational) contribution to $\Delta\Delta G_{comp}$; ^f $\Delta\Delta G_{comp}$ is the relative Gibbs free energy change related to the enzyme-inhibitor FP3-PEP complex formation $\Delta\Delta G_{comp} = \Delta\Delta H_{MM} + \Delta\Delta G_{sol} - \Delta\Delta T S_{vib}$; ^g IC_{50}^{pre} is the predicted inhibition constant towards PfFP-3 calculated from $\Delta\Delta G_{comp}$, using correlation equation ($pIC_{50}^{exp} = \log_{10} IC_{50}^{pre} = -0.8037 \times \Delta\Delta G_{comp} + 4.0009$).

Table 3.16: Table 3.13 continued.

| Analog ^a | M_w ^b (g.mol ⁻¹) | $\Delta\Delta H_{MM}$ ^c (kcal.mol ⁻¹) | $\Delta\Delta G_{sol}$ ^d (kcal.mol ⁻¹) | $\Delta\Delta T S_{vib}$ ^e (kcal.mol ⁻¹) | $\Delta\Delta G_{comp}$ ^f (kcal.mol ⁻¹) | pIC_{50} ^g (nM) |
|---------------------|--|---|--|--|---|---------------------------------|
| PEP 1 | 482 | 0 | 0 | 0 | 0 - | 47230 |
| 12-11-12-1-29 | 494 | -2.66 | 2.76 | -1.03 | 1.13 | 804378.82 |
| 12-11-12-1-13 | 498 | -4.80 | 2.71 | -0.78 | -1.31 | 8907.34 |
| 12-2-1-12-13 | 488 | -8.52 | 1.68 | -0.65 | -6.19 | 1.056 |
| 12-6-2-30-29 | 497 | -1.18 | -0.83 | 1.92 | -3.93 | 69.46 |
| 12-6-1-12-13 | 480 | -3.48 | 1.75 | 2.40 | -4.13 | 47.92 |
| 16-21-30-1-30 | 498 | 2.03 | 1.40 | 4.78 | -1.35 | 8209.10 |
| 16-6-12-1-23 | 499 | -2.96 | 1.02 | 3.71 | -5.65 | 2.86 |
| 6-21-12-1-30 | 481 | -3.85 | 2.14 | 5.50 | -7.21 | 0.16 |
| 7-21-12-1-30 | 499 | -5.90 | 2.76 | 2.79 | -5.93 | 1.72 |
| 8-21-30-1-1 | 500 | -0.44 | 1.24 | 3.63 | -2.84 | 523.39 |
| 11-1-30-30-21 | 494 | -3.30 | 0.08 | 2.44 | -5.66 | 2.84 |
| 11-21-12-1-2 | 498 | -5.74 | 2.44 | 1.36 | -4.66 | 17.92 |
| 12-1-16-30-28 | 495 | -3.01 | 2.67 | 0.32 | -0.65 | 29703.71 |
| 2-1-1-16-2 | 476 | -3.46 | -0.58 | 5.19 | -9.22 | 0.004 |
| 2-1-12-21-30 | 485 | -6.91 | 1.70 | 2.75 | -7.96 | 0.04 |
| 2-1-30-7-2 | 491 | -6.13 | 0.50 | 3.76 | -9.39 | 0.003 |
| 2-2-12-1-30 | 475 | -4.78 | 1.51 | 4.03 | -7.31 | 0.13 |
| 2-6-12-21-30 | 499 | -3.41 | 0.66 | 3.72 | -6.47 | 0.63 |
| 3-1-12-1-13 | 488 | -5.63 | 1.72 | 2.33 | -6.24 | 0.96 |
| 3-24-12-1-30 | 498 | 1.85 | 2.75 | 5.19 | -0.59 | 33461.37 |
| 6-1-30-7-2 | 487 | -6.91 | 0.44 | -1.12 | -5.35 | 5.03 |

^aAnalogue; ^b M_w is molecular mass of the inhibitor; ^c $\Delta\Delta H_{MM}$ is the relative enthalpic contribution to the Gibbs free energy change related to the FP3-PEP complex formation $\Delta\Delta G_{comp}$; ^d $\Delta\Delta G_{sol}$ is the relative solvation Gibbs free energy contribution to $\Delta\Delta G_{comp}$; ^e $\Delta\Delta T S_{vib}$ is the relative entropic (vibrational) contribution to $\Delta\Delta G_{comp}$; ^f $\Delta\Delta G_{comp}$ is the relative Gibbs free energy change related to the enzyme-inhibitor FP3-PEP complex formation $\Delta\Delta G_{comp} = \Delta\Delta H_{MM} + \Delta\Delta G_{sol} - \Delta\Delta T S_{vib}$; ^g IC_{50}^{pre} is the predicted inhibition constant towards P_fFP-3 calculated from $\Delta\Delta G_{comp}$, using correlation equation ($pIC_{50}^{exp} = \log_{10} IC_{50}^{exp} = -0.8037 \times \Delta\Delta G_{comp} + 4.0009$).

3.10 Analysis of new inhibitors

In order to identify the substituents that make the analogues predicted to be active, we have analyzed the frequency of occurrence of certain substituents chosen from figure 3.13, on the predicted active analogues. From the 7 best analogues proposed (seen chemical structure in Figure 3.14) the following R-groups are present 1,2, 3, 6, 7,12, 13, 16,21,24,30 and 31. And Figure 3.15 display the best virtual hit, analogue PEP-2-1-30-7-2 and the least active PEP7 mapped a PH4. Figure 3.16 display 2D schematic interaction diagram of the most potent analogue design at the active site of P_fFP3. Figure 3.17 shows a Connolly surfaces of the active site of P_fFP-3 with the predicted most active PEP inhibitor 2-1-30-7-2.

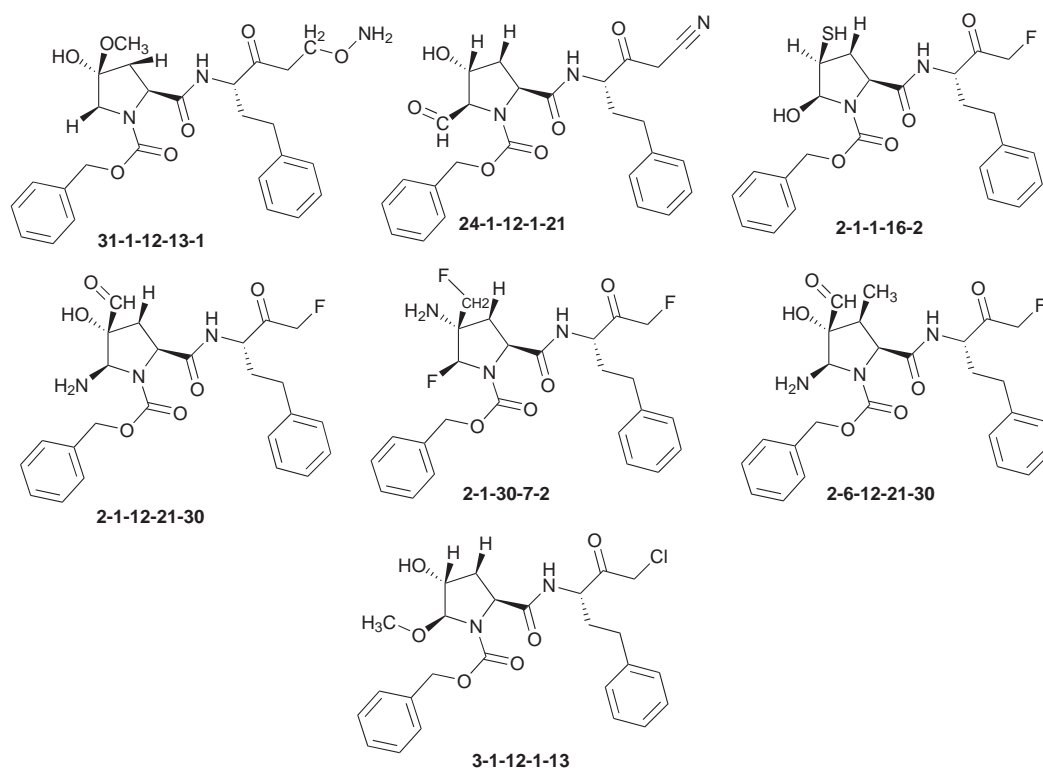


Figure 3.14: Chemical structures towards Pf FP-3 of seven most potent PEP analogues.

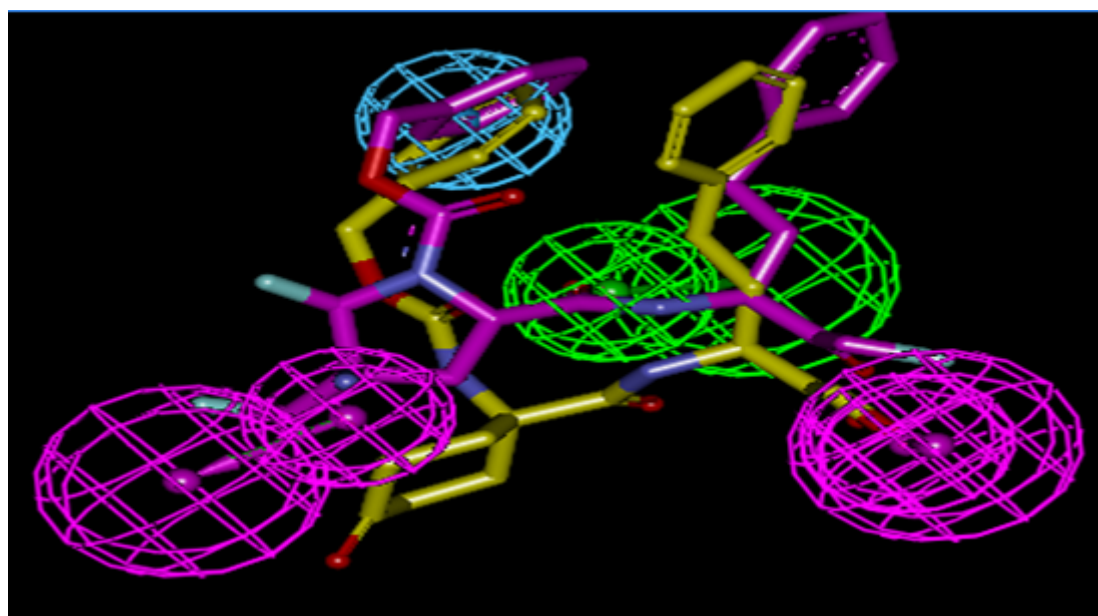


Figure 3.15: The best virtual hit, analogue PEP-2-1-30-7-2 (with purple carbons), the inhibitor PEP7 (Figure 1) shown in yellow carbons, mapped a PH4.

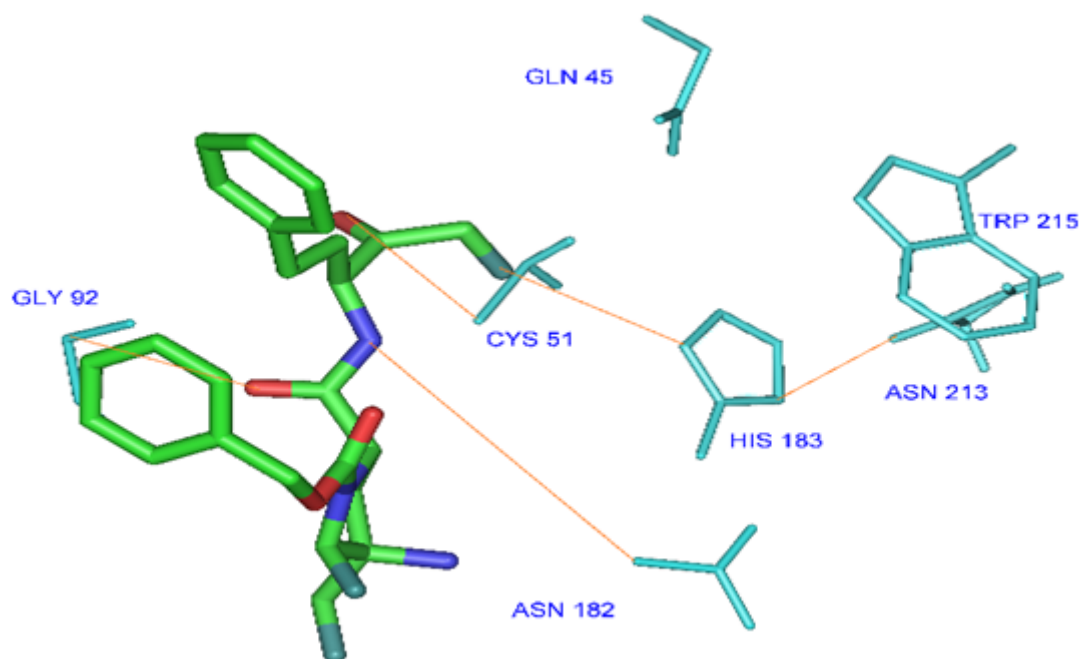


Figure 3.16: 2D schematic interaction diagram of the most potent analogue design PEP-2-1-30-7-2 at the active site of PfFP3.

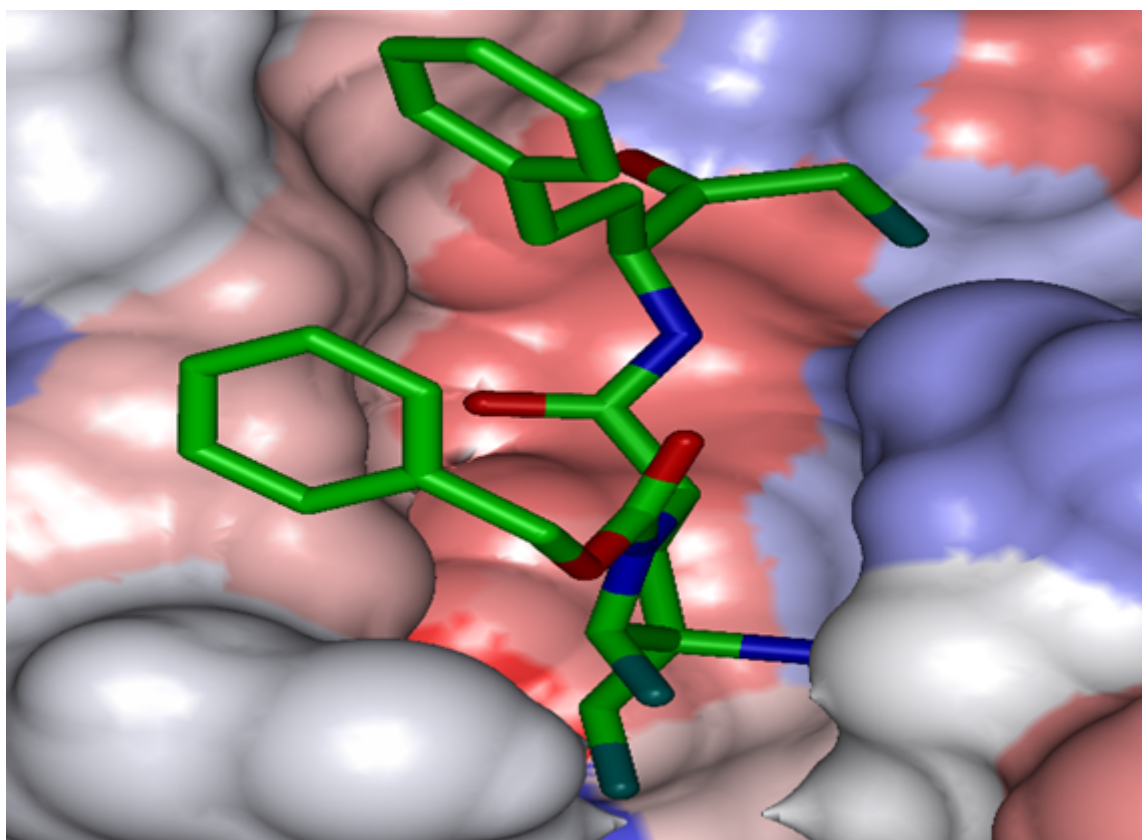


Figure 3.17: Connolly surface of the active site of PfFP-3 with the predicted most active PEP inhibitor PEP-2-1-30-7-2. The binding site surface is colored according to residue hydrophobicity: red - hydrophobic, blue - hydrophilic and white-intermediate.

3.10.1 Predicted ADMET profiles of designed PEPs

The ADMET values for the best active designed PEPs were compared with those computed for drugs used for the treatment of malaria individually or in Artemisinin combined therapy (ACT) or are currently undergoing clinical trials, Table 3.17. Among the 11 analogues compared, 9 display #stars descriptor equal to zero, meaning that the optimal value range of none of the drug-likeness descriptors has been violated. Additionally, oral bioavailability is among the main requirements of the WHO with regard to new drugs in general also anti-malarial drugs. In the last column of Table 3.17, a high level drug-likeness descriptor is displayed. It is the percentage of human oral absorption in gastrointestinal tract (HOA). All the best provisionally active compounds used for comparison have this descriptor comprise between 58.3% and 100% (< 25% - poor, > 80% high) is the range of 95% of drugs. Interesting fact, the #stars is equal to 0 for the designed PEP analogues which is not the case to those of most of the ACT antimalarials.

Table 3.17: Predicted ADME-related properties of the best-designed PEP analogues.

| Analogues ^a | #Star ^b | Mw ^c (g·mol ⁻¹) | Smol ^d [Å ²] | SmolInfo ^e [Å ²] | Vmol ^f [Å ³] | RotB ^g | HBdon ^h | HBacc ⁱ | LogPOW ^j | LogSwat ^k | LogKSHA ^l | LogB/B ^m | BPaco ⁿ [nm·s ⁻¹] | ≠meta ^o | IC ₅₀ ^{pred} | HOA ^p | %HOA ^q |
|------------------------|--------------------|--|-------------------------------------|---|-------------------------------------|-------------------|--------------------|--------------------|---------------------|----------------------|----------------------|---------------------|--|--------------------|----------------------------------|------------------|-------------------|
| 27-1-1-30-6 | 0 | 510.7 | 818 | 632.6 | 1622.7 | 13 | 4 | 9 | 2.71 | -2.8 | 0.22 | -0.7 | 23.4 | 6 | 0.40 | 2 | 59 |
| 31-1-12-13-1 | 0 | 497.6 | 774.4 | 594.2 | 1543.5 | 13 | 2 | 9 | 3.23 | -4.7 | 0.17 | -1.9 | 160.6 | 3 | 0.14 | 3 | 85.3 |
| 12-1-12-13-1 | 0 | 482.6 | 795.5 | 642 | 1538.5 | 12 | 2 | 9 | 3.16 | -4.7 | 0.08 | -1.7 | 245.9 | 3 | 2.02 | 3 | 88.2 |
| 24-1-12-1-21 | 1 | 489.6 | 806.9 | 523.7 | 1561.6 | 12 | 0 | 10 | 2.14 | -5.1 | -0.24 | -2.9 | 21.4 | 5 | 0.28 | 2 | 63.3 |
| 6-21-12-1-30 | 0 | 493.6 | 808 | 591.6 | 1571.8 | 12 | 1 | 9 | 2.72 | -3.9 | 0.26 | -1.9 | 17.1 | 6 | 0.16 | 2 | 65 |
| 2-1-1-16-2 | 1 | 488.6 | 847.4 | 622.5 | 1562.2 | 9 | 1 | 6 | 5.5 | -7.8 | 0.80 | -0.73 | 904 | 4 | 0.004 | 1 | 100 |
| 2-1-12-21-30 | 0 | 497.6 | 866.1 | 607.9 | 1606 | 11 | 2 | 9 | 2.9 | -5.2 | 0.34 | -2.1 | 15.2 | 5 | 0.04 | 2 | 65.1 |
| 2-1-30-7-2 | 0 | 503.6 | 813.6 | 593.7 | 1546.8 | 9 | 2 | 7 | 4.2 | -5.7 | 0.63 | -0.6 | 150.3 | 3 | 0.003 | 3 | 77.4 |
| 2-2-12-1-30 | 0 | 487.6 | 762.4 | 525.9 | 1490.5 | 10 | 2 | 8 | 3.05 | -3.9 | 0.34 | -1.17 | 39.07 | 5 | 0.13 | 3 | 73.3 |
| 2-6-12-21-30 | 0 | 511.6 | 881.1 | 649.9 | 1649.1 | 11 | 2 | 9 | 3.35 | -5.18 | 0.46 | -1.77 | 24.02 | 5 | 0.63 | 2 | 58.3 |
| 3-1-12-1-13 | 0 | 501.1 | 793.1 | 600.6 | 1537.9 | 10 | 1 | 9.7 | 3.67 | -5.47 | 0.11 | -1.08 | 614.1 | 4 | 0.96 | 3 | 85.4 |
| Chloroquine | 1 | 293.7 | 594.1 | 188.9 | 982.9 | 6 | 0 | 3 | 4.56 | -5.3 | 0.41 | -0.1 | 3718.1 | 0 | - | 3 | 100 |
| Amodiaquine | 1 | 333.7 | 603.2 | 131.7 | 1018.7 | 6 | 0 | 5 | 3.61 | -4.4 | -0.02 | -0.4 | 1689.1 | 0 | - | 3 | 100 |
| Dapsone | 1 | 236.2 | 431.6 | 0 | 687.9 | 2 | 0 | 7 | -0.37 | -0.5 | -1.34 | -0.9 | 289.1 | 0 | - | 3 | 68.8 |
| Trimethoprim | 0 | 272.2 | 500.2 | 223.9 | 835.9 | 5 | 0 | 6.5 | 0.59 | -1.5 | -0.91 | -1.2 | 282.8 | 3 | - | 3 | 74.3 |
| Mefloquine | 2 | 362.2 | 533.1 | 0 | 925.1 | 2 | 0 | 4 | 4.14 | -4.9 | 0.15 | 0.5 | 2903.1 | 0 | - | 3 | 100 |
| pamaquine | 0 | 315.5 | 654.8 | 443.4 | 1148.1 | 9 | 1 | 4.75 | 4.02 | -3.8 | 0.43 | 0.2 | 1475.2 | 5 | - | 3 | 100 |
| Sulfametyoprazine | 1 | 268.2 | 473.4 | 77.8 | 773.3 | 4 | 0 | 9 | -1.03 | 0.2 | -1.7 | -1.3 | 195.8 | 1 | - | 2 | 61.9 |
| Quinacrine | 0 | 369.7 | 680.5 | 268.8 | 1163.6 | 7 | 0 | 3.5 | 5.57 | -6.5 | 0.8 | -0.1 | 4435.7 | 1 | - | 1 | 100 |
| Tetracycline | 5 | 422.3 | 604.5 | 173.1 | 1111.8 | 2 | 0 | 7 | -3.43 | 1.1 | -2.5 | -2.6 | 6.8 | 5 | - | 1 | 21.8 |
| Lumefantrine | 5 | 496.7 | 819.1 | 160.7 | 1437.5 | 7 | 0 | 3 | 8.27 | -10.01 | 1.7 | 0.2 | 4337.2 | 0 | - | 1 | 100 |
| Bulaquine | 0 | 369.5 | 560.2 | 360.2 | 1097.8 | 9 | 1 | 5.8 | 3.62 | -2.98 | 0.1 | -0.4 | 3099.7 | 7 | - | 3 | 100 |
| Hydroxychloroquine | 1 | 309.7 | 609.5 | 119.5 | 1006.5 | 6 | 0 | 5 | 3.36 | -4.51 | -0.1 | -0.7 | 1023.7 | 0 | - | 3 | 100 |
| Sulfadoxine | 1 | 296.2 | 510.6 | 152.3 | 849.5 | 5 | 0 | 9.5 | -0.79 | -0.11 | -1.7 | -1.4 | 213.4 | 2 | - | 2 | 63.9 |
| Halofantrine | 5 | 470.2 | 785.4 | 160.2 | 1351.8 | 5 | 0 | 3 | 7.63 | -9.9 | 1.5 | 0.2 | 2844.1 | 0 | - | 1 | 100 |
| Proguanil | 1 | 237.6 | 478.2 | 125.3 | 768.6 | 6 | 0 | 6 | 1.09 | -1.5 | -1.1 | -0.7 | 834.6 | 0 | - | 3 | 85.6 |
| Doxycycline | 4 | 422.3 | 602.2 | 174.1 | 1104.2 | 2 | 0 | 17.2 | -3.99 | 1.73 | -2.88 | -2.45 | 9.17 | 4 | - | 1 | 20.8 |
| arteether | 1 | 312.4 | 531.1 | 506.1 | 970.2 | 2 | 0 | 5.7 | 2.7 | -2.99 | -0.2 | 0.2 | 5731.8 | 0 | - | 3 | 100 |
| dihydroartemisinin | 1 | 284.4 | 477.4 | 395.7 | 864.6 | 1 | 1 | 5.7 | 1.84 | -2.92 | -0.1 | -0.1 | 1664.9 | 0 | - | 3 | 95.4 |
| Artemisinin | 0 | 282 | 456.6 | 380.6 | 848.4 | 0 | 0 | 5.3 | 1.7 | -2.1 | -0.3 | 0.001 | 1886 | 1 | - | 3 | 95.8 |

a) best designed APPs analogues, Table 7; b) drug likeness, number of property descriptors (from 24 out of the full list of 49 descriptors of QikProp, ver. 3.7, release [14]) that fall outside of the range of values for 95% of known drugs; c) molecular weight in g·mol⁻¹ (range for 95% of drugs: 130-725 g·mol⁻¹); d) total solvent-accessible molecular surface, in Å² (probe radius 1.4 Å) (range for 95% of drugs: 300-1000 Å²); e) hydrophobic portion of the solvent-accessible molecular surface, in Å² (probe radius 1.4 Å) (range for 95% of drugs: 0-750 Å²); f) total volume of molecule enclosed by solvent-accessible molecular surface, in Å³ (probe radius 1.4 Å) (range for 95% of drugs: 500-2000 Å³); g) number of non-trivial (not CX3), non-hindered (not alkene, amide, small ring) rotatable bonds (range for 95% of drugs: 0-15); h) estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer (range for 95% of drugs: 0.0-6.0); i) estimated number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer (range for 95% of drugs: 0.0-6.0); j) logarithm of partitioning coefficient between n-octanol and water phases (range for 95% of drugs: -2 - 6.5); k) logarithm of predicted aqueous solubility, log S, in mol dm⁻³ is the concentration of the solute in a saturated solution that is in equilibrium with the crystalline solid (range for 95% of drugs: -6.0 - 0.5); l) logarithm of predicted binding constant to human serum albumin (range for 95% of drugs: -1.5 - 1.5); m) logarithm of predicted brain/blood partition coefficient. Note: QikProp predictions are for orally delivered drugs so, for example, dopamine and serotonin are CNS negative because they are too polar to cross the blood-brain barrier (range for 95% of drugs: -3.0-1.2); n) predicted apparent Caco-2 cell membrane permeability in Boehringer-Ingelheim scale, in nm/s (range for 95% of drugs: < 2500 great); o) number of likely metabolic reactions (range for 95% of drugs: 1-8); p) predicted inhibition constants (IC₅₀^{pred} (nM)). IC₅₀^{pred} was predicted from computed $\Delta\Delta G_{comp}$, using the regression equation shown in Table 5; q) human oral absorption (1 - low, 2 - medium, 3 - high); percentage of human oral absorption in gastrointestinal tract (<25% - poor, >80% high).

(#) Star indicating that the property descriptor value falls outside the range of values for 95% of known drugs.

The following figure show the 2D interaction diagrams between the best previously propose inhibitors and the active site residus of FP-3.

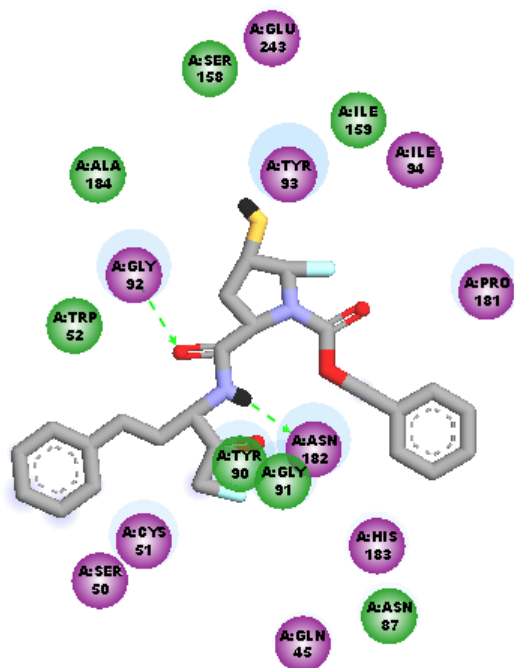


Figure 3.18: 2D interaction diagram of PEP-2-1-1-16-2 hydrogen bonds (HBs) (black, blue and green dashed line)

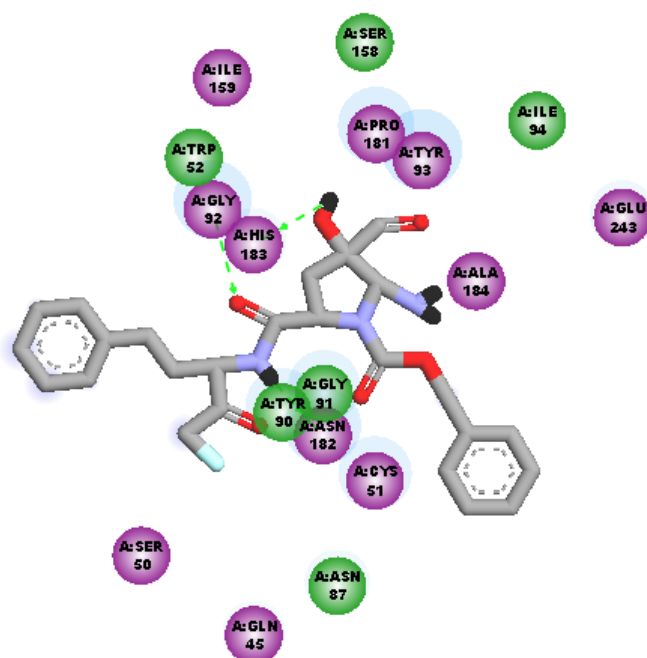


Figure 3.19: 2D interaction diagram of PEP-2-1-12-21-30 hydrogen bonds (HBs) (black, blue and green dashed line).

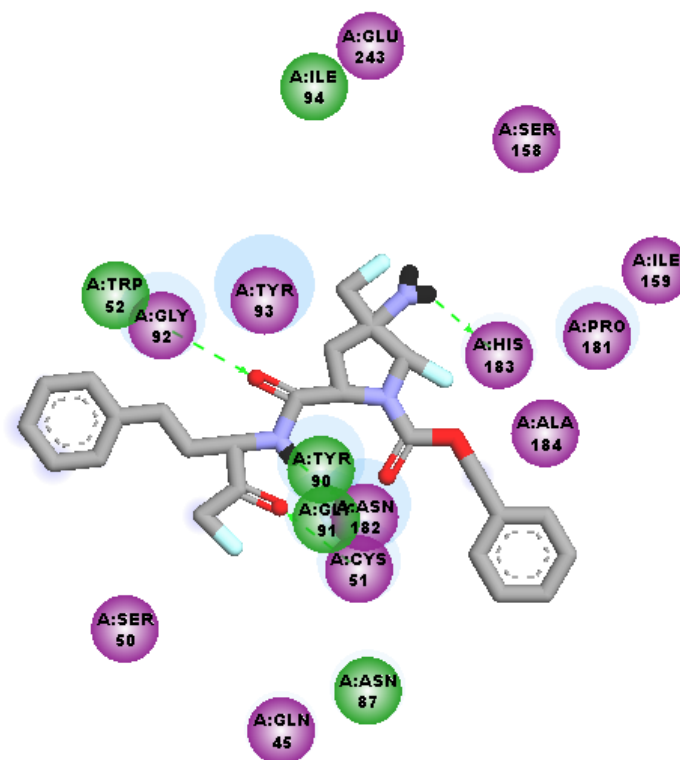


Figure 3.20: 2D interaction diagram of PEP-2-1-30-7-2 hydrogen bonds (HBs) (black, blue and green dashed line).

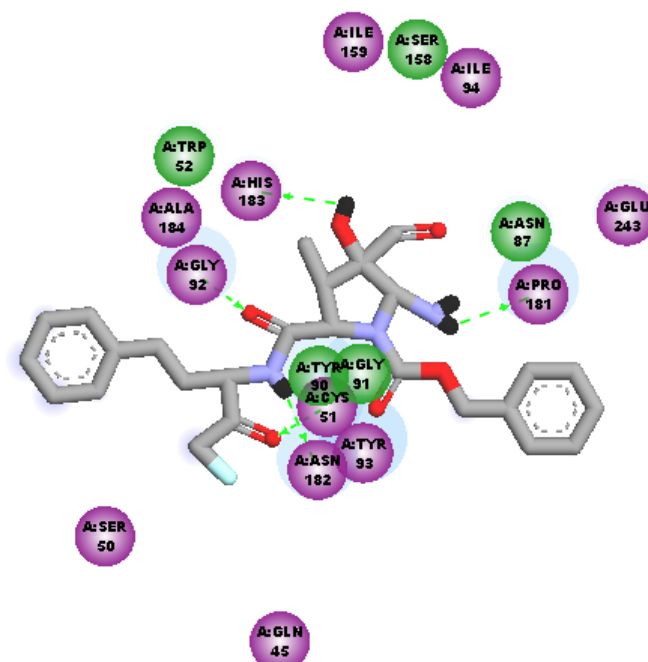


Figure 3.21: 2D interaction diagram of PEP-2-6-12-21-30 hydrogen bonds (HBs) (black, blue and green dashed line).

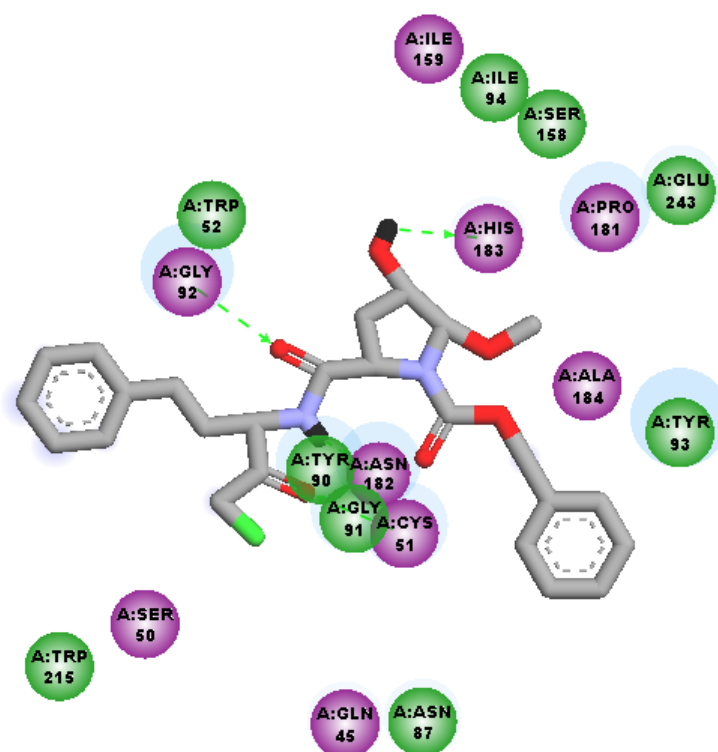


Figure 3.22: 2D interaction diagram of PEP-24-1-12-1-21 hydrogen bonds (HBs) (black, blue and green dashed line).

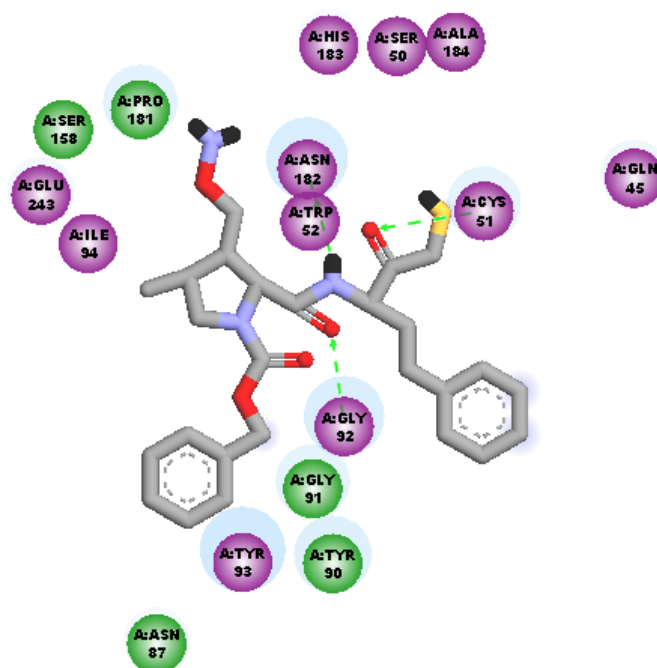


Figure 3.23: 2D interaction diagram of PEP-3-1-12-1-13 hydrogen bonds (HBs) (black, blue and green dashed line).

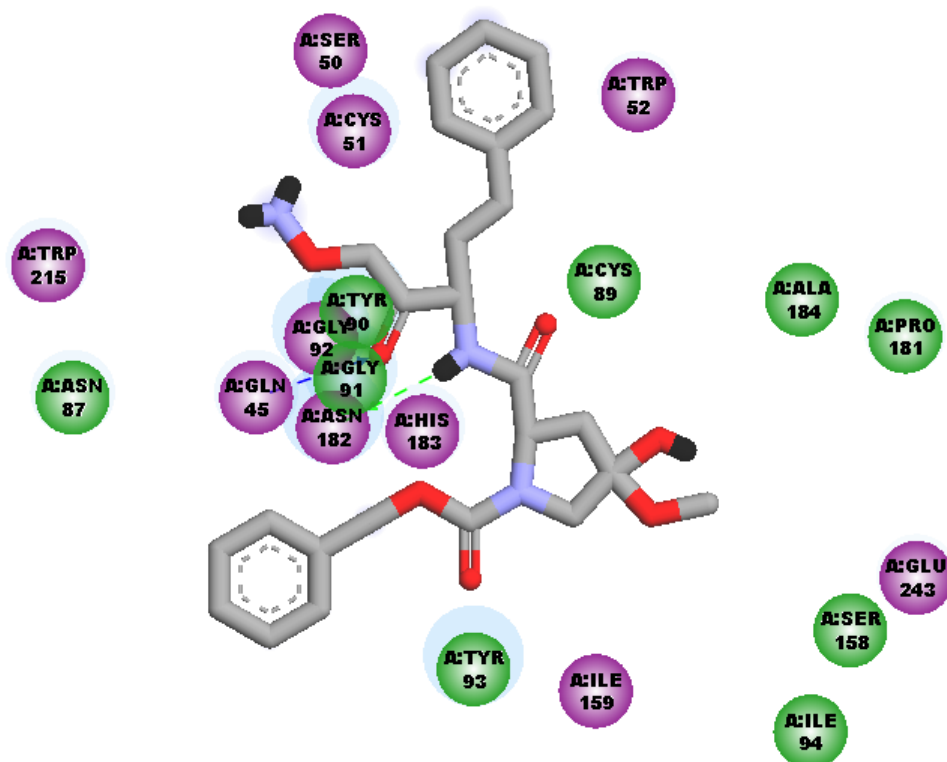


Figure 3.24: 2D interaction diagram of PEP-31-1-12-13-1 hydrogen bonds (HBs) (black, blue and green dashed line).

General Conclusions and Perspectives

In the current work, our focus was to carry out computer-aided design of potential lead compounds for the discovery process against Falcipain-2/3 protein. Thus, firstly, we have attempted to discuss and summarize recent efforts towards the design of malarial cysteine protease inhibitors. It was shown that combining computational techniques with organic synthesis was effective in delivering novel potential inhibitors and providing insight for further improvement of drug candidates. Despite significant efforts from international collaborations and local efforts aimed at controlling the malaria burden, enormous challenges are still being faced by researchers. These include the challenge of developing effective vaccine candidates, together with the drawbacks associated with the use of existing therapies. The emergence and spread of resistance against insecticides, and the available drugs remain another set of challenges. We have attempted to summarize the main finding of several research teams involved in searching for potent cysteine protease FP-2/3 inhibitors against *P. falciparum*.

It appears that the current discussion on efforts towards the design of *P. falciparum* cysteine proteases inhibitors drives us towards combining computational techniques with organic synthesis and using NPs as lead scaffolds for delivering novel potential inhibitors and providing directions for further improvements. Thus, computer-based methods like docking, binding-free energy calculations, virtual screening, and pharmacophore modeling, often used for hit/lead identification along with combinatorial design of novel inhibitors against *P. falciparum* FP-2/3 could be employed. These therefore constitute the available techniques for the development of new drugs against malaria.

Secondly, we have used with success structural information from the crystal structure of FP3-K11017 complex in order to establish a reliable QSAR model of non-covalent inhibition of the FP3 of && by peptidomimetics (PEP) inhibitors. This model correlates the unique descriptor, namely the computed Gibbs free energies (GFE) upon complex formation, with observed inhibitory potencies and is able to identify a few compounds which could be low nM range inhibitors of *P. falciparum*. As GFE is a combined descriptor involving enthalpic gas phase, entropic contributions and solvation free energy, a precise insight into S1' and S2 pockets filling have been performed from the model by analysis of interactions between the enzyme active-site residues and the inhibitor. For this purpose, the breakdown of the interaction energy clearly indicated the residues involved in the affinity with the most active inhibitors. This information has helped to design an initial diversity virtual combinatorial library

of new analogues to be screened by the pharmacophore models derived from the GFE QSAR. The focused library filtered by a set of ADME-related descriptors and screened by matching of the analogues to the PH4 pharmacophore permitted selection of a library subset of orally bioavailable PEP. This subset of 80 best virtual hits was submitted to the computation of predicted *PfFP-3* inhibitory potencies by the formerly established complexation QSAR model. The best analogues reached predicted activities in the low nanomolar concentration range. The most promising dual hits selected for further investigation were, PEP-31-1-12-13-1, PEP-24-1-12-1-21, PEP-2-1-1-16-2, PEP-2-1-12-21-30, PEP-2-1-30-7-2, PEP-2-6-12-21-30, and PEP-3-1-12-1-13, which have respective IC_{50}^{pre} of 0.14, 0.28, 0.004, 0.04, 0.003, 0.63, and 0.96 against *PfFP-3*. These have been proposed for synthesis and biological screening and may lead to a discovery of novel potent orally bioavailable anti-malarial.

In perspectives, we plan to extend the approaches to other diseases provided that the targets are clearly identified.

However, due to the relatively high computational cost of running the software used within this work in terms of CPU time, the realization of the above objectives would require the acquisition of high performance computers and software tools that will allow the development of fast and more efficient methods.

This will probably speed up the drug discovery process against the resistance cases of Malaria and also other neglected tropical diseases currently studied within our research team. In that way, we shall be making our own contribution toward the preclinical drug discovery program as well as in the training of younger scientists engaged in the field of Computer-aided-drug design and discovery.

Bibliography

- [1] World malaria report 2018. Geneva: World Health Organization; 2018. https://www.who.int/malaria/publications/world_malaria_report/en/. Accessed 27 Nov 2019.
- [2] Nwaka S, Hudson A. Innovative lead discovery strategies for tropical diseases. *Nature Reviews Drug Discovery*, **2006**, 5, 941-955.
- [3] Walsh CT. Antibiotics: actions, origins, resistance. *American Society for Microbiology (ASM)*, **2003**.
- [4] Baltz RH. Antibiotic discovery from actinomycetes: will a renaissance follow the decline and fall?. *Sim News*, **2005**, 55, 186-196.
- [5] Baltz RH. Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *Journal of Industrial Microbiology and Biotechnology*, **2006**, 33, 507-513.
- [6] Drlica K. Antibiotic resistance: can we beat the bugs? *Drug Discovery Today*, **2001**, 6, 714-715.
- [7] Böhm HJ, Schneider G. Computer software reviews. *Journal of the American Chemical Society*, **2001**, 123 (29), 7197-7198.
- [8] Lyne PD. Structure-based virtual screening: an overview. *Drug Discovery Today*, **2002**, 7, 1047-1055.
- [9] Pink R, Hudson A, Mouriès MA, Bendig M. Opportunities and challenges in antiparasitic drug discovery. *Nature Reviews Drug discovery*, **2005**, 4, 727-740.
- [10] Aitken SG. Drug discovery in the new millennium. *Chemistry in New Zealand*, **2006**, 88-92.
- [11] Qi D, King RD, Hopkins AL, Bickerton GRJ, Soldatova LN. An ontology for description of drug discovery investigations. *Journal of Integrative Bioinformatics*, **2010**, 7, 126-138.

- [12] Trouiller P, Olliaro P, Torreele E, Orbinski J, Laing R, Ford N. Drug development for neglected diseases: a deficient market and a public policy failure. *The Lancet*, **2002**, 359(9324), 2188-2194.
- [13] Horton J. Drug development for tropical diseases - present situation, future perspectives. *Trends in Parasitology*, **2003**, 19, 1-6.
- [14] Widdus R, White K. Combating diseases associated with poverty: financing strategies for product development and the potential role of public-private partnerships. *Initiative on Public-Private Partnerships for Health*. **2004**, 15-16 April.
- [15] Nwaka S, Ramirez B, Brun R, Maes L, Douglas F, Ridley R. Advancing drug innovation for neglected diseases-criteria for lead progression. *PLoS neglected tropical diseases*, **2009**, 3(8), e440.
- [16] Nwaka S, Ridley RG. Virtual drug discovery and development for neglected diseases through public-private partnerships. *Nature Reviews Drug Discovery*, **2003**, 2(11), 919-928.
- [17] Terstappen GC, Reggiani A. *In silico* research in drug discovery. *Trends in Pharmacological Sciences*, **2003**, 22(1), 23-26.
- [18] Joseph-McCarthy D. Computational approaches to structure-based ligand design. *Pharmacology & Therapeutics*, **1999**, 84(2), 179-191.
- [19] Kantardjieff K, Rupp B. Structural bioinformatic approaches to the discovery of new antimycobacterial drugs. *Current Pharmaceutical Design*, **2004**, 10(26), 3195-3211.
- [20] Xu J, Hagler A. Cheminformatics and drug discovery. *Molecules*, **2002**, 7(8), 566-600.
- [21] DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. *Journal of Health Economics*, **2003**, 22(2), 151-185.
- [22] Rosenthal PJ, Sijwali PS, Singh A, Shenai BR. Cysteine proteases of malaria parasites: targets for chemotherapy. *Current Pharmaceutical Design*, **2002**, 8(18), 1659-1672.
- [23] Yeh I, Altman RB. Drug targets for *Plasmodium falciparum*: a post-genomic review/survey. *Mini Reviews in Medicinal Chemistry*, **2006**, 6(2), 177-202.
- [24] White SW, Zheng J, Zhang YM, Rock CO. The structural biology of type II fatty acid biosynthesis. *Annual Reviews in Biochemistry*, **2005**, 74, 791-831.
- [25] Zhang YM, Lu YJ, Rock CO. The reductase steps of the type II fatty acid synthase as antimicrobial targets. *Lipids*, **2005**, 39(11), 1055-1060.
- [26] Walker NPC, Talanian RV, Brady KD, Dang LC, Bump NJ, Ferenza CR, Franklin S, Ghayur T, Hackett MC, Hammill LD, Herzog L, Hugunin M, Houy W, Mankovich JA, McGuinness L, Orlewicz E, Paskind M, Pratt CA, Reis P, Summani A, Terranova M, Welch JP, Xiong

- L, Möller A, Tracey DE, Kamen R, Wong WW. Crystal structure of the cysteine protease interleukin-1 β -converting enzyme: a (p20/p10)₂ homodimer. *Cell*, **1994**, 78(2), 343-352.
- [27] Thornberry NA. The caspase family of cysteine proteases. *British medical bulletin*, **1996**, 53(3), 478-490.
- [28] Frecer V, Jedinak A, Tossi A, Berti F, Benedetti F, Romeo D, Miertus S. Structure based design of inhibitors of aspartic protease of HIV-1. *Letters in drug design and discovery*, **2005**, 2(8), 638- 646.
- [29] Frecer V, Berti F, Benedetti F, Miertus S. Design of peptidomimetic inhibitors of aspartic protease of HIV-1 containing–Phe Ψ Pro–core and displaying favourable ADME-related properties. *Journal of Molecular Graphics and Modelling*, **2008**, 27(3), 376-387.
- [30] Kerr ID, Lee JH, Farady CJ, Marion R, Rickert M, Sajid M, Pandey KC, Caffrey CR, Legac J, Hansell E, McKerrow JH, Craik CS, Rosenthal PJ, Brinen LS. Vinyl sulfones as antiparasitic agents: a structural basis for drug design. *Journal of Biological Chemistry*, **2009**, 284, 25697-25703.
- [31] Laveran A. Note sur un nouveau parasite trouvé dans le sang de plusieurs malades atteints de fièvre palustre. *Bulletin de l'Académie Nationale de Médecine*, **1888**, 9, 1234-1235.
- [32] Ettari R, Bova F, Zappala M, Grasso S, Micale N. Falcipain-2 inhibitors. *Medicinal Research Reviews*, **2010**, 30, 136-167.
- [33] Faraz Mojab. Antimalarial natural products: a review. *Avicenna Journal of Phytomedicine*, **2012**, 2, 52-62.
- [34] Objectifs du Millénaire pour le développement (<http://www.un.org/fr/millenniumgoals/>).
- [35] WHO. World malaria report 2015. Geneva: World Health Organization; **2015**. https://www.who.int/malaria/publications/world_malaria_report/en/. Accessed December 2015.
- [36] WHO. World malaria report 2016. Geneva: World Health Organization; **2016**. https://www.who.int/malaria/publications/world_malaria_report/en/. Accessed December 2016.
- [37] MacNeil A, Rollin PE. Ebola and Marburg hemorrhagic fevers: neglected tropical diseases. *PLoS Neglected Tropical Diseases*, **2012**, 6, e1546.
- [38] Utzinger J. A research and development agenda for the control and elimination of human helminthiases. *PLoS Neglected Tropical Diseases*, **2012**, 6, e1646.
- [39] Nethavhani SA, van Ree T. Synthesis and antimalarial activities of some novel 2-pyridones. *Arabian Journal for Science and Engineering*, **2014**, 39(9), 6595-6598.

- [40] Kesharwani RK, Singh DV, Misra K. Computation-based virtual screening for designing novel antimalarial drugs by targeting falcipain-III: A structure-based drug designing approach. *Journal of Vector Borne Diseases*, **2013**, 50(2), 93-102.
- [41] Rosenthal PJ. Proteases of malaria parasites: new targets for chemotherapy. *Emerging Infectious Diseases*, **1998**, 4(1), 49-57.
- [42] Ramjee MK, Flinn NS, Pemberton TP, Quibell M, Wang Y, Watts JP. Substrate mapping and inhibitor profiling of falcipain-2, falcipain-3 and berghepain-2: implications for peptidase anti-malarial drug discovery. *Biochemical Journal*, **2006**, 399(1), 47-57.
- [43] Potshangbam AM, Tanneeru K, Reddy BM, Guruprasad L. 3D-QSAR and molecular docking studies of 2-pyrimidinecarbonitrile derivatives as inhibitors against falcipain-3. *Bioorganic and Medicinal Chemistry Letters*, **2011**, 2, 7219-7223.
- [44] Kamkumo RG, Ngoutane AM, Tchokouaha LRY, Fokou PVT, Madiesse EAK, Legac J, Kezetas JJB, Lenta BN, Boyom FF, Dimo T, Mbacham WF, Gut J, Rosenthal PJ. Compounds from *Sorindeia juglandifolia* (Anacardiaceae) exhibit potent anti-plasmodial activities *in vitro* and *in vivo*. *Malaria Journal*, 2012, 11(1), 382.
- [45] World malaria report 2008. Geneva: World Health Organization; **2008**. https://www.who.int/malaria/publications/world_malaria_report/en/. Accessed December 2008.
- [46] Tadigoppula N, Korthikunta V, Gupta S, Kancharla P, Khaliq T, Soni A, Srivastava RK, Srivastava K, Puri SK, Raju KSR, Sijwal PS, Kuma V, Mohammad IS. Synthesis and insight into the structure-activity relationships of chalcones as antimalarial agents. *Journal of Medicinal Chemistry*, **2013**, 56(1), 31-45.
- [47] World malaria report 2017. Geneva: World Health Organization; **2017**. https://www.who.int/malaria/publications/world_malaria_report/en/. Accessed December 2017.
- [48] Cox-Singh J, Davis TM, Lee KS, Shamsul SS, Matusop A, Ratnam S, Rahman HA, Conway DJ, Singh B. *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life threatening. *Clinical Infectious Diseases*, **2008**, 46(2), 165-171.
- [49] Singh B, Sung LK, Matusop A, Radhakrishnan A, Shamsul SS, Cox-Singh J, Thomas A, JConway D. A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *The Lancet*, **2004**, 363(9414), 1017-1024.
- [50] Getachew S, To S, Trimarsanto H, Thriemer K, Clark TG, Petros B, Aseffa A, Price RN, Sarah Auburn. Variation in complexity of infection and transmission stability between neighbouring populations of *Plasmodium vivax* in Southern Ethiopia. *PLoS one*, **2015**, 10(10), e0140780.

- [51] Barnadas C, Timinao L, Javati S, Iga J, Malau E, Koepfli C, Robinson LJ, Senn N, Kinitoboro B, Rare L, Reeder JC, Siba PM, Zimmerman PA, Karunajeewa H, Davis TM, Mueller I. Significant geographical differences in prevalence of mutations associated with *Plasmodium falciparum* and *Plasmodium vivax* drug resistance in two regions from Papua New Guinea. *Malaria Journal*, **2015**, *14*(1), 399.
- [52] Marfurt J, de Monbrison F, Brega S, Barbolat L, Müller I, Sie A, Goroti M, Reeder JC, Beck HP, Picot S, Genton B. Molecular markers of *in vivo Plasmodium vivax* resistance to amodiaquine plus sulfadoxine-pyrimethamine: mutations in *pvdhfr* and *pvmr1*. *The Journal of Infectious Diseases*, **2008**, *198*(3), 409-417.
- [53] Shaw SK, Thomson-Luque R, Obaldía N, Nuñez M, Dutary S, Lim C, Barnes S, Kocken CHM, Duraisingh MT, Adams JH, Pasini EM. Insights into an optimization of *Plasmodium vivax* Sal-1 *in vitro* culture: the Aotus primate model. *PLoS neglected tropical diseases*, **2016**, *10*(7), e0004870.
- [54] Singh B, Daneshvar C. *Plasmodium knowlesi* malaria in Malaysia. *Medical Journal of Malaysia*, **2010**, *65*(3), 166-172.
- [55] Cornejo OE, Escalante AA. The origin and age of *Plasmodium vivax*. *Trends in Parasitology*, **2006**, *22*(12), 558-563.
- [56] Kibret S, Lautze J, McCartney M, Nhamo L, Wilson GG. Malaria and large dams in sub-Saharan Africa: future impacts in a changing climate. *Malaria Journal*, **2016**, *15*(1), P. 448.
- [57] Hoffman SL, Subramanian GM, Collins FH, Venter JC. *Plasmodium*, human and anopheles genomics and malaria. *Nature*, **2002**, *415*(6872), 702-709.
- [58] Enayati A, Hemingway J. Malaria management: past, present, and future. *Annual review of entomology*, **2010**, *55*, 569-591.
- [59] Kilama, Ntoumi F. Malaria: a research agenda for the eradication era. *The Lancet*, **2009**, *374*(9700), 1480-1482.
- [60] Rogier C, Orlandi-Pradines E, Fusai T, Pradines B, Briolant S, Almeras. Malaria vaccines: prospects and reality. *Medecine et Maladies Infectieuses*, **2006**, *36*(8), 414-422.
- [61] Djogbenou L. Vector control methods against malaria and vector resistance to insecticides in Africa. *Medecine Topicale*, **2009**, *69*(2), 160-164.
- [62] Pages F. Vector control for armed forces: a historical requirement requiring continual adaptation. *Medecine Topicale*, **2009**, *69*(2), 165-172.
- [63] <https://www.merriam-webster.com/dictionary/drug>.

- [64] WHO. World malaria report 2003. Geneva: World Health Organization; **2003**. https://www.who.int/malaria/publications/world_malaria_report/en/. Accessed December 2003.
- [65] Mott BT, Tripathi A, Siegler MA, Moore CD, Sullivan DJ. Synthesis and antimalarial efficacy of two carbon-linked, artemisinin-derived trioxane dimers in combination with known antimalarial drugs. *Journal of Medicinal Chemistry*, **2013**, 56(6), 2630-2641.
- [66] Uhlemann AC, Cameron A, Eckstein LU, Fischbarg J, Iserovich P, Zuniga FA, East M, Lee A, Brady L, Haynes RK, Sanjeev Krishna. A single amino acid residue can determine the sensitivity of SERCAs to artemisinins. *Nature structural and molecular biology*, **2005**, 12(7), 628-629.
- [67] Dali B, Keita M, Megnassan E, Frecer V, Miertus S. Insight into selectivity of peptidomimetic inhibitors with modified statine core for plasmepsin II of *Plasmodium falciparum* over human Cathepsin D. *Chemical biology and drug design*, **2012**, 79(4), 411-30.
- [68] Ashley EA, Dhorda M, Fairhurst R M, Amaratunga C, Lim P, Suon S, Sreng MDS, Anderson JM, Mao S, Sam B, Sopha C, Chuor CM, Nguon C, Sovannaroeth S, Pukrittayakamee S, Jittamala P, Chotivanich K, Chutasmit K, Suchatsoonthorn C, Runcharoen R, Hien TT, Thuy-Nhien NT, Thanh NV, Phu NH, Htut Y, Han KT, Aye KH, Mokuolu OA, Olaosebikan RR, Folaranmi OO, Mayxay M, Khanthavong M, Hongvanthong B, Newton PN, Onyamboko MA, Fanello CI, Tshefu AK, Mishra N, Valecha N, Phyo AP, Nosten R, Yi P, Tripura R, Borrmann S, Bashraheil M, Peshu J, Faiz A, Ghose A, Hossain A, Samad R, Rahman R, Hasan M, Islam A, Miotto O, Amato R, MacInnis B, Stalker J, Kwiatkowski DP, Bozdech Z, Jeeyapant A, Cheah PY, Sakulthaew T, Chalk J, Intharabut B, Silamut K, Lee SJ, Vihokhern B, Kunaso Cl, Imwong M, Tarning J, Taylor WJ, Yeung S, Woodrow CJ, Flegg JA, Das D, Smith J, Venkatesan M, Plowe CV, Stepniewska K, Guerin PJ, Dondorp AM, Day NP, White NJ. Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *New England Journal of Medicine*, **2014**, 371(5), 411-423.
- [69] Mishra PS, Sharma H, Mishra R, Gupta SK. A review on antimalarial drug discovery and its screening method. *World Journal of Pharmacy and Pharmaceutical Sciences*, **2014**, 8, 1288-1304.
- [70] Phillipson JD, O'Neill MJ. Antimalarial and amoebicidal natural products. *Biologically active natural products*. Edited by K. Hostettmann and PJ Lea, **1987**.
- [71] Reilly HB, Wang H, Steuter JA, Marx AM, Ferdig MT. Quantitative dissection of clone-specific growth rates in cultured malaria parasites. *International Journal of Parasitology*, **2007**, 37(14), 1599-1607.
- [72] Wongsrichanalai C, Steven RM. Declining artesunate-mefloquine efficacy against falciparum malaria on the Cambodia-Thailand border. *Emerging Infectious Diseases*, **2008**, 14(5), 716-719.

- [73] Pandey KC, Wang SX, Sijwali PS, Lau AL, McKerrow JH, Rosenthal PJ. The *Plasmodium falciparum* cysteine protease falcipain-2 captures its substrate, hemoglobin, via a unique motif. *Proceedings of the National Academy of Sciences*, **2011**, *102*(26), 9138-9143.
- [74] Thomas MB, Godfray HCJ, Read AF, van den Berg H, Tabashnik BE, van Lenteren JC, Waage JK, Takken W. Lessons from agriculture for the sustainable management of malaria vectors. *PLoS Medicine*, **2012**, *9*(7), e1001262.
- [75] Chanda E, Hemingway J, Kleinschmidt I, Rehman AM, Ramdeen V, Phiri FN, Coetzer S, Mthembu D, Shinondo CJ, Chizema-Kawesha E, Kamuliwo M, Mukonka V, Baboo KS, Colema M. Insecticide resistance and the future of malaria control in Zambia. *PLoS One*, **2011**, *6*(9), e24336.
- [76] Lubell Y, Dondorp A, Guérin PJ, Drake T, Meek S, Ashley E, Day NPJ, White NJ, White LJ. Artemisinin resistance- modelling the potential human and economic costs. *Malaria Journal*, **2014**, *13*(1), 1-10.
- [77] Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Perteua M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DMA, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, **2002**, *419*(6906), 498-511.
- [78] Rosenberg R. Malaria: some considerations regarding parasite productivity. *Trends Parasitology*, **2008**, *24*(11), 487-491.
- [79] Kappe SHI, Duffy PE. Malaria liver stage culture: *in vitro* veritas? *The American journal of tropical medicine and hygiene*, **2006**, *74*(5), 706-707.
- [80] Kuehn A, Pradel G. The coming-out of malaria gametocytes. *Journal of Biomedicine and Biotechnology*, **2010**, *2010*, ID 976827.
- [81] Stone WJR, Eldering M, Gemert GJ, Lanke KH, Grignard L, van de Vegte-Bolmer MG, Siebelink-Stoter R, Graumans W, Roeffen WFG, Drakeley CJ, Sauerwein RW, Bousema T. The relevance and applicability of oocyst prevalence as a read-out for mosquito feeding assays. *Scientific Reports*, **2013**, *3*, 3418.
- [82] Kappe SH, Gardner MJ, Brown SM, Ross J, Matuschewski K, Ribeiro JM, Adams JH, Quackenbush J, Cho J, Carucci DJ, Hoffman SL, Nussenzweig V. Exploring the transcriptome of the malaria sporozoite stage. *Proceedings of the National Academy of Sciences*, **2001**, *98*(17), 9895-9900.

- [83] Reilly HB, Wang H, Steuter JA, Marx AM, Ferdig MT. Quantitative dissection of clone-specific growth rates in cultured malaria parasites. *International Journal for Parasitology*, **2007**, 37(14), 1599-1607.
- [84] Dhangadamajhi G, Kar SK, Ranjit M. The survival strategies of malaria parasite in the red blood cell and host cell polymorphisms. *Malaria Research and Treatment*, **2010**.
- [85] Ranford-Cartwright LC, Mwangi JM. Analysis of malaria parasite phenotypes using experimental genetic crosses of *Plasmodium falciparum*. *International Journal for Parasitology*, **2012**, 42, 529-34.
- [86] Teixeira C, Gomes JRB, Couesnon T, Gomes P. Molecular docking and 3D-quantitative structure activity relationship analyses of peptidyl vinyl sulfones: *Plasmodium falciparum* cysteine proteases inhibitors. *Journal of Computer-aided Molecular Design*, **2011**, 25, 763-775.
- [87] Coombs GH, Goldberg DE, Klemba M, Berry C, Kay J, Mottram JC. Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends in Parasitology*, 2001, 17, 532-537.
- [88] Klein EY. Antimalarial drug resistance: a review of the biology and strategies to delay emergence and spread. *International Journal of Antimicrobial Agents*, **2013**, 41, 311-317.
- [89] Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, Moch JK, Muster N, Sacci JB, Tabb DL, Witney AA, Wolters D, Wu Y, Gardner MJ, Holder AA, Sinden RE, Yates JR, Carucci DJ. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature*, **2002**, 419(6906), 520-526.
- [90] Krungkrai SR, Krungkrai J. Insights into the pyrimidine biosynthetic pathway of human malaria parasite *Plasmodium falciparum* as chemotherapeutic target. *Asian Pacific Journal of Tropical Medicine*, **2016**, 9(6), 525-534.
- [91] Sullivan DJ, Gluzman IY, Russell DG, Goldberg DE. On the molecular mechanism of chloroquine's antimalarial action. *Proceedings of the National Academy of Sciences*, **1996**, 93(21), 11865-11870.
- [92] Bray PG, Martin RE, Tilley L, Ward SA, Kirk K, Fidock DA. Defining the role of PfCRT in *Plasmodium falciparum* chloroquine resistance. *Molecular Microbiology*, **2005**, 56(2), 323-333.
- [93] Egan TJ. Haemozoin formation. *Molecular and Biochemical Parasitology*, **2008**, 157(2), 127-136.
- [94] Pisciotta JM, Sullivan D. Hemozoin: oil versus water. *Parasitology International*, **2008**, 57(2), 89-96.

- [95] Dogovski C, Xie SC, Burgio G, Bridgford J, Mok S, McCaw JM, Chotivanich K, Kenny S, Gnädig N, Straimer J, Bozdech Z, Fidock DA, Simpson JA, Tilley L. Targeting the cell stress response of *Plasmodium falciparum* to overcome artemisinin resistance. *PLoS Biology*, **2015**, *13*(4), e1002132.
- [96] O'Neill PM, Victoria EB, Stephen AW. The molecular mechanism of action of artemisinin-the debate continues. *Molecules*, **2010**, *15*(3), 1705-1721.
- [97] Woodrow CJ, Haynes RK, Krishna S. Artemisinins. *Postgraduate Medical Journal*, **2005**, *81*(952), 71-78.
- [98] Drug Resistance at the US National Library of Medicine Medical Subject Headings (MeSH), **2017**.
- [99] Jovel IT, Mejia RE, Banegas E, Piedade R, Alger J, Fontecha G, Ferreira PE, Veiga MI, Enamorado IG, Bjorkman A, Ursing J. Drug resistance associated genetic polymorphisms in *Plasmodium falciparum* and *Plasmodium vivax* collected in Honduras, Central America. *Malaria Journal*, **2011**, *10*(1), p.376.
- [100] Basak SC. Chemobioinformatics: the advancing frontier of computer-aided drug design in the post-genomic era. *Current computer-aided drug design*, **2012**, *8*(1), 1-2.
- [101] Song CM, Lim SJ, Tong JC. Recent advances in computer-aided drug design. *Briefings in Bioinformatics*, **2009**, *10*(5), 579-591.
- [102] Flannery EL, Chatterjee AK, Winzeler EA. Antimalarial drug discovery: approaches and progress towards new medicines. *Nature Reviews Microbiology*, **2013**, *11*(12), 849-862.
- [103] Raphemot R, Posfai D, Derbyshire ER. Current therapies and future possibilities for drug development against liver-stage malaria. *The Journal of Clinical Investigation*, **2016**, *126*(6), 2013-2020.
- [104] World Health Organization. The Global Plan to Stop TB 2011-2015: Transforming the Fight Towards the Elimination of Tuberculosis; WHO: Geneva, Switzerland, **2010**, p. 7.
- [105] Aitken SG. Drug discovery in the new millennium. *Chemistry in New Zealand*, **2006**, 88-92.
- [106] Verlinde CLMJ, Hol WGJ. Structure-based drug design: progress, results and challenges. *Structure*, **1994**, *2*(7), 577-587.
- [107] Schneider G, Baringhaus KH, Molecular Design: Concepts and Applications. John Wiley & Sons, **2008**, pp. 149-160.
- [108] Ntie-Kang F. Computer-aided design and virtual screening of novel inhibitors of pantothenate synthetase, enoyl-acyl carrier protein reductase and multidrug-resistant strains of mycobacterium tuberculosis. Thèse de Doctorat/PhD, CEPAMOQ, Université de Douala, **2014**.

- [109] Macalino SJY, Gosu V, Hong S, Choi S. Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, **2015**, *38*(9), 1686-1701.
- [110] Lavecchia A, Di Giovanni C. Virtual screening strategies in drug discovery: a critical review. *Current Medicinal Chemistry*, **2013**, *20*(23), 2839-2860.
- [111] Grinter SZ, Zou X. Challenges, applications, and recent advances of protein-ligand docking in structure-based drug design. *Molecules*, **2014**, *19*(7), 10150-10176.
- [112] Beddell CR, Goodford PJ, Norrington FE, Wilkinson S, Wootton R. Compounds designed to fit a site of known structure in human haemoglobin. *British Journal of Pharmacology*, **1976**, *57*(2), 201-209.
- [113] Seddon G, Lounnas V, McGuire R, van den Bergh T, Bywater RP, Oliveira L, Vriend G. Drug design for ever, from hype to hope. *Journal of Computer-Aided Molecular Design*, **2012**, *26*(1), 137-150.
- [114] Von Itzstein M, Dyason JC, Oliver SW, White HF, Wu WY, Kok GB, Pegg MS. A study of the active site of influenza virus sialidase: an approach to the rational design of novel anti-influenza drugs. *Journal of Medicinal Chemistry*, **1996**, *39*(2), 388-391.
- [115] Kaldor SW, Kalish VJ, Davies JF, Shetty BV, Fritz JE, Appelt K, Burgess JA, Campanale KM, Chirgadze NY, Clawson DK, Dressman BA, Hatch SD, Khalil DA, Kosa MB, Lubbehusen PP, Muesing MA, Patick AK, Reich SH, Su KS, Tatlock JH. Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. *Journal of Medicinal Chemistry*, **1997**, *40*(24), 3979-3985.
- [116] Blackledge G. New developments in cancer treatment with the novel thymidylate synthase inhibitor raltitrexed ('Tomudex'). *British Journal of Cancer*, **1998**, *77*(S2), 29-37.
- [117] Adkins JC, Faulds D. Amprenavir. *Drugs*, **1998**, *55*(6), 837-842.
- [118] Schames JR, Henchman RH, Siegel JS, Sottriffer CA, Ni H, McCammon JA. Discovery of a novel binding trench in HIV integrase. *Journal of Medicinal Chemistry*, **2004**, *47*(8), 1879-1881.
- [119] Prathipati PAD, Saxena AK. Computer-aided drug design: Integration of structure-based and ligand-based approaches in drug design. *Current Computer-Aided Drug Design*, **2007**, *3*, 133-148.
- [120] Awanish K, Anubhuti J. Drug development strategies. *Anticandidal Agents*, **2007**, 63-71.
- [121] Basso LA, Silva LHPD, Fett-Neto AG, Azevedo JWFD, Moreira ÍDS, Palma MS, Calixto JB, Filho SA, dos Santos RR, Soares MPB, Santos DS. The use of biodiversity as source of new chemical entities against defined molecular targets for treatment of malaria, tuberculosis, and T-cell mediated diseases: a review. *Memórias do Instituto Oswaldo Cruz*, **2005**, *100*(6), 475-506.

- [122] Padmanaban G, Nagaraj VA, Rangarajan PN. Drugs and drug targets against malaria. *Current Science*, **2007**, *92(11)*, 1545-1555.
- [123] Michael EAD, McIntosh T, Dean H, Chen S, Zhang G, Baevova P, Joiner KA. Four distinct pathways of hemoglobin uptake in the malaria parasite *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences*, **2008**, *105(7)*, 2463-2468.
- [124] Miller LH, Su X. Artemisinin: discovery from the Chinese herbal garden. *Cell*, **2011**, *146(6)*, 855-863.
- [125] Liu J, Gluzman IY, Drew ME, Goldberg DE. The role of *Plasmodium falciparum* food vacuole plasmepsins. *Journal of Biological Chemistry*, **2005**, *280(2)*, 1432-1437.
- [126] Blackman MJ. Proteases involved in erythrocyte invasion by the malaria parasite: function and potential as chemotherapeutic targets. *Current Drug Targets*, **2000**, *1(1)*, 59-83.
- [127] Eggleston KK, Duffin KL, Goldberg DE. Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*. *Journal of Biological Chemistry*, **1999**, *274(45)*, 32411-32417.
- [128] Klemba M, Gluzman I, Goldberg DE. A *Plasmodium falciparum* dipeptidylaminopeptidase I participates in vacuolar hemoglobin degradation. *Journal of Biological Chemistry*, **2004**, *279(41)*, 43000-43007.
- [129] Teixeira C, Gomes JRB, Gomes P. Falcipains, *Plasmodium falciparum* cysteine proteases as key drug targets against malaria. *Current Medicinal Chemistry*, **2011**, *18(10)*, 1555-1572.
- [130] Padmanaban G, Nagaraj VA, Rangarajan PN. Drugs and drug targets against malaria. *Current Science*, **2007**, *92(11)*, 1545-1555.
- [131] Kamphuis R, Kalk KH, Swarte MBA, Drenth J. Structure of papain refined at 1.65 Å resolution. *Journal of Molecular Biology*, **1984**, *179(2)*, 233-256.
- [132] Garavito RM, Rossmann MG, Argos P, Eventoff W. Convergence of active center geometries. *Biochemistry*, **1977**, *16(23)*, 5065-5071.
- [133] Malgorzata R, Dorota C, Stec-Niemczyk J. Modes of inhibition of cysteine proteases. *Acta Biochemical Polonica*, **2004**, *51*, 861-873.
- [134] Coterón JM, Catterick D, Castro J, Chaparro MJ, Díaz B, Fernández E, Ferrer S, Gamo FJ, Gordo M, Gut J, de las Heras L, Legac J, Marco M, Miguel J, Muñoz V, Porras E, de la Rosa JC, Ruiz JR, Sandoval E, Ventosa P, Rosenthal PJ, Jose M. Fiandor JM. Falcipain inhibitors: optimization studies of the 2-pyrimidinecarbonitrile lead series. *Journal of Medicinal Chemistry*, **2010**, *53(16)*, 6129-6152.
- [135] Pandey KC, Dixit R. Structure-function of falcipains: malarial cysteine proteases. *Journal of Tropical Medicine*, **2012**, *2012*, ID 345195.

- [136] Kerr ID, Lee JH, Pandey KC, Harrison A, Sajid M, Rosenthal PJ, Brinen LS. Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. *Journal of Medicinal Chemistry*, **2009**, *52*(3), 852-857.
- [137] Shenai BR, Sijwali PS, Singh A, Rosenthal PJ. Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*. *Journal of Biological Chemistry*, **2000**, *275*(37), 29000-29010.
- [138] Sijwali P S, Rosenthal PJ. Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences*, **2004**, *101*(13), 4384-9438.
- [139] Sijwali PS, Koo J, Singh N, Rosenthal PJ. Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*, **2006**, *150*(1), 96-106.
- [140] Powers JC, Asgian JL, Ekici OD, James KE. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chemical Reviews*, **2002**, *102*(12), 4639-4750.
- [141] Vicik R, Busemann M, Baumann K, Schirmeister T. Inhibitors of cysteine proteases. *Current Topics in Medicinal Chemistry*, **2006**, *6*(4), 331-353.
- [142] Lee BJ, Singh A, Chiang P, Kemp SJ, Goldman EA, Weinhouse MI, Vlasuk GP, Rosenthal PJ. Antimalarial activities of novel synthetic cysteine protease inhibitors. *Antimicrobial Agents and Chemotherapy*, **2003**, *47*(12), 3810-3814.
- [143] Schulz F, Gelhaus C, Degel B, Vicik R, Heppner S, Breuning A, Leippe M, Gut J, Rosenthal PJ, Schirmeister T. Screening of protease inhibitors as antiplasmodial agents. Part I: Aziridines and epoxides. *ChemMedChem*, **2007**, *2*(8), 1214-1224.
- [144] Martichonok V, Plouffe C, Storer AC, Menard R, Jones JB. Aziridine analogues of [[trans-(epoxysuccinyl)-L-leucyl]amino]-4-guanidinobutane (E-64) as inhibitors of cysteine proteases. *Journal of Medicinal Chemistry*, **1995**, *38*(16), 3078-3085.
- [145] Bekono BD, Ntie-Kang F, Owono LCO, Megnassan E. Targeting cysteine proteases from *Plasmodium falciparum*: a general overview, rational drug design and computational approaches for drug discovery. *Current Drug Targets*, **2018**, *19*(5), 501-526.
- [146] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. *Nucleic Acids Research*, **2000**, *28*(1), 235-242.
- [147] Wang SX, Pandey KC, Somoza JR, Sijwali PS, Kortemme T, Brinen LS, Fletterick RJ, Rosenthal PJ, McKerrow JH. Structural basis for unique mechanisms of folding and hemoglobin binding by a malarial protease. *Proceedings of the National Academy of Sciences*, **2006**, *103*(31), 11503-11508.

- [148] Hogg T, Nagarajan K, Herzberg S, Chen L, Shen X, Jiang H, Wecke M, Blohmke C, Hilgenfeld R, Schmidt CL. Structural and Functional Characterization of Falcipain-2, a Hemoglobinase from the Malarial Parasite *Plasmodium falciparum*. *Journal of Biological Chemistry*, **2006**, *281(35)*, 25425-25437.
- [149] Wang SX, Pandey KC, Scharfstein J, Whisstock J, Huang RK, Jacobelli J, Fletterick RJ, Rosenthal PJ, Abrahamson M, Brinen LS, Rossi A, Sali A, McKerrow JH. The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family. *Structure*, **2007**, *15(5)*, 535-543.
- [150] Kerr ID, Lee JH, Pandey KC, Harrison A, Sajid M, Rosenthal PJ, Brinen LS. Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. *Journal of Medicinal Chemistry*, **2009**, *52(3)*, 852-857.
- [151] Hansen G, Heitmann A, Witt T, Li H, Jiang H, Shen X, Heussler VT, Rennenberg A, Hilgenfeld R. Structural basis for the regulation of cysteine-protease activity by a new class of protease inhibitors in *Plasmodium*. *Structure*, **2011**, *19(7)*, 919-929.
- [152] Pissurlenkar RR, Shaikh MS, Iyer RP, Coutinho EC. Molecular mechanics force fields and their applications in drug design. *Anti-Infective Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Infective Agents)*, **2009**, *8(2)*, 128-150.
- [153] Debord J. Introduction à la modélisation moléculaire. **2004**, 37-41.
- [154] Foresman JB, Frisch A. Exploring chemistry with electronic structure methods second edition. *Gaussian Inc Pittsburg PA*, **1995**, 254-285.
- [155] Hartree DR. The wave mechanics of an atom with a non-Coulomb central field. Part I theory and methods. *Mathematical Proceedings of the Cambridge Philosophical Society*, **1928**, *24(1)*, 89-110.
- [156] Born M, Oppenheimer JR. On the quantum theory of molecules. *Annalen der Physik*, **1927**, *84*, 457-484.
- [157] Dalgaard E, Jorgensen P. Optimization of orbitals for multi-configurational reference states. *The Journal of Chemical Physics*, **1978**, *69(8)*, 3833-3844.
- [158] Bethe HA, Jackiw R. Intermediate quantum mechanics. *Lecture notes in Physics. Benjamin, New York*, **1958**.
- [159] Andrews DH. The relation between the raman spectra and the structure of organic molecules. *Physical Review*, **1930**, *36(3)*, p.544.
- [160] Lifson S, Warshel A. Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and n-alkane molecules. *The Journal of Chemical Physics*, **1968**, *49(11)*, 5116-5129.

- [161] Bartol P, Comba M, Zimmer M. Conformational searching of transition metal compounds. *Journal of Computational Chemistry*, **1999**, 20(14), 1549-1558.
- [162] Josephson BD, Pallikari-Viras F. Biological utilization of quantum nonlocality. *Foundations of Physics*, 1991, 21(2), 197-207.
- [163] Pearlman DA, Kollman PA. Evaluating the assumptions underlying force field development and application using free energy conformational maps for nucleosides. *Journal of the American Chemical Society*, **1991**, 113(19), 7167-7177.
- [164] Dillen JLM. PEF: a program for the development of empirical force field. *Journal of computational chemistry*, **1992**, 13(3), 257-267.
- [165] Momany FA, McGuire RF, Burgess AW, Sheraga HA. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *The Journal of Physical Chemistry*, **1975**, 79(22), 2361-2381.
- [166] Allinger NL. Conformational Analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *Journal of the American Chemical Society*, **1977**, 99 (25), 8127-8134.
- [167] Ermer O, Lifson S. Consistent force field calculations. III. Vibrations, conformations, and heats of hydrogenation of nonconjugated olefins. *Journal of the American Chemical Society*, **1973**, 95(13), 4121-4132.
- [168] Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, **2004**, 25(13), 1656-1676.
- [169] Warshel A, Lippicirell A. Calculations of ground and excited-state potential surfaces for conjugated heteroatomic molecule. *Journal of the American Chemical Society*, **1981**, 103(16), 4664-4673.
- [170] Lipkowitz KB, Boyd DB. Reviews in computational Chemistry. *Wiley online library*, **1991**, 2, 99-164.
- [171] Chen KH, Allinger NL. Molecular mechanics (MM4) study of saturated four-membered ring hydrocarbons. *Journal of Molecular Structure: Theochem*, **2002**, 581, 215-237.
- [172] Allinger NL, Calculation of molecular structure and energy by force-field methods. *Advances in physical organic chemistry*, **1976**, 13, 1-82.
- [173] Li JH, Allinger NL. Molecular mechanics. The MM3 force field for hydrocarbons.3. The van der Waals' potentials and crystal data for aliphatic and aromatic hydrocarbons. *Journal of the American Chemical Society*, **1989**, 111(23), 8576-8582.

- [174] Allinger NL, Chen K, Lii JH. An Improved Force Field (MM4) for Saturated Hydrocarbons *Journal of computational chemistry*, **1996**, 17, 642-668.
- [175] Weiner SJ, Kollman PA, Nguyent T, Case DA. An all atom force field for simulations of proteins and nucleic acids. *Journal of computational chemistry*, **1986**, 7(2), 230-252.
- [176] Palmer, A. Introduction to molecular modeling, in molecular design and combinatorial chemistry. *Ed. Miertus S, ICS-UNIDO*, **1997**, 1-8.
- [177] Hinchliffe A. Molecular modelling for beginners. *John Wiley and Sons*, **2005**.
- [178] Momany FA, Rone R. Validation of the general purpose QUANTA® 3.2/CHARMm® force field. *Journal of Computational Chemistry*, **1992**, 13(7), 888-900.
- [179] Roterman MH, Lambert KD, Gibson HA, Scheraga. A comparison of the CHARMM, AMBER and ECEPP potentials for peptides. II. - maps for N-acetyl alanine N'-methyl amide: comparisons, contrasts and simple experimental tests. *Journal of Biomolecular Structure and Dynamics*, **1989**, 7(3), 421-453.
- [180] Jing Z, Liu C, Cheng SY, Qi R, Walker BD, Piquemal JP, Ren P. Polarizable force fields for biomolecular simulations: recent advances and applications. *Annual Reviews Biophysics*, **2019**, 6(48), 371-394.
- [181] Patel S, MacKerell Jr AD, Brooks III CL. CHARMM fluctuating charge force field for proteins: I. Parameterization and application to bulk organic liquid simulations. *Journal of Computational Chemistry*, **2004**, 25, 1-15.
- [182] Lemkul JA, Huang J, Roux B, MacKerell AD. An empirical polarizable force field based on the classical drude oscillator model: development history and recent applications. *Chemical Review*, **2016**, 116, 4983-5013.
- [183] Chen J, Martínez TJ. QTPIE: charge transfer with polarization current equalization. A fluctuating charge model with correct asymptotics. *Chemical Physics Letter*, **2007**, 438, 315-20.
- [184] Jorgensen WL, Rives JT. The OPLS (optimized potentials for liquid simulations) potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, **1988**, 110 (6), 1657-1666.
- [185] Hagler A, Lifson S. Energy functions for peptides and proteins. II. Amide hydrogen bond and calculation of amide crystal properties. *Journal of the American Chemical Society*, **1974**, 96(17), 5327-5335.
- [186] Niketic SR, Rasmussen K. The consistent force field: a documentation. *Lecture notes in chemistry*, **1997**.
- [187] Insight-II version 2000 molecular modelling package, discover version 2.98 simulation package, and delPhi version 3.0 solvation program, **2000**, Accelrys, Inc., San Diego, California.

- [188] Dinur U, Hagler AT. New approaches to empirical force fields. *Reviews in computational chemistry*, **1991**, 2, 99-164.
- [189] Discovery studio (version 2.1), Accelrys, Inc., 10188 Telesis Court, Suite 100 San Diego, CA 92121, USA, San Diego, California, **2008**.
- [190] Maple JR, Hwang MJ, Stockfish, TP, Dinur U, Waldman M, Ewing CS, Hagler AT. Derivation of class II force fields. I. Methodology and quantum force field for the alkyl functional group and alkane molecules. *Journal of computational chemistry*, **1994**, 15, 162-182.
- [191] Niketic SR, Rasmussen K. The Consistent force field: a documentation. *Lecture notes in chemistry*, **2012**.
- [192] Discover User Guide, Insight II user guide. *Biosym/MSI, San Diego, CA, USA*, **1995**.
- [193] Molecular Operating Environment Software, *Chemical Computing Group Inc., Montreal, Canada*, **2007**.
- [194] Halgren TA. Merck Molecular Force Field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interaction. *Journal of computational chemistry*, **1996**, 17, 520-552.
- [195] Halgren TA. MMFF VI. MMFF94s option for energy minimization studies. *Journal of computational chemistry*, **1999**, 20, 720-729.
- [196] Halgren TA, Nachbar RB. Merck molecular force field. IV. conformational energies and geometries for MMFF94. *Journal of computational chemistry*, **1996**, 17, 587-615..
- [197] Groenhof G. Introduction to QM/MM simulations. *In Biomolecular Simulations*, **2013**, 924, 43-66.
- [198] Ferrer S, Ruiz-Pernia J, Marti S, Moliner V, Tunon I, Bertran J, Andrés J. Hybrid schemes based on quantum mechanics/molecular mechanics simulations goals to success, problems, and perspectives. *In Advances in protein chemistry and structural biology*, **2011**, 85, 81-142.
- [199] van der Kamp MW, Mulholland AJ. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry*, **2013**, 52(16), 2708-2728.
- [200] Lodola A, De Vivo M. The Increasing Role of QM/MM in Drug. *In Advances in Protein Chemistry and Structural Biology*, **2012**, 87, 337-362.
- [201] Gleeson MP, Gleeson D. QM/MM calculations in drug discovery: a useful method for studying binding phenomena?. *Journal of chemical information and modeling*, **2009**, 49(3), 670-677.
- [202] Menikarachchi LC, Gascon JA. QM/MM approaches in medicinal chemistry research. *Current topics in medicinal chemistry*, **2010**, 10(1), 46-54.
- [203] Le Bel JA. Sur les relations qui existent entre les formules atomiques des corps organiques, et le pouvoir rotatoire de leurs. *Bulletin de la Société chimique de Paris*, **1874**, 22, 337-347.

- [204] van't Hoff JH. Sur les formules de structures dans l'espace. *Bulletin de la Société chimique de Paris*, **1875**, 23, 295-230.
- [205] Bartol J, Comba P, Melter M, Zimmer M. Conformational searching of transition metal compounds. *Journal of computational Chemistry*, **1999**, 20(14), 1549-1558.
- [206] Keseru G, Kolossvary I. Molecular mechanics and conformational analysis in drug design. *Wiley-Blackwell*, **1999**
- [207] Thomas J, Karplus NM. Pseudosystematic conformational search. Application to cycloheptadecane. *Journal of the American Chemical Society*, **1997**, 119(24), 5657-5667.
- [208] Guida WC, Kolossvary I. Low Mode Search. An efficient, automated computational method for conformational analysis: application to cyclic and acyclic alkanes and cyclic peptides. *Journal of the American Chemical Society*, **1996**, 118(21) 5011-5019.
- [209] von der Lieth CW, Kozár T, Hull WE. A (critical) survey of modelling protocols used to explore the conformational space of oligosaccharides. *Journal of Molecular Structure: theochem*, **1997**, 395, 225-244.
- [210] Keserü G, Kolossváry I. Molecular mechanics and conformational analysis in drug design. *Wiley-Blackwell*, **1999**.
- [211] Dugas H. Principe de base en modélisation moléculaire, aspects théoriques et pratiques, Chapitre 3 introduction aux méthodes de minimisation d'énergie. *Librairie de L'Université de Montréal*, **1996**.
- [212] Lipkowitz P. Molecular mechanics in organic synthesis. *Chemical Reviews*, **1993**, 93(7), 2463-2486.
- [213] Jensen F. Introduction to computational chemistry. *John Wiley and Sons: Chichester*, **2007**, 380-420.
- [214] Shewchuk JR. An introduction to the conjugate gradient method. *School of Computer Science Carnegie Mellon University Pittsburgh, Edition 1*, **1994**.
- [215] Kollman PA. Theory of complex molecular interactions: computer graphics, distance geometry, molecular mechanics, and quantum mechanics. *Accounts of Chemical Research*, **1985**, 18(4), 105-111.
- [216] Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. *Science*, **1983**, 220(4598), 671-680.
- [217] Bertsimas D, Tsitsiklis J. Simulated annealing. *Statistical science*, **1993**, 8(1), 10-15.
- [218] Bensaid MO. Thèse de Doctorat. Étude des polymères par dynamiques moléculaire. *Université d'oran des sciences et de la technologie mohamed boudiaf*, **2015**.

- [219] Česlovas V, Krzysztof G, Chulhee K. Sequence-structure mapping errors in the PDB: OB-fold domains. *Protein Sciences*, **2004**, *13*(6), 1594-1602.
- [220] Kerr ID, Lee JH, Farady CJ, Marion R, Rickert M, Sajid M, Pandey KC, Caffrey CR, Legac J, Hansell E, McKerrow JH, Craik CS, Rosenthal PJ, Brinen LS. Vinyl sulfones as antiparasitic agents and a structural basis for drug design. *Journal of Biological Chemistry*, **2009**, *284*, 25697-25703.
- [221] Weldon DJ, Shah F, Chittiboyina AG, Sheri A, Chada, RR, Gut J, Rosenthal PJ, Shivakumar D, Sherman W, Desai P, Jung JC, Avery MA. Synthesis, biological evaluation, hydration site thermodynamics, and chemical reactivity analysis of α -keto substituted peptidomimetics for the inhibition of *Plasmodium falciparum*. *Bioorganic and Medicinal Chemistry Letters*, **2014**, *24*(5), 1274-1279.
- [222] Insight-II and Discover molecular modeling and simulation package, Version 2005, *Accelrys, Inc., San Diego, CA*, **2005**.
- [223] Owono LCO, Keita M, Megnassan E, Frecer V, Miertu S. Design of thymidine analogues targeting thymidilate kinase of *Mycobacterium tuberculosis*. *Tuberculosis Research and Treatment*, **2013**, *2013*, 670836.
- [224] Owono LCO, Ntie-Kang F, Keita M, Megnassan E, Frecer V, Miertus S. Virtually designed triclosan-based inhibitors of enoyl-acyl carrier protein reductase of *Mycobacterium tuberculosis* and of *Plasmodium falciparum*. *Molecular Informatics*, **2015**, *34*(5), 292-307.
- [225] Frecer V, Kabelac M, De Nardi P, Pricl, Miertus S. Structure-based design of inhibitors of NS3 serine protease of hepatitis C virus. *Journal of Molecular Graphics and Modelling*, **2004**, *22*(3), 209-220.
- [226] Frecer V, Jedinak A, Tossi A, Berti F, Benedetti F, Romeo D, Miertus S. Structure based design of inhibitors of aspartic protease of HIV-1. *Letters in Drug Design and Discovery*, **2005**, *2*(8), 638-646.
- [227] Dali B, Keita M, Megnassan E, Frecer V, Miertus S. Insight into selectivity of peptidomimetic inhibitors with modified statine core for plasmepsin II of *Plasmodium falciparum* over human cathepsin D. *Chemical Biology & Drug Design*, **2012**, *79*(4), 411-430.
- [228] Megnassan E, Keita M, Bieri C, Esmel A, Frecer V, Miertus S. Design of novel dihydroxynaphthoic acid inhibitors of *Plasmodium falciparum* lactate dehydrogenase. *Medicinal Chemistry*, **2012**, *8*(5), 970-984.
- [229] Keita M, Kumar A, Dali B, Megnassan E, Siddiqi MI, Frecer V, Miertus S. Quantitative structure-activity relationships and design of thymine-like inhibitors of thymidine monophosphate kinase of *Mycobacterium tuberculosis* with favourable pharmacokinetic profiles. *Royal Society of Chemistry Advances*, **2014**, *4*(99), 55853-55866.

- [230] Frecer V, Seneci P, Miertus S. Computer-assisted combinatorial design of bicyclic thymidine analogues as inhibitors of *Mycobacterium tuberculosis* thymidine monophosphate kinase. *Journal of Computer-aided Molecular Design*, **2011**, 25(1), 31-49.
- [231] Kouassi AF, Kone M, Keita M, Esmel A, Megnassan E, N'Guessan YT, Frecer V, Miertus S. Computer-aided design of orally bioavailable pyrrolidine carboxamide inhibitors of enoyl-acyl carrier protein reductase of *Mycobacterium tuberculosis* with favorable pharmacokinetic profiles. *International Journal of Molecular Sciences*, **2015**, 16(12), 29744-29771.
- [232] Esmel A, Keita M, Megnassan E, Toï B, Frecer V, Miertus S. Insight into binding mode of nitrile inhibitors of *Plasmodium falciparum* Falcipain-3, QSAR and pharmacophore models, virtual design of new analogues with favorable pharmacokinetic profiles. *Journal of Computational Chemistry and Molecular Modelling*, **2017**, 2, 1-22.
- [233] Koehler JEH, Saenger W, van Gunsteren WF. A molecular dynamics simulation of crystalline α -cyclodextrin hexahydrate. *European Biophysics Journal*, **1987**, 15(4), 197-210.
- [234] van Gunsteren WF. The role of computer simulation techniques in protein engineering. *Protein Engineering, Design and Selection*, **1988**, 2(1), 5-13.
- [235] Tirado-Rives J, Jorgensen WL. Molecular dynamics of proteins with the OPLS potential functions. Simulation of the third domain of silver pheasant ovomucoid in water. *Journal of the American Chemical Society*, **1990**, 112(7), 2773-2781.
- [236] McCammon JA, Harvey SC. Dynamics of proteins and nucleic acids. *Cambridge University Press*, **1988**.
- [237] Frenkel D, Smit B, Understanding molecular simulations: From algorithms to applications, *Academic press: San, Diego*, **1996**.
- [238] Warshell A, Levitt M, Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of molecular biology*, **1976**, 103(2), 227-249.
- [239] Bienstoc RJ. Solvation methods for protein-ligand docking. *Fragment-Based Methods in Drug Discovery, Methods in Molecular Biology*, **2015**, 1289, 3-12.
- [240] Abagyan RA, Totrov MM, Kuznetsov DA. ICM: a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry*, **1994**, 15(5), 488-506.
- [241] Honig B, Nicholls A. Classical electrostatics in biology and chemistry. *Science*, **1995**, 268(5214), 1144-11449.
- [242] Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. *Chemical Reviews*, **2005**, 105(8), 2999-3094.

- [243] Brooks BR, Brooks CL III, MacKerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caffisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. CHARMM: The biomolecular simulation program. *Journal Computational Chemistry*, **2009**, *30(10)*, 1545-1614.
- [244] Cossi M, Rega N, Scalmani G, Barone V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *Journal of Computational Chemistry*, **2009**, *24(6)*, 669-681.
- [245] Kleinjung J, Fraternali F. Design and application of implicit solvent models in biomolecular simulations. *Current opinion in Structural Biology*. **2014**, *25*,126-34.
- [246] Guimaraes CRW, Cardozo M. MM-GB/SA resourcing of docking poses in structure- based lead optimization. *Journal of chemical information and modeling*, **2008**, *48(5)*, 958-970.
- [247] Yi CH, Chen JZ, Shi SH, Hu GD, Zhang QG. A computational analysis of pyrazole based inhibitors binding to Hsp90 using molecular dynamics simulation and the MM-GBSA method. *Molecular Simulation*, **2010**, *36(6)*, 454-460.
- [248] Sakalli I, Schöberl J, Knapp EW. mFES: a robust molecular finite element solver for electrostatic energy computations. *Journal of chemical theory and computation*, **2014**, *10(11)*, 5095-5112.
- [249] Cerius 2 Life Sciences molecular simulation software. *Version 4.6, Accelrys Inc., San Diego, CA*, **2002**.
- [250] Gilson MK, Honig B. The inclusion of electrostatic hydration energies in molecular mechanics calculations. *Journal of Computer-Aided Molecular Design*, **1991**, *5(1)*, 5-20.
- [251] Fischer S, Smith JC, Verma CS. Dissecting the vibrational entropy change on protein/ligand binding: burial of a water molecule in bovine pancreatic Trypsin inhibitor. *The Journal of Physical Chemistry B*, **2001**, *105((33))*, 8050-8055.
- [252] Schwarzl SM, Tschopp TB, Smith JC, Fischer S. Can the calculation of ligand binding free energies be improved with continuum solvent electrostatics and an ideal-gas entropy correction? *Journal of Computational Chemistry*, **2002**, *23(12)*, 1143-1149.
- [253] Freccer V, Miertus S. Polarizable continuum model of solvation for biopolymers. *International Journal of Quantum Chemistry*, **1992**, *42(5)*, 1449-1468.
- [254] Copeland R., Lombardo D, Giannaras J, Decicco CP. Estimating Ki values for tight binding inhibitors from dose-response plots. *Bioorganic and Medicinal Chemistry Letters*, **1995**, *5(17)*, 1947-1952.

- [255] Freundlich JS, Yu M, Lucumi E, Kuo M, Tsai HC, Valderramos JC, KaragoyozovL, William R. Jacobs WR, Schiehser GA, Fidock DA, David P. Jacobus DP, Sacchettinid JC. Synthesis and biological activity of diaryl ether inhibitors of malarial enoyl acyl carrier protein reductase. Part 2: 2'-substituted triclosan derivatives. *Bioorganic and Medicinal Chemistry Letters*, **2006**, *16(8)*, 2163- 2169.
- [256] Frecer V, Miertus S, Tossi A, Romeo D. Rational design of inhibitors for drug-resistant HIV-1 aspartic protease mutants. *Drug Design and Discovery*, **1998**, *15(4)*, 211-231.
- [257] Frecer V, Miertus S. Interactions of ligands with macromolecules: Rational design of specific inhibitors of aspartic protease of HIV-1. *Macromolecular chemistry and physics*, **2002**, *203*, 1650-1657.
- [258] Viswanadhan VN, Reddy MR, Bacquet BJ, Erion MD. Assessment of methods used for predicting lipophilicity: application to nucleosides and nucleoside bases. *Journal of computational chemistry*, **1993**, *14(9)*, 1019-1026.
- [259] Discovery studio molecular modeling and simulation program. *Version 2.5, Accelrys, Inc., San Diego, CA, 2009*.
- [259] Ehrlich P. Über den jetzigen stand der Chemotherapie. *Berichte der deutschen chemischen Gesellschaft*, **1909**, *42(1)*, 17-47.
- [260] Schueler FW. Chemobiodynamics and drug design. *Journal of Pharmaceutical Sciences*, **1960**.
- [261] Wermuth CG, Ganellin CR, Lindberg P, Mitscher LA. Glossary of terms used in medicinal chemistry (IUPAC recommendations 1998). *Pure and applied Chemistry*, **1998**, *70(5)*, 1129–1143.
- [262] Hodgson J. ADMET-turning chemicals into drugs. *Nature Biotechnology*, **2001**, *19(8)*, 722-726.
- [263] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, **1997**, *23*, 3-25.
- [264] Lombardo F, Gifford E, Shalaeva MY. *In silico* ADME prediction: data, models, facts and myths. *Mini reviews in medicinal chemistry*, **2003**, *3(8)*, 861-875.
- [265] Gleeson MP, Hersey A, Hannongbua S. *In silico* ADME models: a general assessment of their utility in drug discovery applications. *Current topics in medicinal chemistry*, **2011**, *11(4)*, 358-381.
- [266] Schrödinger: QikProp, version 3.4, LLC, New York, NY, **2011**.

- [267] Cruciani C, Crivori P, Carrupt PA, Test B. Molecular fields in quantitative structure-permeation relationships: the VolSurf approach. *Journal of Molecular Structure: theochem*, **2000**, *503*, 17-30.
- [268] Tetko IV, Bruneau P, Mewes HW, Rohrer DC, Poda GI. Can we estimate the accuracy of ADMET predictions?. *Drug Discovery Today*, **2006**, *11*, 700-707.
- [269] Button WG, Judson PN, Long A, Vessey D. Using absolute and relative reasoning in the prediction of the potential metabolism of xenobiotics. *Journal of Chemical Information and Computer Sciences*, **2003**, *43(5)*, 1371-1377.
- [270] Hou T, Wang J. Structure-ADME relationship: still a long way to go?. *Expert opinion on drug metabolism & toxicology*, **2008**, *4(6)*, 759-770.
- [271] Schrodinger, QikProp (Version 3.7, release 14). New York: *Schrodinger*, **2014**.
- [272] Duffy EM, Jorgensen WL. Prediction of properties from simulations: free energies of solvation in hexadecane, octanol, and water. *Journal of the American Chemical Society*, **2002**, *122*, 2878-2888.
- [273] Jorgensen WL, Duffy EM. Prediction of drug solubility from structure. *Advanced Drug Delivery Reviews*, **2002**, *54*, 355-336.
- [274] Molecular Operating Environment (MOE), **2014**. 10. Chemical Computing Group Inc, 1010 Sherbooke St West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2009.
- [275] Available chemicals directory. <https://blogs.rsc.org/chemical-database-service/>
- [276] Dean PM. Molecular similarity in drug design. *Springer Science and Business Media*, **2012**.
- [277] Lipinski CA. Lead-and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies*, **2004**, *1(4)*, 337-341.

REVIEW ARTICLE

Targeting Cysteine Proteases from *Plasmodium falciparum*: A General Overview, Rational Drug Design and Computational Approaches for Drug Discovery



Boris D. Bekono^{1,†}, Fidele Ntie-Kang^{2,3,*}, Luc C. Owono Owono^{1,*} and Eugene Megnassan⁴

¹University of Yaoundé I, Advanced Teacher Training College, Laboratory for Simulation and Molecular Biophysics, P.O. Box 47 Yaoundé, Cameroon; ²Martin-Luther-Universität Halle-Wittenberg, Institute for Pharmacy, Wolfgang-Langenbeck-Str. 4, 06120 Halle (Saale), Germany; ³University of Buea, Department of Chemistry P.O. Box 63, Buea, Cameroon; ⁴University of Abobo-Adjamé, UFR SFA, Laboratoire de Physique Fondamentale et Appliquée 02 BP 801, Abidjan 02, Cote D'Ivoire

Abstract: Background: The *Plasmodium falciparum* cysteine proteases, also known as falcipains, are involved in different erythrocytic cycle processes of the malaria parasite, e.g. hydrolysis of host haemoglobin, erythrocyte invasion, and erythrocyte rupture. With the biochemical characterization of four falcipains so far, FP-2 (falcipain-2) and FP-3 (falcipain-3), members of the papain-like CAC1 family, are essential haemoglobinases. They could therefore be referred to as potential anti-malarial drug targets in the search for novel therapies, which could ease the burden caused by the increasing resistance to current antimalarial drugs.

Objectives: This review provides a summary of the most important results, highlighting the drug design approaches essential for the understanding of the mechanism of inhibition and discovery of inhibitors against cysteine proteases from *P. falciparum*.

Results: Rational and computer-aided drug discovery approaches for the design of promising falcipain inhibitors are described herein, with a focus on a variety of structure-based and ligand-based modeling approaches. Moreover, the key features of ligand recognition against these targets are emphasized.

Conclusion: This review would be of interest to scientists engaged in the development of drug design strategies to target the cysteine proteases, FP-2 and FP-3.

ARTICLE HISTORY

Received: September 16, 2016
Revised: November 18, 2016
Accepted: December 14, 2016

DOI:
10.2174/1389450117666161221122432

Keywords: Cysteine proteases, computer-aided drug discovery, falcipain, hemoglobin degradation, malaria, *Plasmodium falciparum*.

1. DESCRIPTION OF THE SUBJECT AREA

1.1. Background

The malaria parasite was discovered in 1878 by the French Alphonse Laveran [1]. The incidence of the disease has increased 2- to 3-fold over the last 35 years [2]. In 1982, it was estimated that approximately 215 million people were chronically infected and that some 150 million new cases were reported each year [3-5]. Globally, death cases related to malaria grew rapidly from 1990 and reached a peak of 232 million cases in 2003, with 1.2 million deaths in 2004 alone. In the year 2000, with the adoption of the Millennium Development Goals (MDGs) to improve the state of the world by 2015, one of the eight goals was to reduce the number of

malaria cases by 50% by 2015 [4]. This goal has been achieved. Malaria incidence has indeed declined in recent years in all regions of the world [5]. This is because, between 2000 and 2015, an expansion of interventions against malaria contributed to a 37% decline in malaria mortality rates globally, which reduced an estimated 4.3 million deaths. It is, however, reported that more than 3.2 billion people worldwide are still at risk for malaria, considered high risk for 1.2 billion, with an estimated 214 million malaria cases, leading to approximately 438,000 malaria deaths [5]. Malaria is known to have a large global burden, with strong poverty-promoting effects. It persists as chronic infections, in spite of the available effective medical treatments. Thus, it could be classified as a neglected tropical disease [6, 7]. Diseases like malaria, caused by tropical parasites, represent a major cause of mortality and morbidity in the developing world [8]. Over the last two decades, most of the global funding mechanisms have been directed at HIV/AIDS, malaria, and tuberculosis [9], malaria being one of the deadliest diseases the world is facing. Most of the morbidity cases have been recorded in Africa, South East Asia and South America, the majority being pregnant women and children under five years of age [10].

*Address correspondence to these authors at the University of Yaoundé I, Advanced Teacher Training College, Laboratory for Simulation and Molecular Biophysics, P.O. Box 47 Yaoundé, Cameroon; Tel/Fax: (237) 677300030; E-mail: lcowono@yahoo.fr; and Martin-Luther-Universität Halle-Wittenberg, Institute for Pharmacy, Wolfgang-Langenbeck-Str. 4, 06120 Halle (Saale), Germany; Tel: (237) 677915473; Fax: (237) 333434608; E-mail: ntiekfidele@gmail.com

† These authors contributed equally to this work.

The situation has been worsened by the resistance of the malaria parasites to conventional drugs, particularly in *P. falciparum*, added to the fact that these poverty stricken populations hardly have access to commercialized drugs [11-13]. *P. falciparum* causes hundreds of millions of infections annually and is regarded as the most virulent human malaria pathogen [14-16]. Malaria is a serious problem particularly in Africa, where it is reported to be responsible for 1 in every 5 childhood deaths [17]. It is reported that an African child has between 1.6 and 5.4 episodes of malaria fever each year and every 30 seconds a child dies from malaria [17]. The control of malaria has been hindered by increasing parasite strains resistant to available drugs [18-20]. The history of malaria therapy is long, and a number of small molecules, e.g. quinine, chloroquine, artemisinin and atovaquone, are the treatment of choice [21]. Artemisinin, which is the most efficient currently used anti-malarial drug, is only a single mutation away from emerging drug resistance [22]. Artemisinin combination therapy (ACT), the first-line treatment, was suggested to decrease the transmission of *P. falciparum* and delay the onset of resistance against artemisinin alone in malaria endemic regions [23]. However, the existence of *P. falciparum* strains resistant to ACT was already reported in South East Asia almost a decade ago [24]. This raises concerns about malarial intervention in Africa, where the most virulent and resistant *P. falciparum* strains are prevalent. It is, therefore, imperative to develop other classes of effective anti-malarial drugs, in particular compounds acting against novel biochemical targets. To develop such compounds, it is very important to characterize the structural and biochemical features of new drug targets. Until now, despite the fact that several drug targets have been identified after the unveiling of the *P. falciparum* genome in 2002 [11, 25], only a few anti-malarial targets have been validated *in vivo*. However, as far as availability of new drugs for malaria is concerned, it has to be based on new biological target proteins.

P. falciparum enzymes that break down hemoglobin (Hb) are known to be very promising chemotherapeutic targets, because Hb degradation is a vital process for *P. falciparum* [26]. The protozoan consumes about 75% of the Hb in the infected red blood cells of the human host during the intraerythrocytic stage of the parasite's life cycle [27]. This is the main source of amino acids for the parasite growth and maturation [26]. Falcipains (FPs) are a group of *P. falciparum* cysteine proteases. These are proteolytic enzymes characterized by a cysteine residue in their active site region [28, 29], which are involved in different processes of the erythrocytic cycle of the malaria parasite, e.g. hydrolysis of host hemoglobin, erythrocyte invasion, and erythrocyte rupture [8]. Cysteine proteases are classified into clans and families [30]. Each clan presents different sequences, different structural identities and diverse origins. However, they share the use of a cysteine residue to catalyze the hydrolysis of peptide bonds. Cysteine proteases have been classified into three structurally distinct groups, v.i.z. papain-like (e.g. clan CA, including the cathepsins, FPs, and calpains), ICE-like (e.g. clan CD, including the caspases and gingipains) and picornain-like (e.g. clan PA(C), including the human rhinovirus protease 3C). Different cysteine proteases have been associ-

ated with a variety of human clinical conditions, e.g. caspase-I (psoriasis, inflammation), cathepsin S (autoimmune or inflammatory disorders) and cathepsin K (osteoporosis). They are also present in the life cycle of protozoa, for example, FPs from *P. falciparum* (malaria) and cruzain from *Trypanosoma cruzi* (Chagas disease), as well as in viruses such as HRV 3CLpro from the rhinovirus (common cold) and SARS 3CLpro, an etiological agent responsible for Severe Acute Respiratory Syndrome (SARS) infection. Generally speaking, the mammalian cysteine proteases have been classified into 2 main groups; lysosomal cathepsins (clan CA, family C1) and the caspases (clan CD, family C14). Falcipains from *P. falciparum* are members of clan CA, family C1. Therefore, selectivity of a falcipain inhibitor against the parasite proteases, compared with the corresponding human targets, must be defined as the lack of activity against cathepsin cysteine proteases.

In the search for novel therapies that would ease the burden caused by the increasing resistance to current anti-malarial drugs, cysteine proteases constitute promising targets. Four of the FPs have been characterized biochemically so far. However, as a result of their interesting hemoglobinase capacity and the ample availability of tools, the search for new FP inhibitors has been limited to falcipain-2 (FP-2) and/or falcipain-3 (FP-3) [31]. As a result of the availability of several X-ray structures of FP-2 and FP-3 in the protein data bank (PDB) [32], several potent inhibitors have been developed, including the use of computational approaches. This review will first provide a survey of the available information on the cysteine proteases from the genus *Plasmodium*, then describe progress towards the discovery of promising FP inhibitors, with prospects for the exploitation of this class of enzymes as drug targets. Finally, we would provide highlights on various rational approaches for the discovery of inhibitors against cysteine proteases with examples for FP-2 and FP-3 will be provided. In writing this review, the authors intend to first provide an overview of the general subject, then examine rational approaches to drug design before examining the computational techniques which have been used in this field and presenting the main results.

1.2. *Plasmodium falciparum*: Life Cycle of the Parasite

To rationally develop new anti-malaria agents, it is essential to study the genetics and physiology of *P. falciparum*. It is equally important to understand the *P. falciparum*-host interaction in order to learn how the parasite circumvents host defenses and causes disease. The life cycle of *Plasmodium* parasites is completed through many morphological changes, which alternate between an invertebrate and vertebrate host. *P. falciparum* has three critical points in its life cycle, during which a small number of parasites multiply rapidly in order to generate much larger populations [33]: the mosquito vector stage, liver and blood stages (Fig. 1) [33, 34]. The two last processes take place in the vertebrate (human) host. All malaria types have a similar life cycle. The female and male gametocytes (the only stages within the parasite's life cycle able to mediate the transition from the human host to the insect host [35]) mate inside the gut of the mosquito vector and go through meiosis and thereafter migrate to the midgut wall of the vector mosquito and produce

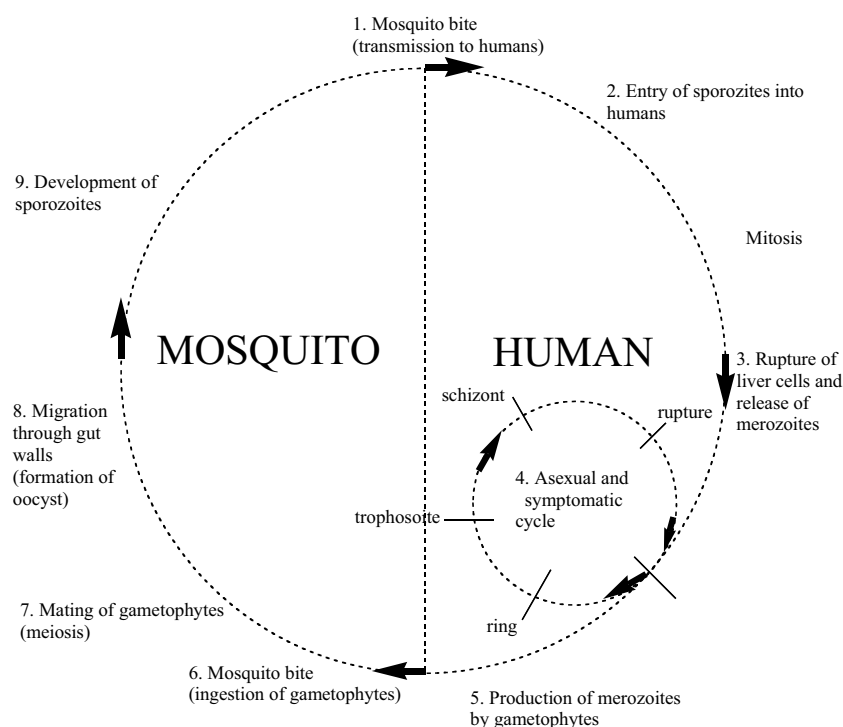


Fig. (1). The malaria parasite life cycle [54].

an oocyst [34], in which many sporozoites develop [36]. Sporozoites, the infectious form of the parasite, are injected into a human host through the saliva of an *Anopheles* mosquito when it bites the human host to feed on blood. These sporozoites travel to the liver, where they invade the human liver cells within a few minutes, thus taking a new form and multiplying (exo-erythrocytic forms). During the process of liver cell rupture, blood stage parasites (also known as), are released, thereby constituting the erythrocytic cycle/asexual blood stage. Each merozoite invades a red blood cell. After two days the parasite starts repeating mitotically. This continues through some stages (*e.g.* ring, trophozoite and schizont), thus producing more merozoites [37, 38]. This is the stage of the life cycle which causes disease, often leading to death if nothing is done [12, 39]. Some merozoites change into gametocytes, a form which does not cause disease but remains in the blood until cleared by drugs or the immune system, or taken up by a mosquito bite. In the mosquito's midgut "male" gametocytes fertilize the "female" to form oocysts [40], which mature into thousands of sporozoites, eventually swimming to the mosquito's salivary glands to initiate another life cycle (Fig. 1). The form of the parasite best studied is the asexual blood stage, largely due to the fact that this is the disease-causing stage. Its experimental accessibility is relatively easy [37]. Thus, most *Plasmodium* proteins that have been well characterized are expressible during the erythrocytic cycle. Although *P. falciparum* proteases may be involved in many processes, only three functions (hydrolysis of host hemoglobin, erythrocyte invasion, and erythrocyte rupture) in erythrocytic parasites are well defined. Therefore, potential anti-malarial drug targets could be broadly classified into three mainly categories: hemoglobin

(Hb) degradation enzymes, signaling and transport protein factors [11]. FP cysteine proteases are mainly involved in Hb degradation and only play a minor role in erythrocyte rupture.

1.3. Hb Degradation by *Plasmodium falciparum*

To obtain the nutrients required for growth during the intra-erythrocytic stage, the parasite has to break down a great amount of hemoglobin from the host cell [41]. The Hb degradation process occurs in food vacuoles of the parasite where 25-75% hemoglobin is digested, thereby liberating the required amino acids for incorporation into parasite proteins [11, 42, 43]. Several proteases are involved in this degradation [11, 44] and they can be divided into two major functional groups:

- proteases involved in the invasion and rupture of erythrocytes
- proteases involved in hemoglobin hydrolysis [45].

In the latter, two relevant protease families: aspartic proteases and cysteine proteases, called plasmepsins and falcipains (FPs), respectively [26, 46], one metalloprotease (falcilysin) [47, 48], and one dipeptidyl aminopeptidase [49] are involved. The degradation process follows an ordered pathway (Fig. 2) [50]. Therefore, blocking the process (or inhibiting any of these enzymes) would be lethal to the parasite. We can therefore conclude that the inhibition of hemoglobin degradation by cysteine protease inhibitors would lead to parasites containing abnormal food vacuoles full of undegraded host hemoglobin.

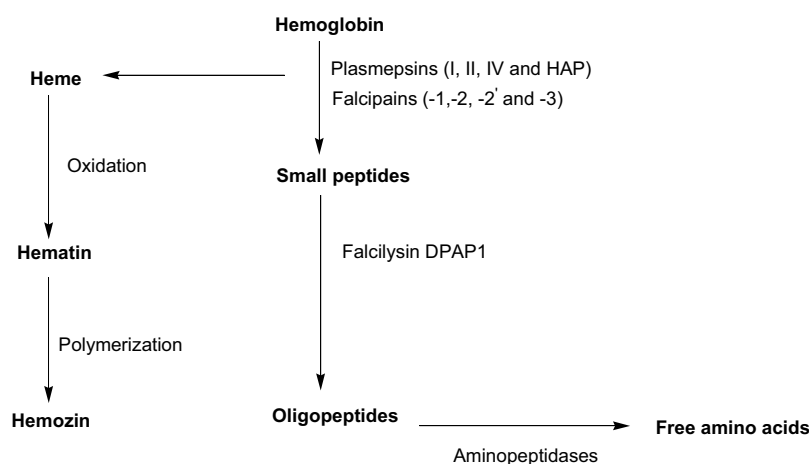


Fig. (2). Summary of the degradation process of hemoglobin in the *P. falciparum* food vacuole [50].

1.4. Drug Targets

The targeting of parasite cysteine proteases with small molecules is a possible approach to developing drugs for the treatment of tropical parasitic diseases, *e.g.* sleeping sickness, Chagas' disease, and malaria [8]. One approach is to search for drug leads among existing small molecule libraries. Many of the currently used anti-malarial drugs were not, however, developed on the basis of rationally identified targets. On the contrary, they were discovered by the identification of the anti-malarial activity of natural products (NPs) (*e.g.* quinine and artemisinin), compounds chemically related to NPs (*e.g.* chloroquine and artesunate), or compounds active against other infectious pathogens (*e.g.* antifolates and tetracyclines [51]). By the mid-1950s, it was confidently expected that malaria would be eradicated worldwide. However, by the mid-1960s these expectations were cut off by the problems of resistance [52]. The vector mosquito developed resistance to potent insecticides such as dichlorodiphenyltrichloroethane and certain strains of *P. falciparum* became resistant to chloroquine treatment, so that by the early 1980s, several strains of *P. falciparum* had become multi-drug resistant [3]. Although resistance to these available drugs tends to be much less widespread geographically, in some areas of the world today the impact of multi-drug resistant malaria can be extensive. Although much of the present research work focuses on the identification of new drugs, researchers have little knowledge of the mode of action and the complete resistance mechanism to some of the known drugs [53]. Malaria drug resistance is often associated with genetic mutations in genes encoding for target proteins. This may result in an unusual expression and folding of these proteins, thus modifying the usual binding site of the drug. Drug resistance in *Plasmodium* is mediated by two processes, which include: (i) the rate of appearance of *de novo* mutations that confer resistance, selected through drug use within an individual; and (ii) the rate of spread of resistant alleles to other individuals [54]. It therefore stands out clearly that the eradication of malaria will require the design and development of new therapeutically active agents against, particularly, the disease causing *P. falciparum* resistant strains.

The development of resistance by the parasite against first line and second line anti-malarial drugs has underscored the importance to develop new drug targets and pharmacophores to treat the disease. With the availability of increasing support for malaria research, a number of drug targets and candidate molecules have been made available for further development. These efforts are, however, still relatively limited. The success rate of candidate molecules to become drugs is also very low, thus making it increasingly necessary to develop other lead compounds, identified on a rational basis, thus facilitating the development of such compounds. It is also important to characterize the biochemical features and biological roles of new drug targets [55].

The design of anti-malarial compounds targeting plasmepsin and FP inhibition by structure-based drug design (SBDD) approaches has received much attention, due to their potential biomedical use. Earlier studies indicated that *P. falciparum* had a limited capacity for *de novo* amino acid synthesis. The parasite therefore depends on hemoglobin hydrolysis as a source of amino acids for protein synthesis [56]. Hemoglobin is initially degraded by aspartic proteases (*e.g.* plasmepsins I, II, and possibly III and IV) located in the digestive vacuole [11, 48, 56]. There are a total of 30 predicted cysteine proteases [14], which have been named according to the function of a catalytic cysteine mediating protein hydrolysis *via* the nucleophilic attack on the carbonyl carbon of a susceptible peptide bond [25, 57]. Four cysteine protease FPs have been identified so far in the *P. falciparum* genome: falcipain-1 (FP-1), FP-2 and -2' (now known as FP-2A and -2B), and FP-3. Only FP-2A, -2B, and -3 have been shown to be involved in Hb degradation, further degrading the plasmepsin-digested products. The generated short peptides are finally digested by the metalloprotease falcilysin [11, 46, 47]. The term "FP" describes a group of papain-family cysteine proteases, which are expressed at different stages during the development cycle of the most virulent human malaria parasite, *P. falciparum*. However, to the best of our knowledge, no commercial drug-targeting FPs have been developed despite their central role in the life cycle of the parasites.

1.4.1. Falcipain-1 (FP-1)

The first cysteine protease gene cloned from *P. falciparum* was named FP-1. The physiological role of FP-1 in asexual blood stages of the parasite is still uncertain, while its primary function is more likely to be in the mosquito-stage parasites [58]. However, some studies suggested that FP-1 acts as a hemoglobinase. Greenbaum *et al.* [59] showed that FP-1 activity was associated with erythrocyte invasion by *P. falciparum* merozoites and specific inhibitors of FP-1 blocked parasite invasion of the host erythrocyte, but had no effect on the parasite Hb-degradation activity. Eksi *et al.* [58] and Sijwali, *et al.* [60] independently carried out knock-out studies, which suggested that FP-1 could play a role in oocyst production. However, FP-1 does not seem to have a critical role in the erythrocytic stages of the *P. falciparum* parasite. FP-1 is encoded on chromosome 14 and is a distant relative of the other falcipains in terms of sequence (showing <40% amino acid sequence identity) and function [61].

1.4.2. Falcipain-2 (FP-2) and falcipain-2' (FP-2')

FP-2 (or FP-2A), is the most extensively studied enzyme within this family. Recent functional and structural data suggest that FP-2 is an attractive target for therapeutic intervention among the identified *P. falciparum* drug targets [62, 63]. The best studied role of FP-2 is its Hb degradation [64]. FP-2 is a principal cysteine protease and essential hemoglobinase of erythrocytic *P. falciparum* trophozoites (which constitutes more than 93% of its role) in the total cysteine protease activity in the food vacuole during the early and late trophozoite stages of the parasite [31]. Many *in vitro* studies have confirmed that inhibitors of FP-2 can block hemoglobin hydrolysis in the parasite, thereby halting the development of culture parasites [65-67]. Evidence of the ability of FP-2 to activate at least one other protease (named plasmepsin II), by cleaving off the propeptide, has been provided [68]. The literature indicate that FP-2' (or FP-2B) is a near-identical copy of FP-2, constituting about 98% identity in mature domains [69, 70]. FP-2' is indistinguishable from FP-2 biochemically, the former also expressed in erythrocytic trophozoites. FP-2' function of is uncertain, although it is hypothesized that a high complementarity exists between FP-2 and -2' in the process of hemoglobin degradation within the host, indicating that the FP-2' is not essential. FP-2 and FP-3 (with 65% amino acid sequence identity) resembled the papain family Clan CA cysteine proteases that cleave native or denatured human hemoglobin and are predominantly expressed in the acidic food vacuole of trophozoites and schizonts [71-73]. Nowadays, the search for potent FP-2 and FP-3 inhibitors is a hot topic in the search for potential malaria drugs.

1.4.3. Falcipain-3 (FP-3)

FP-3, a third putative cysteine protease gene, was identified in the year 2000 prior to its being fully sequenced, described and characterized in 2001 [74]. FP-3 is very similar to FP-2 biochemically. It has been, however, reported that FP-3 is expressed later in the life cycle and appears to be a more efficient haemoglobinase than FP-2 [75]. Thus, cysteine protease inhibitors which inhibit both FP-2 and FP-3 consistently block hemoglobin hydrolysis and parasite development. This suggests that the two proteases are key tar-

get enzymes [76, 77]. More importantly, FP-3 is more stable than FP-2 at moderately acidic pH, and also less stable at neutral pH, suggesting differences in physiological roles for the two proteases [66]. It has been discovered that the concentration of FP-2 in trophozoites is 1.8 times that of FP-3. However, the latter appears to cleave native hemoglobin about twice as rapidly as the former [78]. This indicates that FP-2 and FP-3 contribute almost equally to the digestion of hemoglobin in the food vacuole. The relative contribution of the two enzymes to the hydrolysis of native hemoglobin is, therefore, essentially equivalent, making both target enzymes equally important targets for the inhibition of hemoglobin degradation. Comparison of FP-2 and FP-3, in the presence of smaller ligands, showed that FP-3 is far less active against peptide substrates [13, 74]. This makes the latter target less amenable to inhibition by peptidyl-based small molecules. Additionally, the difference in nature of parasitic cysteine protease and its human orthologues have prompted interest in this class of proteases for drug development [79]. Thus, the significance of FP-2 and FP-3 in the *Plasmodium* life cycle and the availability of several three-dimensional (3D) structures of these enzymes and detailed information available about the interaction between a ligand and its target make both the selected proteins ideal targets for anti-malarial therapy. Thus, the identification of possible molecules that will inhibit the activity of the target protein, using new approaches in rational drug design, is very crucial [80]. Moreover, the combination of high-throughput screening, computer-aided drug discovery (CADD) methods and the availability of FP X-ray structures has contributed to the discovery of potent inhibitors, which are currently able to inactivate these *P. falciparum* target enzymes in a reversible or irreversible manner.

1.5. Domains of Falcipains

It is generally known that FPs have two main domains; the prodomain and the mature domain (Fig. 3). The prodomain has an N-terminal part, required for trafficking FPs to the food vacuole, while the C-terminal part of the prodomain is required for inhibiting the mature domain. The N-terminal part of the mature domain is required for refolding. It has been shown that the prodomain contains unique motifs responsible for targeting the FPs [71]. FP-2 and FP-3 share two unique insertions that distinguish plasmodial cysteine proteases from all other structurally characterized papain-family enzymes [81]. It has been recently demonstrated that the first insertion is located at the N-terminus of the mature enzyme [81]. It has been shown from previous functional studies that the N-terminal extension of FP-2 and FP-3 plays a crucial role in folding of the mature protein into its active conformation [82]. FP-2 and FP-3 efficiently hydrolyze human hemoglobin in the acidic food vacuole of parasite [76, 83]. Pandey *et al.* [71] showed that FPs have a hemoglobin binding domain near the C-terminus end of the mature domain (Fig. 3).

1.6. Key Aspects of FP-2 and FP-3 Active Sites for Enzyme Inhibition

It has been reported that the active sites of FP-2 and FP-3 are located within a cleft between the structurally distinct domains of the papain-like fold and three amino acids residues

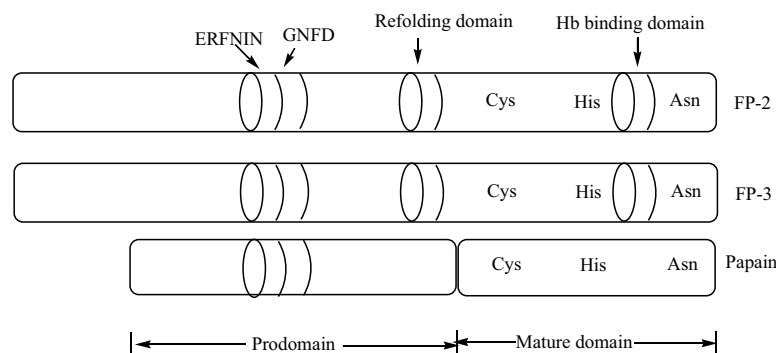


Fig. (3). Domains of falcipains. The active site residues (Cys, His, Asn) are conserved within the papain family, but falcipains have a unique N-terminal extension, which acts as a refolding domain, and a C-terminal inserted as a hemoglobin (Hb) binding domain. The prodomain has the ERFNIN and GNFD motifs, which are conserved in falcipains and papain [71].

Cys42/51, His174/183, Asn204/213, respectively (His159, Cys25 and Asn175 following papain convention, Fig. 3), are responsible of the catalytic activity [71]. The binding site of FP-2 and FP-3 is large enough to be subdivided into pockets S1-S4 for a better analysis of the interactions between the amino acids of each pocket and inhibitors [75]. From the literature, the general structure of a cysteine protease inhibitor (Fig. 4) should comprise an electrophilic moiety for interaction with the cysteine residue of the enzyme and one or two series of substituents, P1-P4, which are responsible for interactions with the different pockets of the protease (Fig. 5).

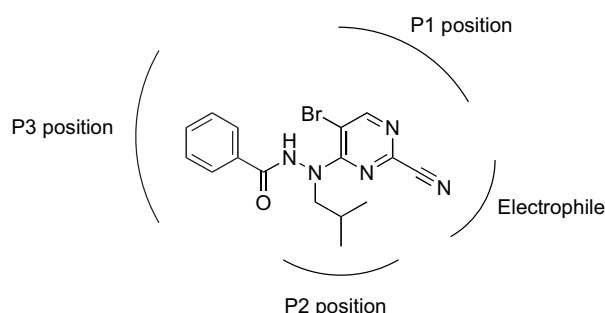


Fig. (4). General structure of 2-pyrimidinecarbonitrile derivatives, showing positions P1, P2 and P3 [75].

1.7. FPs: Crystallized Structures

Traditionally, the first consideration before embarking on a CADD project is whether the detailed 3D structure of the drug target is known. SBDD of anti-malarial compounds targeting FP inhibition is possible due to the availability of several 3D-structures of these enzymes. Four crystal structures of FP-2 from *P. falciparum* have been deposited so far in the protein databank (PDB) [32], three of which correspond to the apo-enzyme (PDB entry: 2GHU, 1YVB, 3PNR) [84, 85], and the other two protein-inhibitor complexes (PDB entry: 3BPF). Meanwhile, there are 4 chains in 3BPF (A, B, C, and D) and epoxysuccinate E64 as an inhibitor on each chain [86, 87]. For FP-3, two crystal structures from *P. falciparum* have been deposited so far, one of which correspond to the apo-enzyme (PDB entry: 3BPM) and the other correspond to the protein-inhibitor complex (PDB entry: 3BWK),

while there are 4 chains (A, B, C, and D) in 3BWK, and leupeptin as an inhibitor on each chain [81, 87]. E64 and leupeptin interact with residues of both enzymes in the S1, S2, and S3 subsites, corresponding to the P1, P2, and P3 positions of the ligands (Fig. 6).

2. DRUG DESIGN APPROACHES AGAINST FALCIPAINS

Nowadays, many compounds have been identified as inhibitors of FPs, which are able to block the enzyme's activity by forming a reversible or irreversible covalent bond with the amino acids of active site cysteine. These FP inhibitors can be broadly divided into three categories [2]:

- i). peptide-based inhibitors,
- ii). non peptidic inhibitors and
- iii). peptidomimetic inhibitors.

Most of the FP inhibitors identified so far are peptide-based inhibitors [2]. Peptide-based inhibitors include: peptidyl fluoromethyl ketones [88], peptidyl vinyl sulfones [89], peptidyl aldehydes and -ketoamide derivatives [90], epoxy-succinyl derivatives [91] and peptidyl azirines [92]. Peptidyl aldehydes and -ketoamides are reversible inhibitors, while the other inhibitors are irreversible. The above compound classes were found to inhibit FP enzymatic activities at very low nanomolar range. However, their utility as therapeutic agents, is limited by their poor absorption through cell membranes and susceptibility to protease degradation [2]. The drug design approaches discussed here include, rational approaches for the discovery of non-peptidic and peptidic inhibitors against cysteine proteases, along with examples of FPs, SBDD and CADD approaches for the design of peptidomimetic. We also include NP-based drug discovery and provide insight for the understanding of drug discovery process against both the targets. We will then review steps towards the discovery of the inhibitors against FP-2 and FP-3 using rational approaches.

2.1. Rational Approaches

2.1.1. Knowledge-based Drug Design

In this context, the concept of drug discovery involves knowledge-based drug design based on the similarity of a

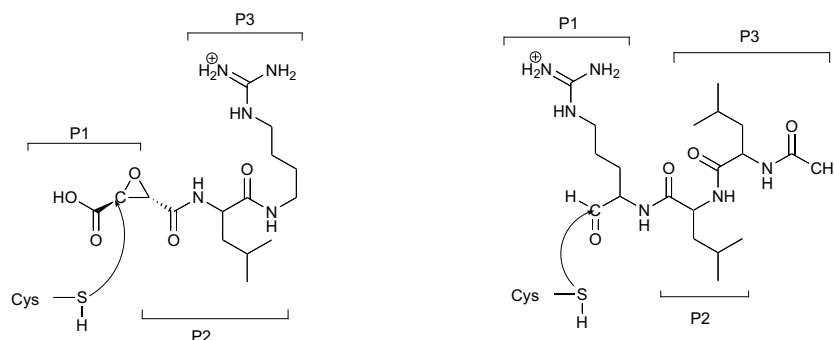


Fig. (5). Chemical structures of E64 and leupeptin. The positions of P1, P2 and P3 of the inhibitors that occupy the S1, S2 and S3 subsites, respectively, of the enzyme, are labeled. Inhibitor and enzyme groups involved in covalent bond formation are shown in red [71].

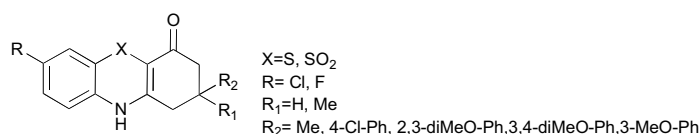


Fig. (6). General structure of phenothiazines.

drug target with other members of the same or different protease family under consideration. This combines information from the knowledge of the substrates and the known inhibitors of the homologous proteases. This is particularly important when the options for treatment against a particular disease are limited. In the case of FPs, drug design efforts exploited previous knowledge of inhibitors of other cysteine proteases [93]. Two major classes of cysteine protease inhibitors are known, small inhibitors like leupeptin, vinyl sulfones, E64, and macromolecular inhibitors like falstatin, PbICP, PyICP, chagasin and cystatin [94]. The design of inhibitors, based on the interactions inhibitor-enzyme complex, is a known approach. Earlier, this approach had deliberately been avoided by small molecule drug developers, largely because of technological hurdles [94].

Dominguez *et al.* [95], designed and evaluated, a series of phenothiazin-4-ones and phenothiazin-4-one-5,5-dioxides as FP inhibitors. The authors were inspired by acridinediones with anti-malarial activity and attempted eliminating their toxicity by replacing the carbonyl group by a sulfur or sulfone isoster [96]. No differences in potencies were observed between the phenothiazine-4-ones and their corresponding 5,5-dioxides. The most potent compounds amongst the phenothiazines showed IC₅₀ values against FP-2 in the lower micromolar range (4–5 μM, Fig. 6). The authors therefore concluded that, although FP inhibition is observed among the phenothiazine derivatives, the antiplasmodial effect could most likely be attributed to the acridinedione-related mechanism.

Greenbaum *et al.* [97] tested pyridyl- and phenylthiosemicarbazone derivatives against both FP-2 and parasites in culture. Most of the compounds failed to show any target enzyme inhibition. Moreover, only weak inhibitors showed no efficacy against the parasite. However, these compounds showed good antiplasmodial activities, which implies that they act through a different mechanism, which the authors attributed to the chelation of endogenous metals. All the thi-

osemicarbazone derivatives tested showed better antiplasmodial activities even though these derivatives proved to be weak FP-2 inhibitors. Chiyanzu *et al.* [98] in the same direction, have designed a new series of compounds based on the isatin scaffold and tested them as FP-2 inhibitors. It was observed that unsubstituted derivatives of isatin exhibited little or no enzymatic activity. *N*-Benzoylation, in combination with substitution at position 5 of the isatin scaffold seemed to be allowed, even though the most potent derivative within this series was the one with an unsubstituted nitrogen (IC₅₀ = 4.4 μM) (compound 1, Fig. 7). On the other hand, Chibale *et al.* [99] further tested a series of hybrid derivatives linking isatin and a thiolactone anti-malarials. All the tested compounds were shown to be inactive or only weak FP-2 inhibitors. Better activities were, however, observed against cultured parasites, suggesting that FP-2 inhibition is not the primary mechanism of action for these compounds. From the above observations, it appears that the reversible cysteine protease inhibitors were the thiosemicarbazones and isatins (cyclic conformationally restricted ketoamides). However, weak FP-2 inhibition was observed despite good antiplasmodial activity, which could be attributed to their ability of the tested compounds to chelate endogenous metals.

Several efforts towards the development of FP-2 inhibitors have focused on peptide based inhibitors [2, 100]. Rockett *et al.* [101] and Rosenthal *et al.* [102] have suggested that peptide fluoromethyl ketones both blocked hemoglobin degradation and parasite growth at nanomolar concentrations. The most potent of these inhibitors (compound 2, Fig. 7) inhibited trophozoite cysteine proteinase proteolytic activity at sub-nanomolar concentration (IC₅₀ = 0.36 nM), with an anti-plasmodial IC₅₀ of 64 nM in the nanomolar range. A related fluoromethylketone, named morpholine urea-Phe-homophenylalanine (hPhe)-CH₂F, was able to cure 80% of the treated *P. vinckei*-infected mice in a murine malaria model [103]. In addition, the compound showed a 90% blockage of the proteolytic activ-

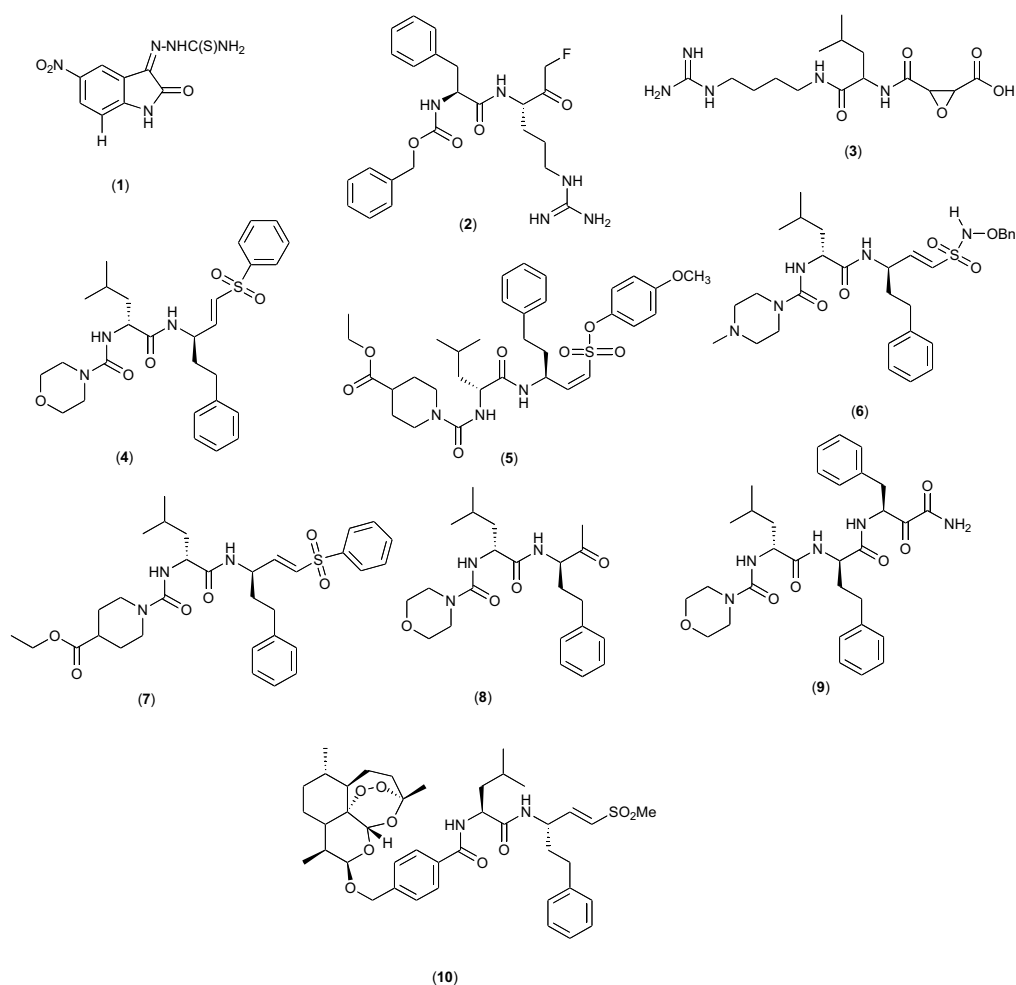


Fig. (7). Examples of active FP-2/3 inhibitor compounds designed using knowledge-based drug design approaches.

ity of *P. vinckei* at 2, 4, 6 and 8h after initial treatment. This offered an early indication that targeting hemoglobin catabolism might be practical. However, severe toxicity in the tested animals has hindered further drug development of the fluoromethyl ketones [66].

E64 (compound 3, Fig. 7) the identified papain-like cysteine protease reversible inhibitor, contains a *trans*-(2*S*,3*S*) configured epoxide ring; meanwhile, the amino acid residue of the peptidyl part of the inhibitor has the *L*-configuration [104]. Epoxysuccinyl peptides derived from E64, which contain *trans*-epoxysuccinic acid as the electrophile, are selective papain superfamily cysteine protease inhibitors. The inhibitory mechanism involves the irreversible activation of the enzyme by alkylation of the active site cysteine [105]. Several reports have been published on the efficacy of E64 and its derivatives as antiplasmodial agents with nanomolar range activities against FPs and parasites in culture [102, 106-108]. However, the lack of selectivity against different members of the papain family has limited the further development of these compounds into drugs. Various mechanism-based vinyl sulfone inhibitors had been reported showing activities against closely related cysteine proteases during

the late 1990s [109]. For example, Rosenthal and coworkers evaluated the anti-malarial activities of a series of vinyl sulfones as cysteine proteases inhibitors [66]. The report showed that some of the vinyl sulfones were non-toxic, yet active against *P. falciparum* FPs, also exhibiting anti-parasitic activity. Several of these compounds exhibited nanomolar activities, among whom compound K11017 (named Mu-Leu-hPh-VSPh, compound 4, Fig. 7) has been co-crystallized with FP-3 [87]. This compound was found to inhibit cysteine proteases of *P. falciparum*, with an activity in the sub-micromolar concentration ($IC_{50} = 0.003 \mu M$), also blocking Hb-degradation and parasite development. Several structural changes have been carried out on K11017, leading to derivatives with better activities than the parent molecule. The study also showed the limited utility of the derived compounds as therapeutic agents. This was due to their susceptibility to protease degradation and poor cell membrane absorption. Olson *et al.* [110] synthesized a series of vinyl sulfones that blocked FP activity, parasite Hb-degradation, and parasite development at nanomolar concentrations. From the results obtained from a study conducted on this set of compounds, it can be observed that a 2-naphthylvinyl sulfone group offers improved activity, particularly against

P. vinckeii *in vivo*, comparable with the phenyl vinyl sulfone both in terms of the inhibition of FP-2 and the inhibition of biological activities of cultured parasites. Inspired by the previously cited work, Shenai *et al.* [111] synthesized and tested 39 new compounds having vinyl sulfones, vinyl sulfonate esters and vinyl sulfonamide as warhead groups to establish the structure activity relationships. Inhibitors with Leu at the P2 position and homophenylalanine (HPhe) at the P1 position displayed potent inhibitory activity against FP-2, FP-3, and cultured *P. falciparum*. These observations point towards the importance of these residues for substrate specificity against FP-2. The vinyl sulfonate esters were the most active (compound 5, Fig. 7), with $IC_{50} = 0.9$ nM against FP-2, followed by the vinyl sulfonamides (compound 7, Fig. 7), with $IC_{50} = 2.3$ nM against FP-2, and finally phenyl vinyl sulfones (compound 7, Fig. 7), with $IC_{50} = 6.9$ nM against FP-2.

Lee and colleagues [90] have developed peptidyl reversible inhibitors of FP-2/3, also showing activity against cultured *P. falciparum* parasites include peptidyl aldehydes and α -ketoamide derivatives developed. The vinyl sulfones, sulfonamides and sulfonate esters were shown to be the most potent, with similar inhibitory activities against the two enzymes and inhibition of parasite development correlated with activity against FP-2. Among all the compounds tested against FP-2 and parasites in culture, Leu showed a marked preference for the P2 position, while morpholino, Z-(benzyloxy) and N-methylpiperazinyl groups preferred the P3 position, yielding good *in vitro* potencies. Representative compounds from the aldehyde (compound 8, Fig. 7), with $IC_{50} = 1$ nM against FP-2, $IC_{50} = 1.1$ nM against the W2 strains of the parasite and ketoamide series (compound 9, Fig. 7), with $IC_{50} = 1$ nM against FP-2, $IC_{50} = 2.9$ nM against the W2 strains of the parasite. Both compounds were active against multiple strains of *P. falciparum*.

Peptidyl aziridine derivatives are known to be selective and irreversible cysteine protease inhibitors. These are aza analogues of epoxides [94]. Schulz *et al.* [91] evaluated the FP-2 and FP-3 inhibitory activities of three classes of compounds; aziridines, aziridine-2-carboxylic acid derivatives, N-acylated aziridine-2,3-dicarboxylic acid derivatives and aziridine-2-carboxylates containing a lysine residue. Beside aziridine-2-carboxylic acid derivatives, there are three types of aziridinyl peptides (Fig. 8), classified to be structurally different [112]. The first type is N-acylated with the aziridine moiety located on the C-terminus of the peptides or amino acids. The second type includes aza-analogs of epoxysuccinyl peptides and has the aziridine moiety at the N-terminus of the amino acid. Meanwhile, the third type are N-acylated (as type I), but with the aziridine moiety located in the middle of the peptide chain. A multiple ligand approach was reported against FP-2. This involved the combination of artemisinin, a sesquiterpene lactone, and a potent anti-malarial compound with dipeptidyl vinyl sulfones *via* a linker to combat multi-drug resistance in *P. falciparum* malaria [113]. Compound 10 (Fig. 7) exhibited potent inhibitory activity of FP-2 (FP-2 $IC_{50} = 0.21$ nM). Additionally, this compound exhibited better activity against multiple strains of malaria, when compared with artemisinin and chloroquine.

2.1.2. Structure-based Drug Design

Batra *et al.* [114] designed several isoquinoline and dihydroisoquinoline analogues, with the goal of further optimizing the activity of compound 11 (Fig. 9), by considering the interaction of designed analogues with the Asp234 of the S2 pocket as essential. Homology models of FP-2 and FP-3 were developed and validated, followed by docking of substrate/known vinyl sulfone inhibitors. The authors further synthesized analogues, which exhibited activities varying within the range of IC_{50} values from 3 to 10 μ M. The most active compound (12, Fig. 9) showed an IC_{50} value of 3 μ M against FP-2. Goud *et al.* [115] also exploited the fact that the thiazoline core favored key interactions and the spatial arrangement of the designed tri-substituted thiazoles in the active site of FP-2, and therefore addressed the *de novo* design of novel non-peptidic inhibitors, which bear the thiazoline scaffold. The authors synthesized and evaluated the biological activities of a series of tri-substituted thiazole analogues against FP-2 and FP-3. This led to the discovery of compound 13 (Fig. 9) as a dual inhibitor of both enzymes ($IC_{50} = 6.6$ μ M against FP-2; $IC_{50} = 29.4$ μ M against FP-3).

2.1.3. Natural Products

Natural products (NPs) show a high degree of structural diversity, coupled with a wide range of biological activities, thereby often constituting a valuable source of lead structures for drug discovery [116]. However, further development of NPs into drugs is often hampered by lack of knowledge of their modes of action. Several strategies for the identification of the direct targets of active NPs have been developed in the past few years, including the FPs [117], *e.g.* following the discovery of the anti-malarial activity of lipochalcone A, Chen *et al.* [118] developed several chalcone derivatives (*e.g.* compound 14, Fig. 10), from Chinese liquorice root. These compounds exhibited potent anti-malarial activity [119-121]. Molecular modeling studies, as well as experimental evaluation of the chalcone derivatives, showed an anti-malarial activity *via* multiple mechanisms of action, including FP-2 inhibition. Moreover, compound 15 (Fig. 10, with $IC_{50} = 1.8$ μ M against FP-2) was shown to be a promising compound within this series [122]. The work reported by Stolze *et al.* [123] served as a valuable guide for the future rational design of potent FP-2/3 inhibitors by providing the mechanistic basis for the observed anti-malarial potency of the marine NP symplostatatin 4 (compound 16, Fig. 10).

Isatin, a NP from the plants of the *Isatis* genus, proved to be a useful starting point for the design and synthesis of chemical libraries with activities against cysteine and serine proteases, *e.g.* FP-2 [124, 125]. The derived compounds showed a broad range of activities, from higher to lower micromolar levels. Compound 17 ($IC_{50} = 4.4$ μ M against FP-2) and compound 18 ($IC_{50} = 0.375$ μ M against FP-2) are the most promising members of this class FP-inhibiting of NP derivatives (Fig. 10). Both compounds also inhibited cysteine proteases of trypanosomes, rhodesin and cruzain.

2.1.4. Peptidomimetic Design

Peptidomimetic FP-2/3 inhibitors exhibiting anti-malarial activity in the micromolar range have been designed and

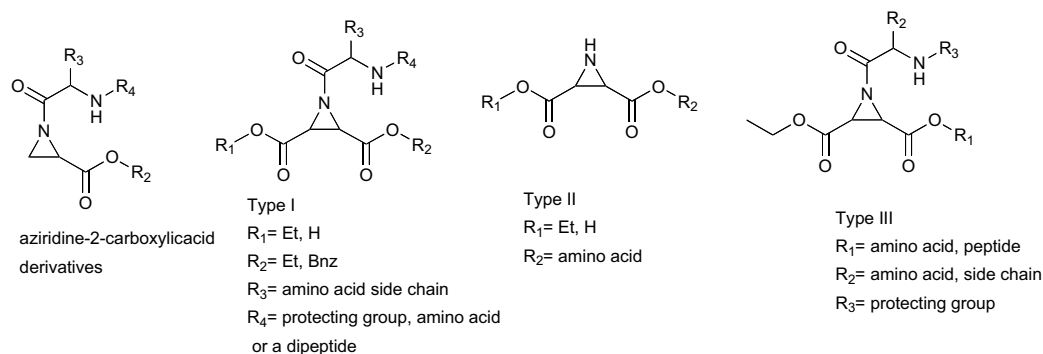


Fig. (8). Types of aziridinyl peptides [114].

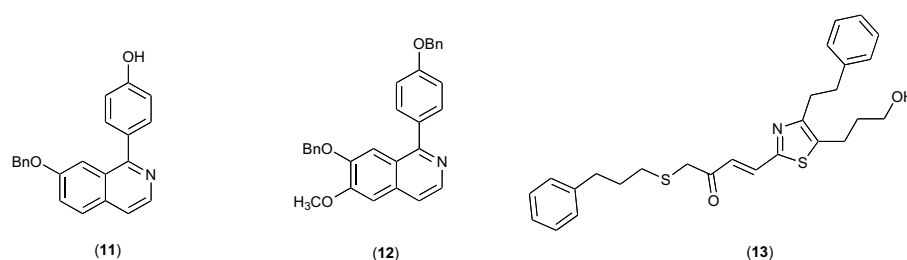


Fig. (9). Some potent non-peptidic inhibitors developed by structure-based drug design.

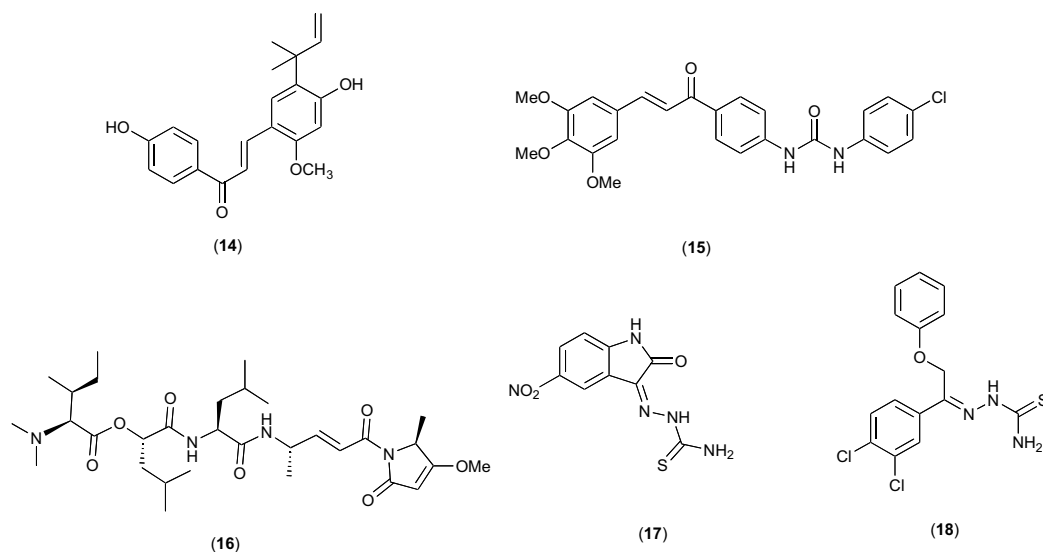


Fig. (10). Some natural product-based inhibitors of FP-2/3.

synthesized by O'Neill and co-workers [79]. Molecular modeling was used to compare the low energy conformations of the proposed peptidomimetics, by substituting leucine at P2 of vinyl sulfone with pyridine of the vinyl sulfone inhibitor K11017. The derived peptidomimetic showed *in vitro* anti-malarial inhibition, with IC_{50} values that vary from 6 to 40 μM . Once more, the best inhibitor of this series displayed hPhe at P1, having a phenyl vinyl sulfone electrophile. The presence of Phe at the P1 position showed lower FP-inhibitory activities, which confirmed the benefit on antiplasmodial activity of longer chains at P1. In terms of the

electrophilic groups, phenyl vinyl sulfones proved to be more potent inhibitors, with only minor noticeable differences between the activities of the aldehydes and α,β -unsaturated esters.

Further synthesis and biological activities of a new class of peptidomimetic FP-2 inhibitors have been evaluated by Ettari *et al.*, based on a 1,4-benzodiazepine scaffold and combining various α,β -unsaturated electrophilic groups, *e.g.* vinyl-ketones, -amides, -esters, and -nitriles with a vinyl ester moiety [100, 126, 127]. Some of these FP-2 inhibitors

exhibited antiplasmodial activities with IC₅₀ values in the low micromolar range. This is because the unsaturated ester function, which functions as a Michael acceptor, is capable of trapping the active site Cys42 thiol group through a covalent bond [128, 129]. However, information from the literature indicates that most of the reported FP-2 inhibitors are mainly derived from peptide analogues, with nanomolar IC₅₀ values. The latter form covalent interactions between the electrophilic groups, *e.g.* aldehydes, nitriles, vinyl sulfones and epoxides with thiolate of the catalytic cysteine amino acid [130, 131].

2.1.5. Coupling Chemical Synthesis and In Silico Design

During the last two decades, there have been several attempts that combine synthesis of derivatives of hit compounds with *in silico* screening approaches to develop and optimize new FP-2/3 inhibitors. In addition to the experimental *in vitro* structure-based screening of compound libraries, various computer-assisted drug design (CADD) approaches have been utilized in efforts towards identifying novel anti-malarial agents in general and FP-2/3 inhibitors in particular. CADD has directly or indirectly led to the identification of about 50 clinical candidates together with numerous drug approvals [132]. As the terminology suggests, CADD refers to the utilization of computer technology and software to underpin drug discovery efforts. Computational approaches often enhance the efficiency of drug discovery processes. This is by rendering the design of new drug candidates more rapid and cost-efficient, since nowadays CADD has the ability of predicting the biological activity of a potential drug even before it is synthesized and tested. Thus, CAAD, also known as computer-aided rational drug design strategies could help in developing hypotheses for quantitative structure-activity relationships (QSAR). Moreover, docking approaches have emerged as promising tools to predict the binding mode of a drug molecule in the target binding site. *In silico* techniques therefore help to support, guide, and streamline the processes of drug discovery, design, development, and optimization. We now focus our discussion on two *in silico* techniques; structure-based and ligand-based approaches. Structure-based drug design (SBDD) procedures can be applied (such as docking simulations) if the target is known from experimental (*e.g.* from X-ray crystallographic or NMR studies) or theoretical sources (*e.g.* by homology modeling). This approach relies on information retrieved from the 3D structure of the drug target, which is used in the search for new compounds that could bind to a particular target. The advantages of the structure- or target-based approaches include the carefully guided step by step rational search for the desirable steric and electronic complementarity between the target and designed ligands. Ligand-based drug design methods can be used (*e.g.* employing QSAR or 3D-QSAR models, pharmacophore models, *etc.*), which are rather based on the analysis of a number of ligands known to act with a common mechanism of action to a given receptor [133]. These approaches could be very useful in cases where the 3D crystal structure of the target is unknown. FP-2 and FP-3 have a good number of crystallized target structures, which have been used in CADD attempts.

3. INPUT FROM COMPUTATIONAL STUDIES

3.1. Overview of Computational Methods in Drug Development

Molecular mechanics (MM), molecular dynamics (MD) and quantum mechanics (QM) are the three main computer simulation techniques used to study protein-ligand interactions. The last mentioned approach is based on attempting to find the solutions to the Schrödinger equation. QM methods are often subcategorized into semi-empirical molecular orbital methods and *ab initio* methods. *Ab initio* is the most accurate QM method and calculates properties of a molecular system by solving the Schrödinger equation numerically from first principles. All electrons are treated explicitly and the interactions between atoms are determined solely from their electronic configurations and position of the atoms. Despite its accuracy, the method is quite computationally expensive, making it only realistic for small sized molecular systems and/or for very short time scales when using molecular dynamics. Semi-empirical methods make use of various simplifications or approximations (by only treating the valence electrons explicitly and by fitting several parameters from experimental data or previously computed parameters from *ab initio* calculations. This makes semi-empirical methods to be applicable to larger system sizes and longer time scales, when compared with *ab initio* methods. MM approaches are often used for simulations involving very large molecules, *e.g.* proteins and nucleic acids which typically contain several hundreds or thousands of atoms. In such simulations, the microscopic state of the system is investigated by using only the positions and forces acting on the atoms. Inter-atomic potentials (referred to as force fields) are used to model the interaction between atoms. Force field parameters are typically derived from either highly accurate *ab initio* calculations or from fitted to experimental data. When MM is applied to MD, the studied system can be investigated a time scale up to a microsecond.

3.2. Molecular Dynamics Simulations

MD can be used to validate the docking simulation and to analyze the dynamics action of the ligand within the binding pocket. In drug discovery, MD simulations have been widely employed to gain insight into receptor-drug interactions at an atomic level [134]. An important application of MD simulations in drug discovery is in the calculation of the binding energies corresponding to ligand-receptor interactions. These free energy calculations are carried out using the MM Poisson-Boltzmann surface area approach [135]. Examples of the application of MD simulations in anti-malarial CADD are presented in this subsection. Cieplak *et al.* [136] have performed MD studies on FPs-inhibitors complexes. These MD studies allowed the authors to calculate the low-energy states of the complexes over time and evaluate the binding modes of the inhibitors in these states, leading to the identification of three potential inhibitors (compounds **19-21**, Fig. 11).

Combined QM/MM approaches have also been carried out, by treating the reacting species (drug molecule) with a QM method and including the enzyme environment with MM [137, 138], *e.g.* Du *et al.* [139] have studied catalytic interactions between the octapeptide (AVLQSGFR) and the

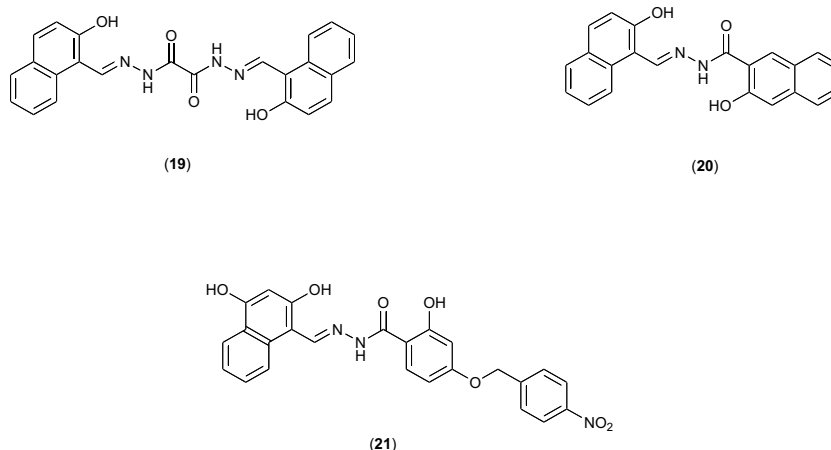


Fig. (11). Molecular structures used in MD studies on model complexes of FP inhibitors [138].

SARS 3CL^{pro} target, by using combined MM and QM. Arafet *et al.* [140] have worked with a FP-2- E64 complex to study the inhibition of FP-2 by E64, coupling QM/MM and MD simulation in order to obtain a complete picture of the key states of low free energy of this complex over time in order to understand the inhibition mechanism of FP-2 by E64. The authors observed that the irreversible attack of Cys42 on E64 could take place on both carbon atoms of the epoxy ring. It was also noted that certain residues, *e.g.* Gln171, Asp170, Gln36, Trp43, Asn81, and His174 could anchor the inhibitor in the right orientation to facilitate the reaction. Additionally, the MD simulations led to the detection of two different conformations of a product-like structure, deferring in their relative orientation of E64 within the active site of the protein:

- i). one conformation was similar to the initial X-ray structure (PDB entry: 3BPF), in which the guanidine group of E64 interacts with Tyr78, and
- ii). the other conformation shows the guanidine group of E64 interacting with Asp170.

Analysis of the results reveals that both structures were found to be stable and the last appears to be energetically favored. Thus, these observations may be useful for the rational design of new compounds with higher inhibitory activity against FP-2. Omotuyi and Hamada [141], have processed a FP-2-hemoglobin complex and determined the states of low energy of this complex *via* MM/QM MD simulation techniques. They observed that the interactions between FP-2 and hemoglobin are made between cys42, His174, Glu36, Asp-173/204 in subsites S1, S1', and S3 of FP-2, comprising residues 34 to 51 of hemoglobin. The authors have then determined the structure of this complex which precedes hydrolysis of hemoglobin and notes that in this conformation the hemoglobin links preferentially with the amino acids in the subsites S1 and S3 of FP-2, reducing his flexibility conformation and facilitating the formation of a stabilizing β -sheet between Leu172 and Val176 of FP-2 and the Phe45 and Asp47 of Hb. These observations show that the interaction of inhibitors with the amino acids of subsites S1 and S3 is a guide to optimize the discovery of FP-2 inhibitors and prevent degradation of the hemoglobin.

The mechanism could be explained by the fact that during Hb degradation, free heme is released and oxidized from the ferrous state (Fe^{2+}) to the ferric heme (Fe^{3+}) [142]. Both heme and hematin are known to be potentially toxic to the parasite [143]. In order to counter this, the parasite has developed a detoxification system which results in the formation of an inert crystalline hematin polymer (β -hemozoin pigment) [142, 144]. Docking and MD simulation techniques have been used by Marques *et al.* [145] in order to study the interactions between heme and porphyrin analogues with FP-2. This is based on the fact that fixation of heme with FP-2 prevents the presence of heme in the parasite active site. This promotes the polymerization of heme in the hemozoin which is not toxic to the parasite. Careful observation of the interactions of porphyrin analogs with FP-2 shows that iron has a low contribution in the interactions of the FP2-heme complex and revealing that it is important to have a propionic and ethenyl groups in structure of heme analogues for to prevent the formation of the complex FP-2-heme. The studies of docking and MD simulation unveiled a novel, inducible heme-binding moiety in FP-2, which is adjacent to the catalytic site. The general conclusion of their study is that binding of heme with FP-2 can limit the accumulation of free heme in the food vacuole of parasite, allowing polymerization of heme black pigment which is not toxic to the parasite and called hemozoin.

Saddala *et al.* [146] treated the crystal structure of FP-2 (PDB entry: 2GHU) by the methods of energy minimization and made a study by molecular dynamics using NAMD 2.9 software with the CHARMM like force field. Knowing that the drug mefloquine (compound **22**, Fig. 12) shows a high inhibition against FP-2, the authors made a virtual screening of 15560 compounds analogous of mefloquine from the PubChem database. These compounds were docked in the active site of FP-2 with the Autodock software, and five compounds (**23-27**, Fig. 12), which showed the lowest binding energies of -9.2, -9.1, -8.5, -8.1 and -7.1 kcal/mol, respectively were identified as inhibitors of FP-2. So these five compounds can serve as a starting point for the CADD more potent inhibitors against FP-2. Very recently Musyoka *et al.* [147] docked the South African natural compound library towards FP-2/3 and other nine cysteine proteases. These

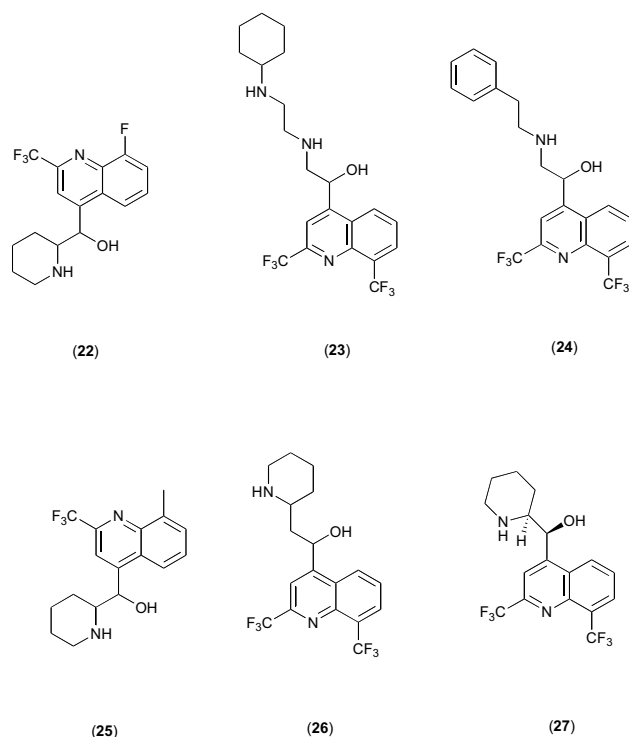


Fig. (12). Best binding affinity lead compounds.

docking studies led the authors to the identification of the compound 5PGA (compound **28**, Fig. **13**) as a potential hit. The authors further screened 186 compounds analogues of this potential hit from the ZINC database and docked these compounds into the active site of FP-2/3. Five hits (compound **29-33**, Fig. **13**) and 5PGA were selected for further analysis, based on the docking energy results (lowest energy conformations of the complex with FP-2/3). Thus these complexes can be look at as the starting structures for further MD simulations. Moreover, a careful study of these complexes also allowed the authors to determine the number of hydrogen bonds and their occupancy during the MD simulations. They have observed that the hydrogen bond occupancy was higher in all complexes, except in FP-2 when in complex with 5PGA. They find more than the hydrogen bonds formed between FP-2 and FP-3 with compounds called ZINC03869631 (compound **30**), ZINC04532950 (compound **31**) and ZINC05247724 (compound **33**) showed greater stability. The above observations can be useful in the search of dual potent inhibitors of FP-2/3.

3.3. Computational Structure-based Drug Design

3.3.1. Structure-based Virtual Screening (SBVS)

VS of large chemical libraries have proven to be a successful approach for lead identification. Different docking programs which employ various scoring functions to predict ligand binding to target sites have been employed to quantify different aspects of ligand binding to targets. SBVS approaches have by far been the most utilized methods in drug discovery efforts for FP-2 and FP-3 inhibitors. As an example, Ring *et al.* [148] and Schirmeister *et al.* [149], carried

out a study, which was based on the X-ray structures of papain and actinidin. This led to the design of a model structure for FP. This was later used as a receptor for ligand docking in order to search a library of about 55,000 commercially available small compounds. The 31 identified compounds chosen were tested against FP-2 inhibition. The most active compound (compound **19**, Fig. **11**) inhibited FP-2 with an IC_{50} value of 6 μ M. The study of the complex forming between FP-2 with this active compound after docking showed that the hydrophobic S2 subsite is occupied by one naphthol group, while the second naphthol group showed a π -stacking interaction with Trp177 of the S1' subsite. Additionally, it was observed that each hydroxyl group on the naphthol rings formed hydrogen-bonds with Ser160 at S2 and Gln19 at the S1' subsite. Based on these observations, Li *et al.* [150] then used the most active compound as a starting structure for various chemical modifications with the goal of improving the anti-malarial activity of this family of compounds. Joachimiak *et al.* [151] rather based their VS study on the X-ray structure of cathepsin K zymogen (PDB entry: 1BY8) for homology model generation for FP-2. This led to the identification of 8 non peptidic FP-2 inhibitors (compounds **34-41**, Fig. **14**), having IC_{50} range between 1 and 7 μ M in the enzyme assay. Three of these compounds (**36**, **37**, and **40**) were also active against the W2 (chloroquine resistant) strain, with IC_{50} range < 26 μ M. Thus, these compounds could be further developed into new FP-2 inhibitors. Desai and coworkers [152-155] screened the ChemBridge database against homology models of FP-2 and FP-3, leading to the identification of 12FP-2/3 dual inhibitors, amongst which compounds **42-48** (Fig. **14**), having IC_{50} values ranging between 1 and 63.4 μ M for FP-2, 4.9 and 62.2 μ M for FP-3.

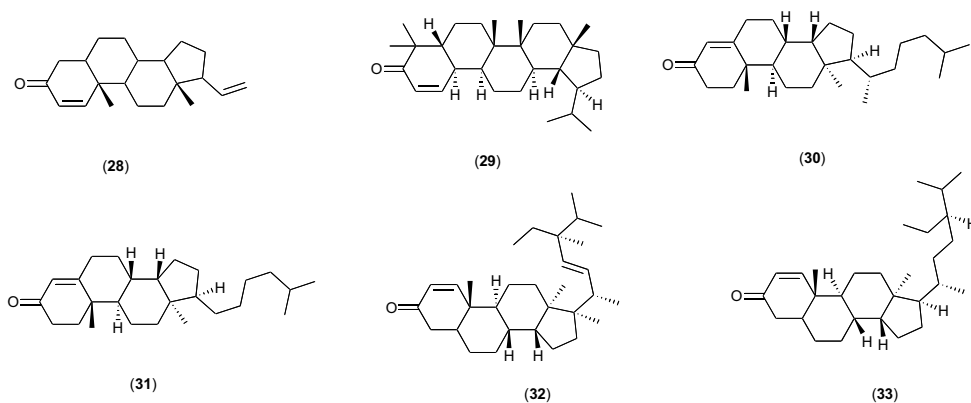


Fig. (13). Non-peptidic structures of some FP-2/3 inhibitors identified from the ZINC database and using molecular dynamics simulation.

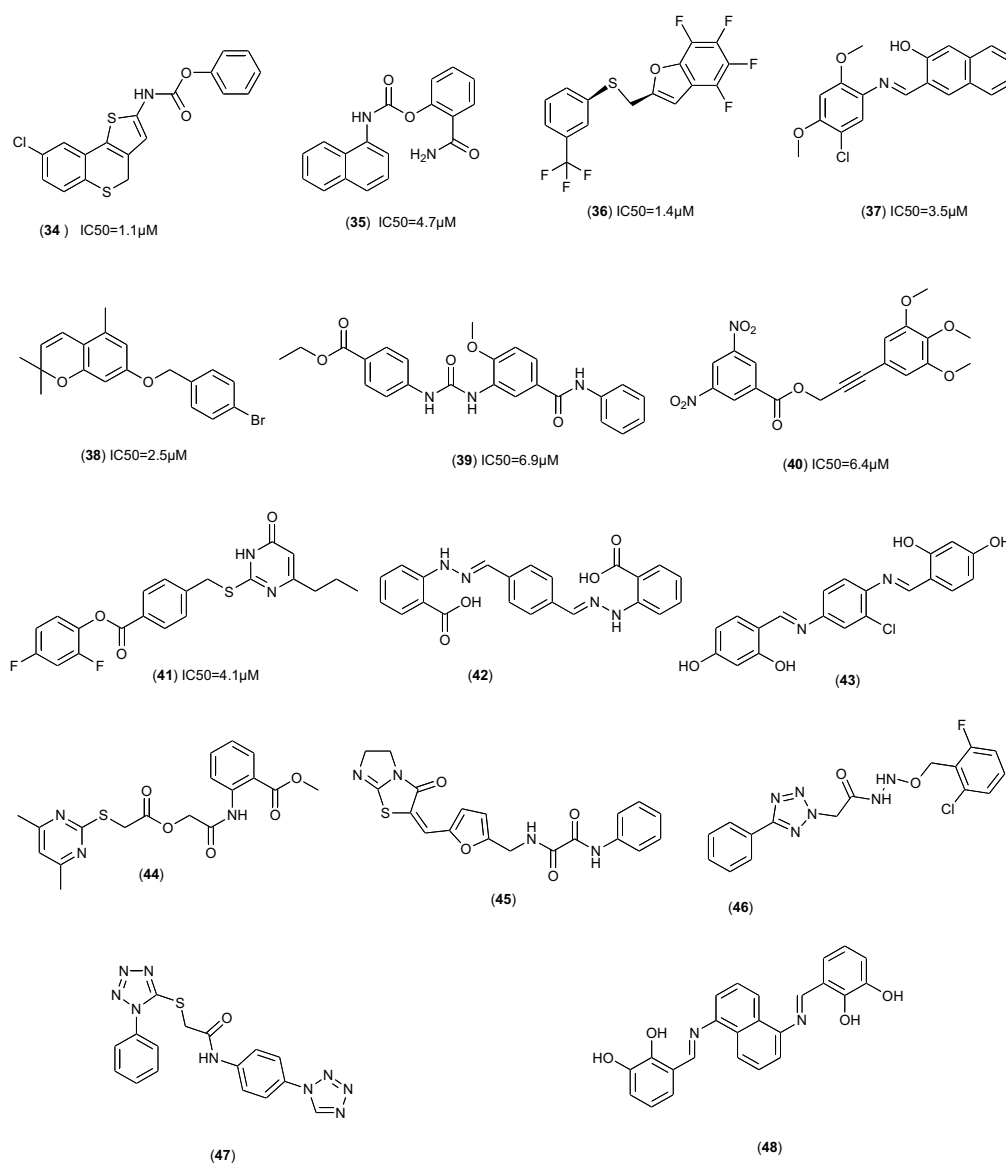


Fig. (14). Non-peptidic structures of some FP-2 inhibitors identified using virtual screening.

Compound **42** showed the highest dual activity, having an excellent profile against falcipains (FP-2 IC_{50} = 4.6 μ M, FP-3 IC_{50} = 5.8 μ M), the chloroquine-sensitive (D6) and chloroquine resistant (W2) strains of *P. falciparum*, (W2 IC_{50} = 6.6 μ M, D6 IC_{50} = 6.9 μ M). In the same order of ideas and in another work, Desai and colleagues suggest that compound **48**, along with other compounds (e.g. compounds **42** and **43**), give the possibility of developing broad spectrum anti-protozoal drugs [154]. Honglin *et al.* [156] tested 81 compounds retrieved after docking the SPECS database towards the crystal structure of *P. falciparum* (PDB entry: 2GHU) [157, 158]. The identified compounds exhibited IC_{50} values between 2.4 and 54.2 μ M, while the most active (compound **49**, Fig. 15), showed IC_{50} = 2.4 μ M against FP-2, prompting it to be used for further substructure searches and synthetic derivation. Shah *et al.* [157] have conducted a similar study by using SBVS on a focused cysteine protease inhibitor library leading to the identification of 50 compounds, among which 21 non-peptidic inhibitors were identified with IC_{50} values range from 1.4 to 49 μ M against FP-2. Four inhibitors of this series (**50-53**, Fig. 15) were found to be dual inhibitors of FP-2 and FP-3 (compound **50**, IC_{50} = 2.18 μ M against FP-2 and 4.95 μ M against FP-3); (compound **51**, IC_{50} = 7.51 μ M against FP-2 and 30.27 μ M against FP-3); (compound **52**, IC_{50} = 11.58 μ M against FP-2 and 45.64 μ M against FP-3); (compound **53**, IC_{50} = 20.87 μ M against FP-2 and 36.05 μ M against FP-3). These results suggest that these non-peptidic scaffolds could serve for further evaluation of structure activity relationships, which could help in the development of dual potent inhibitors of FP-2/3, eventually useful for anti-malarial chemotherapy. In another study Chintakrindi *et al.* [158] designed inhibitors of *P. falciparum* FP-3 by using the crystal structure of FP-3 in complex with its inhibitor K11017. The complex served to report the effect of substitution at the 7-amino position of the coumarin scaffold on the inhibition of FP-3 and a set of novel non-peptidic inhibitors were obtained by *de novo* design. The authors then have docked these compounds, predicted and validated the activities of designed inhibitors using various 3D-QSAR techniques. Thus, one compound which was prioritized for synthetic derivation has been identified and predicted as a potential FP-3 inhibitor (compound **54**, Fig. 15), although no biological data were experimentally obtained. Zhu *et al.* [159] designed and synthesized a series of novel small molecule FP-2 inhibitors based on the study of three optimizations areas of the lead (*R*)-2-phenoxy-carboxamido-3-(1*H*-indol-3-yl)-*N*-benzylpropanamide, which was identified based on SBVS. Among compounds of this series, four compounds (**55-58**, Table 1) exhibited FP-2 inhibitory activity, with IC_{50} values between 10.0 and 39.4 μ M, while the most active compound of this series (compound **57**) showed an inhibitory activity three-fold better than that of the parent and may prove useful for the development of potent FP-2 inhibitors. Ang *et al.* [8] screened a cathepsin inhibitor library against 5 parasite cysteine proteases, known to be involved in tropical parasitic diseases. These inhibitors have been grouped into chemotypes and the inhibitory activity of each chemotype was evaluated against FP-3 over time by the authors. It was observed that irreversible inhibitors such as fluoromethylketones and vinyl sulfones showed time-dependent inhibition while those reversible inhibitors such as nitriles showed no

time dependence. The authors have then tested all FP-3 inhibitors against FP-2. They identified a significant number of potent FP-2/3 inhibitors, where three ketobenzoxazole-based analogs (**59-61**, Table 2), identified as potent inhibitors of FP-2/3 and *P. falciparum* growth in culture. Thus, these compounds could be developed by computational approaches in order to become powerful drugs against malaria and away from resistant. Kesharwani *et al.* [11] have used the LigBuilder software and refined the compounds, based on Lipinski rule to design 104 anti-malarial compounds, analogs of leupeptin via structure-based drug designing approach. These compounds were docked in the active sites of the crystal structures of FP-2 and FP-3 (entry code 3BPF and 3BWK respectively). The docking results suggest that some of these compounds (e.g. compound **62**, **63**, Fig. 16) shows good binding with FP-2 than the native ligand E64. Moreover, 15 of the compounds have showed better binding affinity compared to the native inhibitor of FP-3 leupeptin. This series of compounds show the new leupeptin analogs that can be developed by the searchers involved in the discovery of new malaria by using computational approaches. Coteron *et al.* [75] began to carry out optimization studies with the aryl nitriles series following an empirical approach in which optimization of the initial hits was made by considering the interactions between the active sites of FP-2/3 and the P1-P3 positions of the hits (Fig. 4). Some of these compounds presented excellent *in vitro* activities against FP-2 or FP-3 and other are dual inhibitors of both enzymes (e.g. compounds **64-77**, Tables 3 and 4). Others compounds presented excellent *in vitro* enzymatic and anti-parasitic activities included compounds (**78-86**, Table 5). These optimizations studies have been carried out in same laboratory and in the same experimental conditions. Thereby, correlation studies between the structures and activities of these compounds can be conducted to predict the activities of the similar compounds which are not yet synthesized and can present a better profile as dual inhibitors of both enzymes. Musyoka *et al.* [160] have screened after docking scores, a diverse set of non-peptidic compounds (5-substituted-2-cyanopyrimidines derivatives, chalcones, isoquinolines and thiosemicarbazones) from the literature whose activities against FP-2 and/or FP-3 and their plasmodial homologs (targets) and the human papain-like cathepsins (off-targets) have been determined experimentally. The authors have then determined the mode of interaction of these non-peptidic compounds with FP-2 and FP-3, through by molecular docking and characterization between key amino acid residues and potential inhibitors.

3.3.2. General Observations

It was generally observed that:

- i). there are differences in amino acid compositions within the subsites of human cathepsins and plasmodial FP-2 and FP-3 homologs, important for inhibitor design;
- ii). Among the four non-peptidic compound classes studied, the 5-substituted-2-cyanopyrimidines derivatives optimized by Coteron *et al.* [75] are the best docked low binding free energy FP-2/3 inhibitors of *P. falciparum*, when compared with human cathepsins.

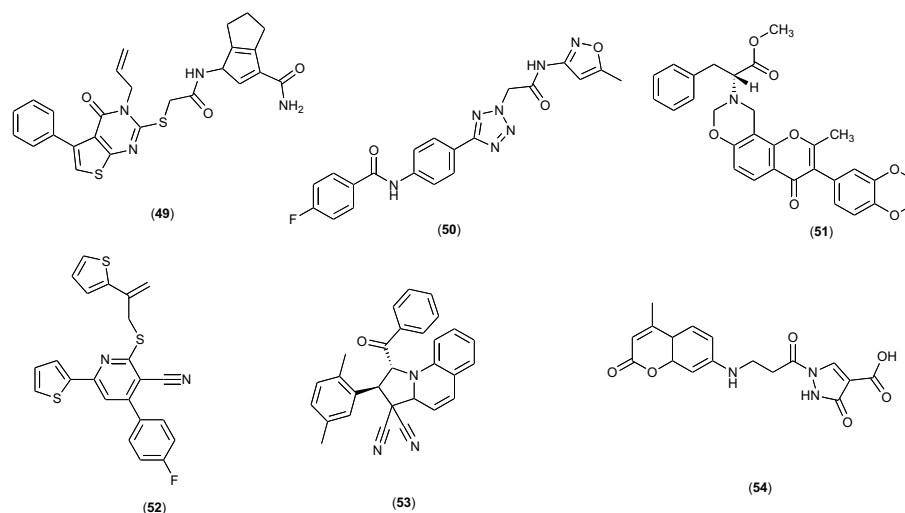
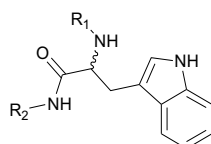


Fig. (15). The most active inhibitors of FP-2/3 identified from the SPECS database.

iii). the substituent chemical groups in the main 5-substituted-2-cyanopyrimidines derivatives scaffold determine the overall potency of the individual derivatives. Thus, the novel results obtained could be further exploited in SBVS and computer-aided drug design studies. Other docking studies with the FP-2/3 crystal structure were performed for the aldehyde and ketone based inhibitors [161].

Table 1. Chemical structures of compounds 55 to 58, along with their inhibitory activities against FP-2.



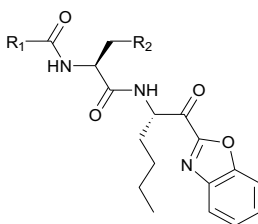
| Compound | R ₁ | R ₂ | FP-2 IC ₅₀ (μM) |
|----------|----------------|----------------|----------------------------|
| 55 | | | 29.7 |
| 56 | | | 39.4 |
| 57 | | | 10.0 |
| 58 | | | 13.3 |

3.4. Ligand-based Computational Drug Discovery Efforts

3.4.1. Two-dimensional (2D) and Three-dimensional (3D) QSAR Modeling

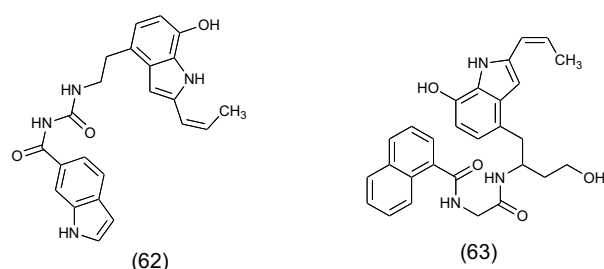
The QSAR is a valuable tool in drug design and medicinal chemistry for the development of new drug molecules.

This technique provides a complete set of the data, without necessity of performing expensive laboratory experiments [162]. Very recently, the structure of 3-phenylquinoxaline 1,4-di-N-oxide derivative has been used as template by Soni and Patidar [163] for QSAR studies and the classical Hansch approach was carried out to establish correlations between *P. falciparum* inhibitory activity and various substituent constants at position R₂, R₆, R₇ and W of molecule (Compound 87, Fig. 17). The statistical parameters of the QSAR model obtained showed a strong correlation between the structure and activity of these compounds. The model so obtained could be effectively used as a channeling tool for further structure modification and designing of some novel potent anti-malarial compounds. Moreover, the most popular QSAR techniques, *e.g.* comparative molecular field analysis (CoMFA) and comparative similarity indices analysis (CoMSIA) approaches incorporate ligand 3D information by searching for sites on the ligand which could be modified in order to obtain more specific ligands. This approach has been pursued by several research groups. One such study was reported by Potshangbam *et al.* [14], who performed CoMFA, CoMSIA, and docking studies of 2-pyrimidinecarbonitrile-based derivatives on FP-3 of *P. falciparum*. These results have provided important clues for the design of new compounds with high predicted FP-3 inhibitory activities. Similar studies have been reported by Ghasemi and Shiri [164]. More recently Wang *et al.* [165] have used the X-ray crystal structure of FP-3 (PDB entry: 3BPM) for docking studies of 247 2-pyrimidinecarbonitrile derivatives known as FP-3 inhibitors and have developed a set of ligand- and target-based 3D-QSAR models (CoMFA and CoMSIA) for these compounds, in order to understand the structure-activity correlation of FP-3 inhibitors. Finally, they submitted the models to MD simulation studies in order to elucidate the probable binding modes of the ligand in the FP-3 binding site. The models thus obtained by the authors; can help in designing and predicting new potent inhibitors of FP-3 in other work of search for new anti-malarial. Ehmke *et al.* [166], in order to optimize the use of triazine nitriles as proposed by Oballa *et al.* [167] as FP-2 inhibitors, designed a small library of drug-

Table 2. Ketobenzoxazole-based analogues: potent FP-2/3 inhibitors, and their inhibitory activities against FP-2/3.

| Compound | R ₁ | R ₂ | FP-2 IC ₅₀ (μM) | FP-3 IC ₅₀ (μM) | P. falc.W2 GI ₅₀ (μM) |
|----------|----------------|----------------|----------------------------|----------------------------|----------------------------------|
| 59 | Morpholine | i-Pr | 0.0021 | 0.0083 | 1.2 |
| 60 | Morpholine | Ph | 0.0044 | 0.0049 | 2.9 |
| 61 | Me | i-Pr | 0.066 | 0.017 | 4.8 |

like derivatives aided by computational modeling. Mundra and Mahesh [168] have used ligand-based methods to design and synthesis of 1,4 substituted piperazine derivatives as novel FP-2 inhibitors. These compounds have been tested *in vitro* by the authors for determine their activity against FP-2. A total of 17 novel compounds were retain, based on their activity. Three compounds (**88-90**, Table 6) of this series have shown good inhibitory activity. Furthermore, the authors performed docking studies of this compounds in order to get an idea of how this active analogues bind with the target. It was generally observed that the amino acid residues Trp206, Ile85, Leu84 and Val152 were most commonly involved in hydrophobic interactions, while Asn173, Cys42 and Gln6, were involved in hydrogen bonding interactions. Based on the interactions between the aforementioned residues, the most potent compound from this series (compound **88**) was selected to be used as a starting point for the design of new candidates, with the aim of optimizing the FP inhibitory potencies of this class of compounds.

**Fig. (16).** Some potent FP-2/3 inhibitors designed and subjected to virtual screening *via* docking simulation against FP-2 and FP-3 receptors.

3.4.2. Pharmacophore Modeling Approaches

The perception of a pharmacophore is often an essential step towards understanding the interaction between a ligand and its receptor. Pharmacophore modeling is a successful yet very diverse subfield of CADD [169]. The pharmacophore concept has been used widely in rational drug design. Pharmacophores are often used to represent molecules at 2D or

3D levels, by depicting schematically the most important elements for molecular recognition. This concept is also useful for modeling ADME-tox properties, predicting side effects, target identification and off-target predictions. Pharmacophores are often combined with molecular docking simulations in CADD in order to improve VS performance. The use of pharmacophore-based methods has also gained greater traction in anti-plasmodial drug discovery [170, 171]. Desai *et al.* [154], for example, attempted to identify the common pharmacophore features of diverse FP-2 hits. A four feature hypothesis, which correlated well with the ligand-target interactions observed in the docking studies, was selected and two hydrogen bond acceptors and two hydrophobic regions have been identified as key features. Based on this pharmacophore hypothesis, the authors were able to identify more than 60% of the FP-2 hits from the Chembridge database, although the model did not lead to the identification compounds exhibiting dual FP-2/3 inhibition. Energy-based pharmacophore modeling is a hybrid approach or ligand-/ structure-based technique which uses docking energy scores to find active components of ligands against the target. This approach has been used by Kesharwani *et al.* [11], starting with the X-ray structure of the FP-3 protein (PDB entry: 3BPM). The key interaction site for leupeptin at the active site of FP-3 was identified. Based on these observations, a pharmacophore model was suggested, leading to the design of novel leupeptin analogues by using two basic fragments of leupeptin and by defining growing points. Mundra and Mahesh used the previously cited approach to select some promising potential FP-2 inhibitors (with IC₅₀ value ranging from 1 to 10.9 μM) and developed a pharmacophore model of these FP-2 inhibitors.

The core features of the selected inhibitors include:

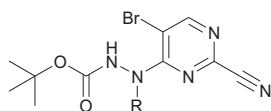
- i) a hydrophobic moiety; commonly an aromatic residue,
- ii) an aromatic moiety (monocyclic or bicyclic)
- iii) a linker; hydrogen bond donor and acceptor atom(s), attaching the hydrophobic moiety to the aromatic residue.

After examining these common features towards the design of a basic backbone as a FP-2 inhibitor, the authors built

a three component pharmacophore model, and a series of novel FP-2 inhibitors were designed and synthesized on the basis of this three component pharmacophore model. Chakka *et al.* [172], in order to identify a novel class of FP-2 inhibitors as potential anti-malarial agents, have used the structure of FP-2 (PDB entry: 1YVB) as a template for selected the pharmacophore features for *in silico* screening a library of commercially available compounds against the active site of FP-2. Initial hits were selected based on following criteria:

- i) at least two hydrophobic features in two of the three sub-sites S1', S1, and S2.
- ii) at least one additional feature, such as a hydrogen bond donor, hydrogen bond acceptor, cationic feature, or other hydrophobic features.

Table 3. Some aryl nitriles having different alicyclic substituents addressing the P2 pocket, with their effects on FP-2/3 activities.

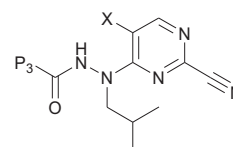


| Compound | R | FP-2 IC ₅₀ (μM) | FP-3 IC ₅₀ (μM) |
|----------|---|----------------------------|----------------------------|
| 64 | | 0.001 | 0.010 |
| 65 | | 0.001 | 0.008 |
| 66 | | 0.002 | 0.028 |
| 67 | | <0.0005 | 0.033 |

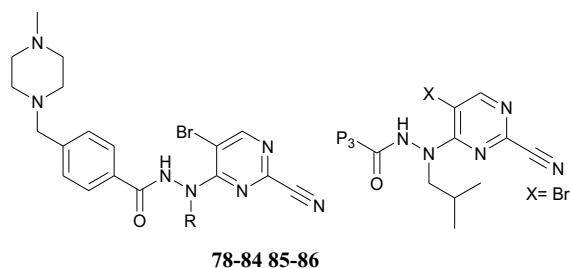
Based on this criteria, 2084 initial hits were obtained and subjected to Surflex-Dock™- based docking studies in the active site of FP-2 and 200 compounds, with their best pose based on the best-fit score, were selected. Finally, only seven compounds were acquired from commercial vendors and then the ability of these compounds to inhibit FP-2 activity *in vitro* was evaluated. Among them, compound (91, Fig. 18) exhibited moderate inhibition of FP-2, with a K_i of 27 ± 6 μM. Based on this molecule, a total of 23 compounds were design and evaluated in the protease assay by the authors where the most potent inhibitors from this enzymatic assay ($K_i < 10$ μM) were further evaluated in the parasite cultures as anti-malarial agents. The authors noted that some of these compounds presented excellent sub-micromolar inhibition against FP-2 with K_i values ranging from 0.2 ± 0.1 μM to 0.8

± 0.1 μM, and other inhibited FP-2 moderately, with K_i values ranging from 1 to 10 μM. Three compounds presented excellent inhibition of FP-2, and anti-parasitic (strain 3D7, Dd2) activities included compounds (92-94, Table 7). Thus, the results obtained suggest that this novel series of compounds should be further investigated as potential anti-malarial agents.

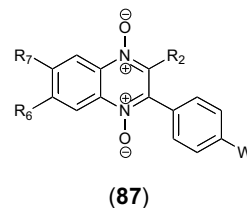
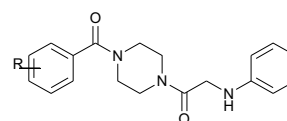
Table 4. Chemical structures of some potent aryl nitrile FP-2/3 inhibitors, along with their effects when different fused (hetero) aromatic ring substituents are placed at position P₃.



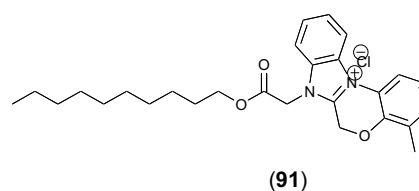
| Compound | X | P ₃ | FP-2 IC ₅₀ (μM) | FP-3 IC ₅₀ (μM) |
|----------|----|----------------|----------------------------|----------------------------|
| 68 | Br | | 0.0002 | 0.027 |
| 69 | Br | | 0.0002 | 0.040 |
| 70 | Br | | <0.0005 | 0.015 |
| 71 | Br | | <0.0005 | 0.012 |
| 72 | Br | | <0.0005 | 0.004 |
| 73 | Cl | | 0.0008 | 0.026 |
| 74 | Br | | 0.0015 | 0.006 |
| 75 | Cl | | 0.001 | 0.005 |
| 76 | Br | | 0.003 | 0.004 |
| 77 | Br | | 0.0005 | 0.008 |

Table 5. Some potent aryl nitrile FP-2/3 inhibitors and their *in vitro* anti-malarial activities (78-86).

| Compound | R | FP-2 IC ₅₀ (μM) | FP-3 IC ₅₀ (μM) | pfW ₂ IC ₅₀ (μM) |
|----------|---|----------------------------|----------------------------|--|
| 78 | | 0.0004 | 0.055 | 0.015 |
| 79 | | 0.0009 | 0.035 | 0.024 |
| 80 | | 0.0005 | 0.006 | 0.033 |
| 81 | | <0.0005 | 0.011 | 0.024 |
| 82 | | 0.0005 | 0.006 | 0.027 |
| 83 | | 0.0009 | 0.017 | 0.059 |
| 84 | | <0.0005 | 0.026 | 0.028 |
| 85 | | <0.0005 | 0.003 | 0.001 |
| 86 | | <0.0005 | 0.004 | 0.002 |

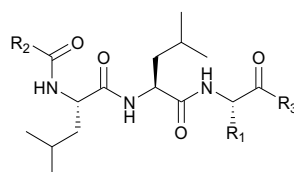
**Fig. (17).** Scaffold of 3-phenylquinoxaline 1,4-di-N-oxide derivatives.**Table 6.** Inhibition of falcipain-2 activity (IC₅₀) by compounds 88-90.

| Compound | R | FP-2 IC ₅₀ (μM) |
|----------|--|----------------------------|
| 88 | <i>p</i> -CH ₃ | 45.1 |
| 89 | <i>p</i> -CH ₂ CH ₃ | 48.7 |
| 90 | <i>o</i> -OCH ₂ CH ₃ | 72.2 |

**Fig. (18).** Structure of anti-FP active hit from *in silico* screening.

3.5. Integrated Structure and Ligand-based Approaches

Coupling structure and ligand-based methods is theoretically associated with greater efficiency in hit identification. In our literature review, two recent applications of this approach were encountered to identify *P. falciparum* FP-2 and FP-3 inhibitors. Shah *et al.* [77] performed combined ligand and SBVS, with ChemBridge and Asinex commercial databases to identify the non-peptidic inhibitors of FP-2 of *P. falciparum*. These inhibitors were the analogs of the hits discovered in their laboratory by virtual screening studies. Approximately, a total of 1500 hits were obtained and subsequently docked with the FP-2 crystal structure (PDB entry: 3BPF) using Glide XP protocol [157]. The top 300 molecules were selected based on the Emodel score. 69 compounds selected based on visual inspection criteria such as formation of hydrogen bonds by ligand atoms with key residues, ligand geometry, and commercial availability and submitted in biological test against FP-2. Finally 28 compounds showed inhibition of FP-2 in the low micromolar range with IC₅₀ values of 5 to 48 μM. Among them, some were active against cultured malaria parasites with IC₅₀ < 10 μM and the compound **95** (Fig. 19) was the most active, with an IC₅₀ of 4.64 μM. Mugumbate *et al.* [170]

Table 7. FP-2 inhibitory activity (K_i) and *P. falciparum* 3D7, Dd2 (IC_{50}) *in vitro* growth inhibition by compounds 92–94.

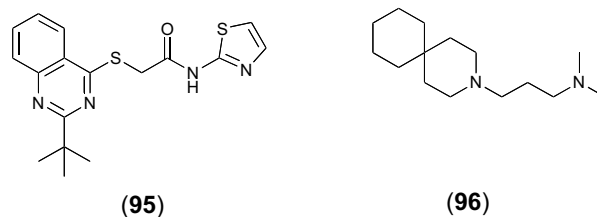
| Compound | R ₁ | R ₂ | R ₃ | FP-2 K_i (μ M) | pf3D7 IC_{50} (μ M) | pfDd2 IC_{50} (μ M) |
|----------|----------------|----------------|----------------|-----------------------|----------------------------|----------------------------|
| 92 | | | | 0.2 ± 0.1 | 5.5 ± 0.3 | 1.7 ± 0.1 |
| 93 | | | | 0.8 ± 0.1 | 5.2 ± 0.5 | 1.9 ± 0.3 |
| 94 | | | | 1.1 ± 0.1 | 0.9 ± 0.1 | 7.5 ± 0.3 |

screened more than 8.5 million compounds of the ZINC database using two different protocols involving structure and ligand-based methods. In first protocol, the Lipinski's "rule of five" was first used to screen the ZINC database for drug-like properties, and 1.1 million compounds were selected and clustered to avoid very similar compounds. The authors then subjected the structurally diverse of compounds to structure-based screening with the FP-2 crystal structure using the GOLD program and 26 compounds have been selected after visual inspection. In the second protocol, 25000 drug-like compounds were extracted from a drug-like subset using the ZINC search tool. The data set was further filtered on the basis of the properties of nine non-peptide inhibitors of FP-2 using the Open eye drug-like filter program. 100 compounds were selected on the basis of similarity scores and subjected to structure-based screening using Autodock. Finally 34 best ligands were selected on the basis of the docking score and visual inspection using such criteria as ligand adopted position and orientation in the binding site, formation of hydrogen bonds, and chemical diversity. Both protocols led to identified 60 compounds and purchased, despite that only 19 of them showed higher micromolar inhibitory activity against cultured chloroquine-resistant *P. falciparum* W2 strain. Compound **96** (Fig. 19) was one of the most active compounds, with an IC_{50} value of 1.32μ M against cultured parasites.

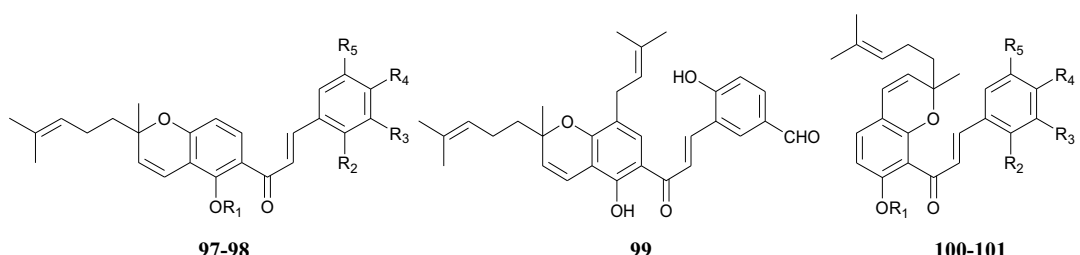
3.6. Natural Products

Many NPs and their derivatives are potential inhibitors of FP-2 [173]. The importance of incorporating modern drug discovery techniques into natural product drug discovery is

self-evident. These modern approaches include, among others, the adoption of computational drug discovery approaches. Whereas VS, molecular docking, and other computational techniques have predominantly been applied to collections of synthetic compounds, efforts have been made to leverage them to complement and enrich NP drug discovery. In what is an innovative application of established computational approaches to NPs, *in silico* tools were used to characterize the chemical space of a diverse group of NPs with previously reported *in vitro* activity against *P. falciparum* strains.

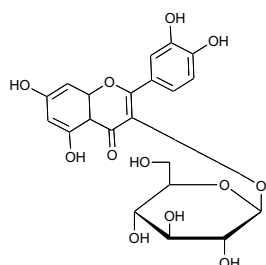
**Fig. (19).** Inhibitors of FP-2 and FP-3 obtained by integrated structure- and ligand-based approaches.

Liu *et al.* [174] designed and synthesized a series of 18 dihydroartemisinin derivatives based on a (thio) semicarbazone scaffold as potential FP-2 inhibitors. The authors performed molecular docking studies aimed at investigating the binding modes and interaction energies of the inhibitors. It was observed that this new class of dihydroartemisinin derivatives could be identified as potential FP-2 inhibitors. Preliminary SARs could serve as a foundation for further

Table 8. Inhibitory activities against FP-2 and *in vitro* anti-malarial activities of chromenochalcones 97-101.


| Compound | R ₁ | R ₂ | R ₃ | R ₄ | R ₅ | FP-2 IC ₅₀ (μM) | pf3D7(μM) |
|----------|----------------|----------------|----------------|----------------|----------------|----------------------------|-----------|
| 97 | H | OH | H | H | CHO | 13.5 | 1.19 |
| 98 | H | H | CHO | OH | H | 5.5 | 0.59 |
| 99 | - | - | - | - | - | ND | 0.25 |
| 100 | H | H | CHO | OH | H | ND | 0.39 |
| 101 | H | OH | H | H | CHO | 4.6 | 0.31 |

investigation of anti-malarial drugs. Tadigoppula *et al.* [17] synthesized a library of 88 chalcones with diverse structural features, *e.g.* prenylated chalcones, chromenochalcones, and chromenodihydrochalcones, based on the NP lead molecule (lipochalcone) isolated from the roots of Chinese licorice and evaluated for their *in vitro* anti-malarial activity. Among these compounds, **97-101** (Table 8) exhibited promising *in vitro* anti-malarial activities against the 3D7 strains of *P. falciparum*. Docking studies showed the possible binding orientation of these inhibitors in active sites of the FP-2 enzyme. Compounds **97**, **98**, and **101** showed modest inhibitory activity against the major Hb-degrading cysteine protease FP-2 and can be used for the development of new potent FP-2 inhibitors. Wang *et al.* [173] identified 10 NPs as *P. falciparum* FP-2 inhibitors from an in house NPs database using SBVS. The identified compounds showed moderate inhibitory activities against FP-2 with IC₅₀ values in the range of 3.18 to 68.19 μM. Compound **102** (Fig. 20) IC₅₀ = 5.54 μM against FP-2, exhibited *in vitro* anti-plasmodial activities against a chloroquine-sensitive strain (*P. falciparum* 3D7; IC₅₀ = 5.54 μM) and a chloroquine-resistant strain (*P. falciparum* Dd2; IC₅₀ = 4.05 μM).

**(102)****Fig. (20).** A natural product *P. falciparum* FP-2 inhibitor.

CONCLUSION

We have attempted to discuss and summarize recent efforts towards the design of malarial cysteine protease inhibitors. It was shown that combining computational techniques with organic synthesis was effective in delivering novel potential inhibitors and providing insight for further improvement of drug candidates. Despite significant efforts from international collaborations and local efforts aimed at controlling the malaria burden, enormous challenges are still being faced by researchers. These include the challenge of developing effective vaccine candidates, together with the drawbacks associated with the use of existing therapies. The emergence and spread of resistance against insecticides, and the available drugs remain another set of challenges. In this review, we have attempted to summarize the main findings of several research teams involved in searching for potent cysteine protease FP-2/3 inhibitors against *P. falciparum*. It appears that the current discussion on efforts towards the design of *P. falciparum* cysteine proteases inhibitors drives us towards combining computational techniques with organic synthesis and using NPs as lead scaffolds for delivering novel potential inhibitors and providing directions for further improvements. Thus, computer-based methods like docking, binding-free energy calculations, virtual screening, and pharmacophore modeling, often used for hit/lead identification along with combinatorial design of novel inhibitors against *P. falciparum* FP-2/3 could be employed. These therefore constitute available techniques for the development of new drugs against malaria.

LIST OF ABBREVIATIONS

| | | |
|----------|---|--|
| ACT | = | Artemisinin combination therapy |
| ADME-tox | = | Absorption, distribution, metabolism, excretion and toxicity |
| AIDS | = | Acquired immune-deficiency syndrome |

| | | |
|------------------|---|--|
| CADD | = | Computer-aided drug discovery |
| CQ | = | Chloroquine |
| FP | = | Falcipain |
| Hb | = | Hemoglobin |
| HIV | = | Human immunodeficiency virus |
| MD | = | Molecular dynamics |
| MM | = | Molecular mechanics |
| NP | = | Natural product |
| PDB | = | Protein databank |
| QM | = | Quantum mechanics |
| QSAR | = | Quantitative structure-activity relationship |
| VS | = | Virtual screening |
| SBD | = | Structure-based drug design |
| SBVS | = | Structure-based virtual screening |
| WHO | = | World health organization |
| List of symbols | | |
| IC ₅₀ | = | Concentration of a competitive tight binding inhibitor that causes 50% reduction of the rate of catalytic substrate conversion |
| K _i | = | Inhibitory constant |
| μM | = | Micromolar concentration (1 μM = 10 ⁻⁶ mol dm ⁻³) |
| nM | = | Nanomolar concentration (1 nM = 10 ⁻⁹ mol dm ⁻³) |
| ns | = | Nanosecond (1 ns = 10 ⁻⁹ s) |
| ps | = | Picosecond (1 ps = 10 ⁻¹² s) |

CONSENT FOR PUBLICATION

All authors agree to the contents of the submitted manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

FUNDING

FNK is currently a Georg Forster fellow of the Alexander von Humboldt Foundation, Germany.

AUTHORS' CONTRIBUTIONS

LCOO and EM conceived the idea. FNK and BDB participated in the data collection and analysis and developed the structure of the manuscript. BDB wrote the first draft of

the paper under the close supervision of FNK. All authors read and approved the final version of the manuscript before submission.

REFERENCES

- [1] Laveran A. Note sur un nouveau parasite trouvé dans le sang de plusieurs malades atteints de fièvre palustre. *Bull Acad Nat Med* 1888; 9: 1234-5.
- [2] Ettari R, Bova F, Zappala M, Grasso S, Micale N. Falcipain-2 inhibitors. *Med Res Rev* 2010; 30: 136-67.
- [3] Faraz Mojab. Antimalarial natural products: a review. *Avicenna J Phytomed* 2012; 2: 52-62.
- [4] Objectifs du Millénaire pour le développement http://www.undp.org/content/undp/fr/home/mdgoverview/mdg_goals/mdg6. Accessed on 25 Nov 2016
- [5] World Malaria Report. World Health Organization, Geneva, Switzerland, 2015.
- [6] MacNeil A, Rollin PE. Ebola and Marburg hemorrhagic fevers: neglected tropical diseases. *PLoS Negl Trop Dis* 2012; 6(6): e1546.
- [7] Utzinger J. A research and development agenda for the control and elimination of human helminthiasis. *PLoS Negl Trop Dis* 2012; 6(4): e1646.
- [8] Ang KKH, Ratnam J, Gut J, et al. Mining a cathepsin inhibitor library for new antiparasitic drug leads. *PLoS Negl Trop Dis* 2011; 5(5): e1023.
- [9] O'Neill HG, Mzilahowa T, de Deus N, Njenga SM, Mmbaga EJ, Kariuki TM. Evaluation of the European foundation initiative into African research in neglected tropical diseases by the African fellows. *PLoS Negl Trop Dis* 2013; 7(3): e2019.
- [10] Nethavhani SA, van Ree T. Synthesis and antimalarial activities of some novel 2-pyridones. *Arab J Sci Eng* 2014; 39: 6595-6598.
- [11] Kesharwani RK, Singh DV, Misra K. Computation-based virtual screening for designing novel antimalarial drugs by targeting falcipain-III: A structure-based drug designing approach. *J Vector Borne Dis* 2013; 50: 93-102.
- [12] Rosenthal PJ. Proteases of malaria parasites: new targets for chemotherapy. *Emerg Infect Dis* 1998; 4: 49-57.
- [13] Ramjee MK, Flinn NS, Pemberton TP, Quibell M, Wang Y, Watts JP. Substrate mapping and inhibitor profiling of falcipain-2, falcipain-3 and berghepain-2: implications for peptidase anti-malarial drug discovery. *Biochem J* 2006; 399: 47-57.
- [14] Potshangbam AM, Tanneeru K, Reddy BM, Guruprasad L. 3D-QSAR and molecular docking studies of 2-pyrimidinecarbonitrile derivatives as inhibitors against falcipain-3. *Bioorg Med Chem Lett* 2011; 2: 7219-23.
- [15] Kamkumo RG, Ngoutane AM, Tchokouaha LR, et al. () Compounds from *Sorindeia juglandifolia* (Anacardiaceae) exhibit potent anti-plasmodial activities *in vitro* and *in vivo*. *Malar J* 2012 Nov 21; 11: 382.
- [16] World Malaria Report. World Health Organization, Geneva, Switzerland, 2008.
- [17] Tadigoppula N, Korthikunta V, Gupta S, et al. Synthesis and insight into the structure-activity relationships of chalcones as anti-malarial agents. *J Med Chem* 2013; 56: 31-45.
- [18] Pandey KC, Wang SX, Sijwali PS, Lau AL, McKerrow JH, Rosenthal PJ. The *Plasmodium falciparum* cysteine protease falcipain-2 captures its substrate, hemoglobin, via a unique motif. *Proc Natl Acad Sci USA* 2011; 102: 9138-43.
- [19] Thomas MB, Godfray HCJ, Read AF, et al. Lessons from agriculture for the sustainable management of malaria vectors. *PLoS Med* 2012; 9(7): e1001262.
- [20] Chanda E, Hemingway J, Kleinschmidt I, et al. Insecticide resistance and the future of malaria control in Zambia. *PLoS ONE* 2011; 6: e24336.
- [21] Mott BT, Tripathi A, Siegler MA, Moore CD, Sullivan DJ. Synthesis and antimalarial efficacy of two carbon-linked, artemisinin-derived trioxane dimers in combination with known antimalarial drugs. *J Med Chem* 2013; 56: 2630-41.
- [22] Uhlemann AC, Cameron A, Eckstein L, et al. A single amino acid residue can determine the sensitivity of SERCAs to artemisinins. *Nat Struct Mol Biol* 2005; 12: 628-29.
- [23] Dali B, Keita M, Megnassan E, Frecer V, Mierts S. Insight into selectivity of peptidomimetic inhibitors with modified statine core

- for plasmepsin II of *Plasmodium falciparum* over human Cathepsin D. *Chem Biol Drug Des* 2012; 79: 411-30.
- [24] Wongsrichanalai C, Steven RM. Declining artesunate mefloquine efficacy against *falciparum* malaria on the Cambodia-Thailand border. *Emerg Infect Dis* 2008; 14: 716-9.
- [25] Teixeira C, Gomes JRB, Couesnon T, Gomes P. Molecular docking and 3D-quantitative structure activity relationship analyses of peptidyl vinyl sulfones: *Plasmodium falciparum* cysteine proteases inhibitors. *J Comput Aided Mol Des* 2011; 25: 763-75.
- [26] Coombs GH, Goldberg DE, Klemba M, Berry C, Kay J, Mottram JC. Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends Parasitol* 2001; 17: 532-37.
- [27] Gil LA, Valiente PA, Pascutti PG, Pons T. Computational perspectives into plasmepsins structure-function relationship: implications to inhibitors design. *J Trop Med Article* 2011; 657483: 657483.
- [28] Tyagi C, Grover S, Dhanjal JK, Goyal S, Goyal M, Grover A. Mechanistic insights into mode of action of novel natural cathepsin L inhibitors. *BMC Genomics* 2013; 14(Suppl 8): S10.
- [29] Rawlings ND, Morton FR, Barrett AJ. MEROPS: the peptidase database. *Nucleic Acids Res* 2006; 34: D270-72.
- [30] Leung-Toung R, Zhao Y, Li W, Tam TF, Karimian K, Spino M. Thiol proteases: inhibitors and potential therapeutic targets. *Curr Med Chem* 2006; 13: 547-81.
- [31] Marco M, Miguel CJ. Falcipain inhibition as a promising antimalarial target. *Curr Top Med Chem* 2012; 12: 408-44.
- [32] Berman HM, Westbrook J, Feng Z, et al. The protein data bank. *Nucl Acids Res* 2000; 28 (1): 235-42.
- [33] Rosenberg R. Malaria: some considerations regarding parasite productivity. *Trends Parasitol* 2008; 24: 487-91.
- [34] Kappe SHI, Duffy PE. Malaria liver stage culture: *in vitro* veritas? *Am J Trop Med Hyg* 2006; 74: 706-7.
- [35] Kuehn A, Pradel G. The coming-out of malaria gametocytes. *J Biomed Biotechnol* 2010; 2010: 976827.
- [36] Stone WJR, Eldering M, Gemert GJ, et al. The relevance and applicability of oocyst prevalence as a read-out for mosquito feeding assays. *Sci Rep* 2013 Dec 4; 3: 3418.
- [37] Kappe SH, Gardner MJ, Brown SM, et al. Exploring the transcriptome of the malaria sporozoite stage. *Proc Natl Acad Sci USA* 2001; 98: 9895-900.
- [38] Reilly HB, Wang H, Steuter JA, Marx AM, Ferdig MT. Quantitative dissection of clone-specific growth rates in cultured malaria parasites. *Int J Parasitol* 2007; 37: 1599-607.
- [39] Dhangadamajhi G, Kar SK, Ranjit M. The survival strategies of malaria parasite in the red blood cell and host cell polymorphisms. *Malar Res Treatment* 2010; 2010: 973094.
- [40] Ranford-Cartwright LC, Mwangi JM. Analysis of malaria parasite phenotypes using experimental genetic crosses of *Plasmodium falciparum*. *Int J Parasitol* 2012; 42: 529-34.
- [41] Michael EAD, McIntosh T, Dean H, et al. Four distinct pathways of hemoglobin uptake in the malaria parasite *Plasmodium falciparum*. *Proc Natl Acad Sci USA* 2008; 105: 2463-8.
- [42] Miller LH, Su X. Artemisinin: discovery from the Chinese herbal garden. *Cell* 2011; 146: 855-863.
- [43] Francis SE, Sullivan DJ, Goldberg DE. Haemoglobin metabolism in the malaria parasite *Plasmodium falciparum*. *Annu Rev Microbiol* 1997; 51: 97-123.
- [44] Liu J, Gluzman IY, Drew ME, Goldberg DE. The role of *Plasmodium falciparum* food vacuole plasmepsins. *J Biol Chem* 2005; 280: 1432-7.
- [45] Blackman MJ. Proteases involved in erythrocyte invasion by the malaria parasite: function and potential as chemotherapeutic targets. *Curr Drug Targets* 2000; 1: 59-83.
- [46] Rosenthal PJ. Protease Inhibitors. In *Antimalarial chemotherapy: mechanisms of action, resistance, and new directions in drug discovery*; Rosenthal PJ (Ed.). Springer: New York, 2001; pp. 325-46.
- [47] Guruprasad L, Tanneeru K, Guruprasad K. Structural rationale for the recognition of arginine at P₃ in PEXEL motif containing proteins of *Plasmodium falciparum* by plasmepsin V. *Protein Pept Lett* 2011; 18: 634-41.
- [48] Eggleston KK, Duffin KL, Goldberg DE. Identification and characterization of falcylisin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*. *J Biol Chem* 1999; 274: 32411-7.
- [49] Klemba M, Gluzman I, Goldberg DE. A *Plasmodium falciparum* dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation. *J Biol Chem* 2004; 279: 43000-7.
- [50] Teixeira C, Gomes JRB, Gomes P. Falcipains, *Plasmodium falciparum* cysteine proteases as key drug targets against malaria. *Curr Med Chem* 2011; 18: 1555-72.
- [51] Mishra PS, Sharma H, Mishra R, Gupta SK. A review on antimalarial drug discovery and its screening method. *World J Pharm Pharmaceut Sci* 2014; 3: 1288-304.
- [52] Phillipson JD, O'Neill MJ. Antimalarial and amoebicidal natural products. In: Hostettmann K, Lea PJ (Eds.) *Biologically Active Natural Products*. 1987; pp. 49-64. Oxford, UK, Clarendon Press.
- [53] O'Neill PM, Barton VE, Ward SA, Chadwick J. 4-Aminoquinolines: chloroquine, amodiaquine and next-generation analogues. In: *Treatment and Prevention of Malaria, Part of the series Milestones in Drug Therapy* 2012; pp. 19-44 Springer, Basel, Switzerland.
- [54] Klein EY. Antimalarial drug resistance: a review of the biology and strategies to delay emergence and spread. *Int J Antimicrob Agents* 2013; 41: 311-7.
- [55] Subramanian S, Hardt M, Choe Y, et al. Hemoglobin cleavage site-specificity of the *Plasmodium falciparum* cysteine proteases falcipain-2 and falcipain-3. *PLoS ONE* 2009; 4(4): e5156
- [56] Padmanaban G, Nagaraj VA, Rangarajan PN. Drugs and drug targets against malaria. *Curr Sci* 2007; 92: 1545-55.
- [57] Barrett AJ, Rawlings ND. Evolutionary lines of cysteine peptidases. *Biol Chem* 2001; 382: 727-33.
- [58] Eksi S, Czesny B, Greenbaum DC, Bogyo M, Williamson KC. Targeted disruption of *Plasmodium falciparum* cysteine protease, falcipain 1, reduces oocyst production, not erythrocytic stage growth. *Mol Microbiol* 2004; 53: 243-50.
- [59] Greenbaum DC, Baruch A, Grainger M, et al. A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* 2002; 298: 2002-6.
- [60] Sijwali PS, Kato K, Seydel KB, Gut J, Lehman J, et al. *Plasmodium falciparum* cysteine protease falcipain-1 is not essential in erythrocytic stage malaria parasites. *Proc Natl Acad Sci USA* 2004; 101: 8721-6.
- [61] Goh LL, Sim TS. Characterization of amino acid variation at strategic positions in parasite and human proteases for selective inhibition of falcipains in *Plasmodium falciparum*. *Biochem Biophys Res Commun* 2005; 335: 762-70.
- [62] Wang WM, Ge G, Lim NH, Nagase H, Greenspan DS. TIMP-3 inhibits the procollagen N-proteinase ADAMTS-2. *Biochem J* 2006; 398: 515-9.
- [63] Korde R, Bhardwaj A, Singh R, et al. Prodomain peptide of *Plasmodium falciparum* cysteine protease (falcipain-2) inhibits malaria parasite development. *J Med Chem* 2008; 51: 3116-23.
- [64] Pandey KC, Wang SX, Sijwali P S, Lau AL, McKerrow J H, Rosenthal PJ. The *Plasmodium falciparum* cysteine protease falcipain-2 captures its substrate, hemoglobin, via a unique motif. *Proc Natl Acad Sci USA* 2005; 102: 9138-43.
- [65] Zhu J, Chen T, Liu J, Ma R, Lu W, et al. 2-(3,4-Dihydro-4-oxothieno[2,3-d]pyrimidin-2-ylthio) acetamides as a new class of falcipain-2 inhibitors. 3. Design, synthesis and biological evaluation. *Molecules* 2009; 14: 785-97.
- [66] Rosenthal PJ, Olson JE, Lee GK, Palmer JT, Klaus JL, Rasnick D. Antimalarial effects of vinyl sulfone cysteine proteinase inhibitors. *Antimicrob Agents Chemother* 1996; 40: 1600-3.
- [67] Sajid M, McKerrow J. Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol* 2002; 120: 1-21.
- [68] Drew ME, Banerjee R, Uffman EW, Gilbertson S, Rosenthal PJ, Goldberg DE. *Plasmodium* food vacuole plasmepsins are activated by falcipains. *J Biol Chem* 2008; 283: 12870-6.
- [69] Singh N, Sijwali PS, Pandey KC, Rosenthal PJ. *Plasmodium falciparum*: biochemical characterization of the cysteine protease falcipain-2'. *Exp Parasitol* 2006; 112: 187-92.
- [70] Jeong JJ, Kumar A, Hanada T, et al. Cloning and characterization of *Plasmodium falciparum* cysteine protease, falcipain-2B. *Blood Cells Mol Dis* 2006; 36: 429-35.
- [71] Pandey KC, Dixit R. Structure-function of falcipains: malarial cysteine proteases. *J Trop Med* 2012; 2012: 345195.
- [72] Rosenthal PJ. Cysteine proteases of malaria parasites. *Int J Parasitol* 2004; 34: 1489-99.
- [73] Sijwali PS, Koo J, Singh N, Rosenthal PJ. Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of *Plasmodium falciparum*. *Mol Biochem Parasitol* 2006; 150: 96-106.

- [74] Sijwali PS, Shenai BR, Gut J, Singh A, Rosenthal PJ. Expression and characterization of the *Plasmodium falciparum* haemoglobinase falcipain-3. *Biochem J* 2001; 360: 481-9.
- [75] Coter JM, Catterick D, Castro J, et al. Falcipain inhibitors: optimization studies of the 2-pyrimidinecarbonitrile lead series. *J Med Chem* 2010; 53: 6129-52.
- [76] Sijwali P S, Rosenthal PJ. Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*. *Proc Natl Acad Sci USA* 2004; 101: 4384-9.
- [77] Shah F, Gut J, Legac J, et al. Computer-aided drug design of falcipain inhibitors: virtual screening, structure-activity relationships, hydration site thermodynamics, and reactivity analysis. *J Chem Inf Model* 2012; 52: 696-710.
- [78] Sabnis YA, Desai PV, Rosenthal PJ, Avery MA. Probing the structure of falcipain-3, a cysteine protease from *Plasmodium falciparum*: comparative protein modeling and docking studies. *Protein Sci* 2003; 12: 501-9.
- [79] Verissimo E, Berry N, Gibbons P, et al. Design and synthesis of novel 2-pyridone peptidomimetic falcipain 2/3 inhibitors. *Bioorg Med Chem Lett* 2008; 18: 4210-4.
- [80] Xu J, Hagler A. Cheminformatics and drug discovery. *Molecules* 2002; 7: 566-600.
- [81] Kerr ID, Lee JH, Pandey KC, et al. Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. *J Med Chem* 2009; 52: 852-7.
- [82] Shenai BR, Sijwali PS, Singh A, Rosenthal PJ. Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*. *J Biol Chem* 2000; 275: 29000-29010.
- [83] Sijwali PS, Koo J, Singh N, Rosenthal PJ. Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of *Plasmodium falciparum*. *Mol Biochem Parasitol* 2006; 150: 96-106.
- [84] Hogg T, Nagarajan K, Herzberg S, et al. Structural and functional characterization of falcipain-2, a hemoglobinase from the malarial parasite *Plasmodium falciparum*. *J Biol Chem* 2006; 281: 25425-37.
- [85] Wang SX, Pandey KC, Somoza JR, et al. Structural basis for unique mechanisms of folding and hemoglobin binding by a malarial protease. *Proc Natl Acad Sci USA* 2006; 103: 11503-8.
- [86] Hansen G, Heitmann A, Witt T, et al. Structural basis for the regulation of cysteine-protease activity by a new class of protease inhibitors in *Plasmodium*. *Structure* 2011; 19: 919-29.
- [87] Kerr ID, Lee JH, Farady CJ, et al. Vinyl sulfones as antiparasitic agents and a structural basis for drug design. *J Biol Chem* 2009; 284: 25697-703.
- [88] Powers JC, Asgian JL, Ekici OD, et al. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev* 2002; 102: 4639-750.
- [89] Vicik R, Busemann M, Baumann K, Schirmeister T. Inhibitors of cysteine proteases. *Curr Top Med Chem* 2006; 6: 331-53.
- [90] Lee BJ, Singh A, Chiang P, et al. Antimalarial activities of novel synthetic cysteine protease inhibitors. *Antimicrob Agents Chemother* 2003; 47: 3810-4.
- [91] Schulz F, Gelhaus C, Degel B, et al. Screening of protease inhibitors as antiplasmodial agents. Part I: Aziridines and epoxides. *Chem Med Chem* 2007; 2: 1214-24.
- [92] Martichonok V, Plouffe C, Storer AC, Menard R, Jones JB. Aziridine analogs of [[*trans*-(epoxysuccinyl)-L-leucyl]amino]-4-guanidinobutane (E-64) as inhibitors of cysteine proteases. *J Med Chem* 1995; 38: 3078-85.
- [93] Shah F, Mukherjee P, Desai P, Avery M. Computational approaches for the discovery of cysteine protease inhibitors against malaria and SARS. *Curr Comput Aided Drug Des* 2010; 6: 1-23.
- [94] Pandey KC. Macromolecular inhibitors of malarial cysteine proteases - an invited review. *J Biomed Sci Eng* 2013; 6: 885-95.
- [95] Dominguez JN, Lopez S, Charris J, et al. Synthesis and antimalarial effects of phenothiazine inhibitors of a *Plasmodium falciparum* cysteine protease. *J Med Chem* 1997; 40: 2726-32.
- [96] Raether W, Enders B, Hofmann J, et al. Antimalarial activity of new floxacrine-related acridinedione derivatives: studies on blood schizontocidal action of potential candidates against *P. berghei* in mice and *P. falciparum* *in vivo* and *in vitro*. *Parasitol Res* 1989; 75: 619-26.
- [97] Greenbaum DC, Mackey Z, Hansell E, et al. Synthesis and structure-activity relationships of parasitocidal thiosemicarbazone cysteine protease inhibitors against *Plasmodium falciparum*, *Trypanosoma brucei* and *Trypanosoma cruzi*. *J Med Chem* 2004; 47: 3212-9.
- [98] Chiyanzu I, Hansell E, Gut J, Rosenthal PJ, McKerrow JH, Chibale K. Synthesis and evaluation of isatins and thiosemicarbazone derivatives against cruzain, falcipain-2 and rhodesain. *Bioorg Med Chem Lett* 2003; 13: 3527-30.
- [99] Hans RH, Wiid IJF, van Helden PD, et al. Novel thiolactone-isatin hybrids as potential antimalarial and antitubercular agents. *Bioorg Med Chem Lett* 2011; 21: 2055-8.
- [100] Ettari R, Micale N, Schirmeister T, et al. Novel peptidomimetics containing a vinyl ester moiety as highly potent and selective falcipain-2 inhibitors. *J Med Chem* 2009; 52: 2157-60.
- [101] Rockett KA, Playfair JHL, Targett GAT, Angliker H, Shaw E. Inhibition of intraerythrocytic development of *Plasmodium falciparum* by proteinase inhibitors. *FEBS Lett* 1990; 259: 257-9.
- [102] Rosenthal PJ, Wollish WS, Palmer JT, Rasnick D. Antimalarial effects of peptide inhibitors of a *Plasmodium falciparum* cysteine proteinase. *J Clin Invest* 1991; 88: 1467-72.
- [103] Rosenthal PJ, Lee GK, Smith RE. Inhibition of a *Plasmodium vinckeii* cysteine proteinase cures murine malaria. *J Clin Invest* 1993; 91: 1052-6.
- [104] James KE, Asgian JL, Li ZZ, et al. Design, synthesis and evaluation of aza-peptide epoxides as selective and potent inhibitors of caspases-1, -3, -6 and -8. *J Med Chem* 2004; 47: 1553-74.
- [105] Schirmeister T, Klockow A. Cysteine protease inhibitors containing small rings. *Mini Rev Med Chem* 2003; 3: 585-96.
- [106] Rosenthal PJ, McKerrow JH, Aikawa M, Nagasawa H, Leech JH. A malarial cysteine proteinase is necessary for hemoglobin degradation by *Plasmodium falciparum*. *J Clin Invest* 1988; 82: 1560-16.
- [107] Shenai BR, Sijwali PS, Singh A, Rosenthal PJ. Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*. *J Biol Chem* 2000; 275: 29000-10.
- [108] Wickham ME, Culvenor JG, Cowman AF. Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. *J Biol Chem* 2003; 278: 37658-63.
- [109] Palmer JT, Rasnick D, Klaus JL, Bromme D. Vinyl sulfones as mechanism-based cysteine protease inhibitors. *J Med Chem* 1995; 38: 3193-6.
- [110] Olson JE, Lee GK, Semenov A, Rosenthal PJ. Antimalarial effects in mice of orally administered peptidyl cysteine protease inhibitors. *Bioorg Med Chem* 1999; 7: 633-8.
- [111] Shenai BR, Lee BJ, Alvarez-Hernandez A, et al. Structure activity relationships for inhibition of cysteine protease activity and development of *Plasmodium falciparum* by peptidyl vinyl sulfones. *Antimicrob Agents Chemother* 2003; 47: 154-60.
- [112] Powers JC, Asgian JL, Ekici OD, James KE. Irreversible inhibitors of serine, cysteine and threonine proteases. *Chem Rev* 2002; 102: 4639-750.
- [113] Capela R, Oliveira R, Goncalves LM, et al. Artemisinin-dipeptidyl vinyl sulfone hybrid molecules: design, synthesis and preliminary SAR for antiplasmodial activity and falcipain-2 inhibition. *Bioorg Med Chem Lett* 2009; 19: 3229-32.
- [114] Batra S, Sabnis YA, Rosenthal PJ, Avery MA. Structure-based approach to falcipain-2 inhibitors: synthesis and biological evaluation of 1,6,7-trisubstituted dihydroisoquinolines and isoquinolines. *Bioorg Med Chem* 2003; 11: 2293-9.
- [115] Goud PM, Sheri A, Desai PV, et al. Design, synthesis and evaluation of trisubstituted thiazoles targeting *Plasmodium falciparum* cysteine proteases. *Med Chem Res* 2005; 14: 74-105.
- [116] Mayer AM, Glaser KB, Cuevas C, et al. The odyssey of marine pharmaceuticals: a current pipeline perspective. *Trends Pharmacol Sci* 2010; 31: 255-65.
- [117] Lomenick B, Olsen RW, Huang, J. Identification of direct protein targets of small molecules. *ACS Chem Biol* 2011; 6: 34-46.
- [118] Chen M, Theander TG, Christensen SB, Hviid L, Zhai L, Kharazmi A. Licochalcone A, a new antimalarial agent, inhibits *in vitro* growth of the human malaria parasite *Plasmodium falciparum* and protects mice from *P. yoelii* infection. *Antimicrob Agents Chemother* 1994; 38: 1470-5.
- [119] Li R, Chen X, Gong B, et al. *In vitro* antimalarial activity of chalcones and their derivatives. *J Med Chem* 1995; 38: 5031-7.

- [120] Dominguez JN, Charris JE, Lobo G, *et al.* Synthesis of quinolinyl chalcones and evaluation of their antimalarial activity. *Eur J Med Chem* 2001; 36: 555-60.
- [121] Liu M, Wilairat P, Go ML. Antimalarial alkoxylated and hydroxylated chalcones: structure-activity relationship analysis. *J Med Chem* 2001; 44: 4443-52.
- [122] Dominguez JN, Leon C, Rodrigues J, Gamboa de Dominguez N, Gut J, Rosenthal PJ. Synthesis and evaluation of new antimalarial phenylurenyl chalcone derivatives. *J Med Chem* 2005; 48: 3654-8.
- [123] Stolze SC, Deu E, Kaschani F, *et al.* The antimalarial natural product symplostatins 4 is a nanomolar inhibitor of the food vacuole falcipains. *Chem Biol* 2012; 19: 1546-55.
- [124] Webber SE, Tikhe J, Worland ST, *et al.* Design, synthesis, and evaluation of nonpeptidic inhibitors of human rhinovirus 3C protease. *J Med Chem* 1996; 39: 5072-82.
- [125] Shuttleworth SJ, Nasturica D, Gervais C, Siddiqui MA, Rando RF, Lee N. Synthesis of isatin-based serine protease inhibitors. *Bioorg Med Chem Lett* 2000; 10: 2501-4.
- [126] Ettari R, Micale N, Schirmeister T, *et al.* Novel peptidomimetics containing a vinyl ester moiety as highly potent and selective falcipain-2 inhibitors. *J Med Chem* 2009; 52: 2157-60.
- [127] Grazioso G, Legnani L, Toma L, Ettari R, Micale N, De Micheli C. Mechanism of falcipain-2 inhibition by α,β -unsaturated benzo[1,4]diazepin-2-one methyl ester. *J Comput Aided Mol Des* 2012; 26: 1035-43.
- [128] Ettari R, Micale N, Grazioso G, *et al.* Synthesis and molecular modeling studies of derivatives of a highly potent peptidomimetic vinyl ester as falcipain-2 inhibitors. *Chem Med Chem* 2012; 7: 1594-600.
- [129] Choi HJ, Cui M, Li DY, Song HO, Kim HS, Park H. Anti-malarial activity of new *N*-acetyl-L-leucyl-L-leucyl-L-norleucinal (ALLN) derivatives against *Plasmodium falciparum*. *Bioorg Med Chem Lett* 2013; 23: 1293-6.
- [130] Mallik SK, Li DY, Cui M, Song HO, Park H, Kim HS. Synthesis and evaluation of peptidyl α,β -unsaturated carbonyl derivatives as anti-malarial calpain inhibitors. *Arch Pharm Res* 2012; 35: 469-79.
- [131] Ettari R, Zappalà M, Micale N, *et al.* Peptidomimetics containing a vinyl ketone warhead as falcipain-2 inhibitors. *Eur J Med Chem* 2011; 46: 2058-65.
- [132] Jorgensen WL. The many roles of computation in drug discovery. *Science* 2004; 303: 1813-8.
- [133] Njogu PM, Guantai EM, Pavadai E, Chibale K. Computer-aided drug discovery approaches against the tropical infectious diseases malaria, tuberculosis, trypanosomiasis, and leishmaniasis. *ACS Infect Dis* 2016; 2: 8-31.
- [134] Nair PC, Miners JO. Molecular dynamics simulations: from structure-function relationships to drug discovery. *In Silico Pharmacol* 2014; 2: 1-4.
- [135] Borhani DW, Shaw DE. The future of molecular dynamics simulations in drug discovery. *J Comput Aided Mol Des* 2012; 26: 15-26.
- [136] Cieplak P, Kollman PA. A technique to study molecular recognition in drug design: preliminary application of free energy derivatives to inhibition of a malarial cysteine protease. *J Mol Recognit* 1996; 9: 103-15.
- [137] Senn HM, Thiel W. QM/MM methods for biomolecular systems. *Angew Chem Int Ed* 2009; 48: 1198-229.
- [138] van der Kamp MW, Mulholland AJ. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* 2013; 52: 2708-28.
- [139] Du QS, Wang SQ, Zhu Y, *et al.* Polyprotein cleavage mechanism of SARS CoV Mpro and chemical modification of the octapeptide. *Peptides* 2004; 25: 1857-64.
- [140] Arafet K, Ferrer S, Martí S, Moliner V. Quantum mechanics/molecular mechanics studies of the mechanism of falcipain-2 inhibition by the epoxysuccinate E64. *Biochemistry* 2014; 53: 3336-46.
- [141] Omotuyi IO, Hamada T. Dynamical footprint of falcipain-2 catalytic triad in hemoglobin- β -bound state. *J Biomol Struct Dyn* 2015; 33: 1027-36.
- [142] Francis SE, Sullivan DJ, Goldberg DE. Hemoglobin metabolism in the malaria parasite *Plasmodium falciparum*. *Annu Rev Microbiol* 1997; 51: 97-123.
- [143] Atamna H, Ginsburg H. Origin of reactive oxygen species in erythrocytes infected with *Plasmodium falciparum*. *Mol Biochem Parasitol* 1993; 61: 231-72.
- [144] Pagola S, Stephens PW, Bohle DS, Kosar AD, Madsen SK. The structure of malaria pigment beta-haematin. *Nature* 2000; 404: 307-17.
- [145] Marques AF, Gomes PSFC, Oliveira PL, Rosenthal PJ, Pascutti PG, Lima LM. Allosteric regulation of the *Plasmodium falciparum* cysteine protease falcipain-2 by heme. *Archiv Biochem Biophys* 2015; 573: 92-6.
- [146] Saddala MS, Babu DK, Bhavani G, Ayisha S, Usha RA. Molecular dynamic simulations docking inhibitors of falcipain-2. *Online J Bioinform* 2014; 15: 98-105.
- [147] Musyoka TM, Kanzi AM, Lobb KA, Bishop ÖT. Structure based docking and molecular dynamic studies of plasmodial cysteine proteases against a South African natural compound and its analogs. *Sci Rep* 2016; 6: 23690.
- [148] Ring CS, Sun E, McKerrow JH, *et al.* Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *Proc Natl Acad Sci USA* 1993; 90: 3583-7.
- [149] Schirmeister T, Kaeppeler U. Non-peptidic inhibitors of cysteine proteases. *Mini Rev Med Chem* 2003; 3: 361-73.
- [150] Li R, Kenyon GL, Cohen FE, *et al.* *In vitro* antimalarial activity of chalcones and their derivatives. *J Med Chem* 1995; 38: 5031-7.
- [151] Joachimiak MP, Chang C, Rosenthal PJ, Cohen FE. The impact of whole genome sequence data on drug discovery a malaria case study. *Mol Med* 2001; 7: 698-710.
- [152] Sabnis Y, Rosenthal PJ, Desai P, Avery MA. Homology modeling of falcipain-2: validation, *de novo* ligand design and synthesis of novel inhibitors. *J Biomol Struct Dyn* 2002; 19: 765-74.
- [153] Sabnis YA, Desai PV, Rosenthal PJ, Avery MA. Probing the structure of falcipain-3, a cysteine protease from *Plasmodium falciparum*: Comparative protein modeling and docking studies. *Protein Sci* 2003; 12: 501-9.
- [154] Desai PV, Patny A, Sabnis Y, *et al.* Identification of novel parasitic cysteine protease inhibitors using virtual screening. 1. the Cambridge database. *J Med Chem* 2004; 47: 6609-15.
- [155] Desai PV, Patny A, Gut J, *et al.* Identification of novel parasitic cysteine protease inhibitors by use of virtual screening. 2. The available chemical directory. *J Med Chem* 2006; 49: 1576-84.
- [156] Li H, Huang H, Chen L, *et al.* Identification of novel falcipain-2 inhibitors as potential antimalarial agents through structure-based virtual screening. *J Med Chem* 2009; 52: 4936-40.
- [157] Shah F, Mukherjee P, Gut J, *et al.* Identification of novel malarial cysteine protease inhibitors using structure-based virtual screening of a focused cysteine protease inhibitor library. *J Chem Inf Model* 2011; 51: 852-64.
- [158] Chintakrindi AS, Shaikh MS, Coutinho EC. *De novo* design of 7-aminocoumarin derivatives as novel falcipain-3 inhibitors. *J Mol Model* 2012; 18: 1481-93.
- [159] Zhu J, Chen T, Chen L, *et al.* 2-Amido-3-(1*H*-Indol-3-yl)-*N*-substituted-propanamides as a new class of falcipain-2 inhibitors. 1. Design, synthesis, biological evaluation and binding model studies. *Molecules* 2009; 14: 494-508.
- [160] Musyoka TM, Kanzi AM, Lobb KA, Bishop ÖT. Analysis of non-peptidic compounds as potential malarial inhibitors against plasmodial cysteine proteases *via* integrated virtual screening workflow. *J Biomol Struct Dyn* 2016; 34(10): 2084-101.
- [161] Gibbons P, Verissimo E, Araujo NC, *et al.* Endoperoxide carbonyl falcipain 2/3 inhibitor hybrids: toward combination chemotherapy of malaria through a single chemical entity. *J Med Chem* 2010; 53: 8202-6.
- [162] Ibezim E, Duchowicz PR, Ortiz EV, Castro EA. QSAR on arylpiperazine derivatives with activity on malaria. *Chemometrics and Intelligent Laboratory Systems* 2012; 110: 81-8.
- [163] Soni LK, Patidar K. QSAR analysis of 3-phenylquinoxaline (1,4)-di-*N*-oxide derivatives as antimalarial agents. *Pharmaceut Sci* 2015; 3: 6-8.
- [164] Ghasemi JB, Shiri F. Molecular docking and 3D-QSAR studies of falcipain inhibitors using CoMFA, CoMSIA, and Open 3D QSAR. *Med Chem Res* 2012; 21: 2788-806.
- [165] Wang J, Li F, Li Y, Yang Y, Zhang S, Yang L. Structural features of falcipain-3 inhibitors: an *in silico* study. *Mol BioSyst* 2013; 9: 2296-310.
- [166] Ehmke V, Heind C, Rottmann M, *et al.* Potent and selective inhibition of cysteine proteases from *Plasmodium falciparum* and *Trypanosoma brucei*. *ChemMedChem* 2011; 6: 273-78.
- [167] Oballa RM, Truchon JF, Bayly CI, *et al.* Generally applicable method for assessing the electrophilicity and reactivity of diverse



- nitrile-containing compounds. *Bioorg Med Chem Lett* 2007; 17: 998-1002.
- [168] Mundra S, Mahesh R. Evaluation of novel 1-(4-(substituted) piperazin-1-yl)-2-(phenylamino)ethanone derivatives as falcipain-2 inhibitors. *J Young Pharm* 2015; 7: 96-105.
- [169] Qing X, Lee XY, Raeymaeker JD, Tame JRH, Zhang KYJ. Pharmacophore modeling: advances, limitations, and current utility in drug discovery. *Journal of Receptor, Ligand and Channel Research* 2014; 7: 81-92.
- [170] Mugumbate G, Newton AS, Rosenthal PJ, *et al.* Novel antiplasmodial hits identified by virtual screening of the ZINC database. *J Comput Aided Mol Des* 2013; 27: 859-71.
- [171] Mustata G, Follis AV, Hammoudeh DI, *et al.* Discovery of novel Myc-Max heterodimer disruptors with a three-dimensional pharmacophore model. *J Med Chem* 2009; 52: 1247-50.
- [172] Chakka SK, Kalamuddin M, Sundararaman S, *et al.* Identification of novel class of falcipain-2 inhibitors as potential antimalarial agents. *Bioorg Med Chem* 2015; 23: 2221-40.
- [173] Wang L, Zhang S, Zhu J, *et al.* Identification of diverse natural products as falcipain-2 inhibitors through structure-based virtual screening. *Bioorg Med Chem Lett* 2014; 24: 1261-4.
- [174] Liu Y, Cui K, Lu W, *et al.* Synthesis and antimalarial activity of novel dihydro-artemisinin derivatives. *Molecules* 2011; 16: 4527-38.

REVIEW

Open Access



The potential of anti-malarial compounds derived from African medicinal plants: a review of pharmacological evaluations from 2013 to 2019

Boris D. Bekono^{1†}, Fidele Ntie-Kang^{2,3,4*†} , Pascal Amoa Onguéné⁵, Lydia L. Lifongo², Wolfgang Sippl³ , Karin Fester⁶ and Luc C. O. Owono^{1*}

Abstract

Background: African Traditional Medicine (ATM) is used for the healthcare of about 80% of the rural populations of the continent of Africa. The practices of ATM make use of plant-products, which are known to contain plant-based secondary metabolites or natural products (NPs), likely to play key roles in drug discovery, particularly as lead compounds. For various reasons, including resistance of strains of *Plasmodium* to known anti-malarial drugs, local African populations often resort to plant-based treatments and/or a combination of this and standard anti-malarial regimens. Emphasis has been laid in this review to present the anti-malarial virtue of the most recently published phytochemicals or natural products, which have been tested by in vitro and in vivo assays.

Methods: The data was based on the current version of the African Compound Libraries, which are constantly being updated based on inputs from journal articles and student theses (M.Sc/Ph.D) from African University libraries. Emphasis was laid on data published after 2012. In order to carry out the original data collection, currently being included in the African Compounds Database, individual journal websites were queried using the country names in Africa as search terms. Over 40,000 articles "hits" were originally retrieved, then reduced to about 9000 articles. The retained articles/theses was further queried with the search terms "malaria", "malarial", "plasmodium", "plasmodial" and a combination of them, resulting in over 500 articles. Those including compounds with anti-malarial activities for which the measured activities fell within the established cut off values numbered 55, which were all cited in the review as relevant references.

Results and discussion: Pure compounds derived from African medicinal plants with demonstrated anti-malarial/antiplasmodial properties with activities ranging from "very active" to "weakly active" have been discussed. The majority of the 187 natural products were terpenoids (30%), followed by flavonoids (22%), alkaloids (19%) and quinones (15%), with each of the other compound classes being less than 5% of the entire compound collection. It was also observed that most of the plant species from which the compounds were identified were of the families Rubiaceae,

*Correspondence: ntiékfidele@gmail.com; lcowono@yahoo.fr

†Boris D. Bekono and Fidele Ntie-Kang contributed equally and should be regarded as joint first authors

¹ Department of Physics, Ecole Normale Supérieure, University of Yaoundé I, P. O. Box 47, Yaoundé, Cameroon

² Department of Chemistry, Faculty of Science, University of Buea, P. O. Box 63, Buea, Cameroon

Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Meliaceae and Asphodelaceae. The review is intended to continue laying the groundwork for an African-based anti-malarial drug discovery project.

Keywords: Africa, Malaria, Medicinal plants, Natural products, Traditional medicine

Background

Malaria is an endemic disease in most tropical countries (Africa, Asia, and Latin America), with about half of the world's population at risk of infection according to the World Health Organization (WHO) [1]. According to the latest World Malaria Report, released in December 2019, there were 228 million cases of malaria in 2018, and the estimated number of malaria deaths stood at 405,000. The causative agents for malaria infections are *Plasmodium* protozoans (i.e. *Plasmodium falciparum*, *Plasmodium malariae*, *Plasmodium ovale*, and *Plasmodium vivax*), although most severe infections are caused by *P. falciparum* [2–4]. Most deaths are recorded among African children below the age of 5 years [1–4]. This calls for an urgent need for new anti-malarial therapies for any one of the following reasons:

- The development of resistance against insecticides (e.g. dichlorodiphenyltrichloroethane, DDT) by the disease vectors (female anopheline mosquitoes) [5–7].
- The inefficacy of chemoprophylaxis, which has often resulted in poor results [1, 8–10].
- The development of resistance by *Plasmodium* protozoans against most of the drugs currently used to treat malaria (e.g. chloroquine, artemisinin and its derivatives) [11–13].

Plants are known to be a rich reservoir of bioactive secondary metabolites (or natural products, NPs), for example, the anti-malarial drugs quinine and artemisinin (AT) are both of plant origin [14]. The benefits of plants containing bioactive anti-malarial compounds, particularly the bitter principles (alkaloids and terpenoids), include their use in the preparation of traditional remedies against malaria, fever, and inflammation [15]. In fact, more than 80% of the local populations of most tropical countries, including African populations, are dependent on medicinal plants for the treatment of most diseases, including malaria, despite the current wide availability of standard malaria treatments for populations in the rural areas, as well as those in cities [16, 17]. It has become of interest to summarize the major findings regarding the most promising secondary metabolites with proven in vitro and in vivo potencies, so as to pave the way for further development with compounds from African sources as leads for anti-malarial drug discovery. Recent

reviews have either emphasized plants used in specific countries or regions for the treatment of malaria [18–21], secondary metabolites from selected plant species [22] or families of species [23], plants used as repellents against the mosquito vectors [24], or reports on the analysis of components of improved traditional preparations against malaria [25, 26].

Previous reviews have described the in vitro and in vivo potencies of compounds that have been isolated from African floral matter published data before 2013 [27, 28]. These reviews had previously described over 500 NPs, within the major NP classes, including alkaloids, terpenoids, flavonoids, coumarins, phenolics, polyacetylenes, xanthenes, quinones, steroids, and lignans. These compounds were described in the literature as exhibiting from weak to very good in vitro anti-malarial activities, based on well-established cut-off values [29–31]. Besides, a cheminformatic analysis of the aforementioned dataset, with a focus on molecular descriptors related to “drug-likeness”, drug metabolism and pharmacokinetics (DMPK), and some rules of thumb such as the Lipinski “Rule of Five” [32], showed that over 50% of the anti-malarial compounds had physicochemical properties that fell within the range of “drug-like” molecules [33].

The present review focuses on compounds with tested activities against various malaria parasites derived from a literature survey from 2013 to 2019 [29–31]. A total of 187 NPs belonging to diverse classes, including alkaloids, flavonoids, phenolics, flavonoids, steroids, and terpenoids are described. These compounds have been identified from 45 plant species belonging to 23 families. It is hoped that the results summarized will help for lead compound identification and for further anti-malarial drug discovery. The review describes the NPs with potential anti-malarial properties from African medicinal plants, arranged alphabetically according to the main NP compound classes.

Materials and methods

Data collection

In this review, an attempt has been made to document the anti-malarial activities of NPs derived from African medicinal plants. The data was based on the current version of the African compound libraries [34–37], which are constantly being updated based on inputs from journal articles and student theses (M.Sc/Ph.D.) available in African University libraries. Emphasis was

Table 1 Summary of testing methodologies and parasite strains described in this report

| Murine (in vivo) models | Strains | Parasite name | Origin | Assay description | References | |
|---------------------------------|-----------------|-----------------------------------|--|--|--|---------|
| CQ-sensitive | NK 173 | <i>Plasmodium berghei</i> | Not reported | Not reported | [38–40] | |
| | ANKA | <i>P. berghei</i> | Not reported | Not reported | | |
| | | <i>Plasmodium vinckei petteri</i> | Not reported | Not reported | | |
| <i>In vitro</i> models | | | | | | |
| CQ-sensitive | 3D7 | <i>P. falciparum</i> | Not reported | Parasite lactate dehydrogenase (pLDH) assay | [41, 42] | |
| | | | Not reported | Parasite growth inhibition assay | [43] | |
| | | | Not reported | Translation inhibitory assay | [44] | |
| | D6 | <i>P. falciparum</i> | Sierra Leone | Incorporated G- ³ H hypoxanthine assay | [41, 42, 45, 46] | |
| | | | | Parasite lactate dehydrogenase (pLDH) assay | [47, 48] | |
| | | | | Non-radioactive Malaria SYBR Green I assay | [49, 50] | |
| | | | | Modified non-radioactive Malaria SYBR Green I assay | [42, 49, 51] | |
| | D10 | <i>P. falciparum</i> | Not reported | pLDH assay | [42, 52] | |
| | F ₃₂ | <i>P. falciparum</i> | Tanzania | Not reported | | |
| | FCA20 | <i>P. falciparum</i> | Ghana | Not reported | | |
| | CQ-resistant | K1 | <i>P. falciparum</i> | Thailand | Modified [³ H]-hypoxanthine incorporation assay and [³ H]-hypoxanthine incorporation assay | [53–56] |
| | | | | | | |
| | | NF54 | <i>P. falciparum</i> | Not reported | [³ H]-hypoxanthine incorporation assay | [53–56] |
| Dd2 | | <i>P. falciparum</i> | Not reported | Non-radioactive Malaria SYBR Green I assay | [49, 50] | |
| FcM ₂₉ | | <i>P. falciparum</i> | Cameroon | Not reported | | |
| FcB1 | | <i>P. falciparum</i> | Colombia | [³ H]-hypoxanthine incorporation assay | [41] | |
| CQ- and pyrimethamine-resistant | | K1 | <i>P. falciparum</i> | Thailand | Modified [³ H]-hypoxanthine incorporation assay and [³ H]-hypoxanthine incorporation assay | [53–56] |
| | | | | | | |
| | W2 | <i>P. falciparum</i> | Indochina | Modified non-radioactive Malaria SYBR Green I assay | [42, 49, 51] | |
| | | | | Incorporated G- ³ H hypoxanthine assay | [45, 46] | |
| | | | | Non-radioactive Malaria SYBR Green I assay | [49, 50] | |
| Multidrug-resistant | NF54 | <i>P. falciparum</i> | | [³ H]-hypoxanthine incorporation assay | [41] | |
| | | | | | | |
| | K1 | <i>P. falciparum</i> | Thailand | Modified [³ H]-hypoxanthine incorporation assay and [³ H]-hypoxanthine incorporation assay | [53–55] | |
| | | | | | | |
| | | | | [³ H]-hypoxanthine incorporation assay | [54] | |
| | | | | [³ H]-hypoxanthine incorporation assay | [53–56] | |
| | | | Parasite lactate dehydrogenase (pLDH) assay | [52, 57] | | |
| | | | [³ H]-hypoxanthine incorporation assay | [53–56] | | |
| | | | [³ H]-hypoxanthine incorporation assay | [45, 46] | | |
| | | | [³ H]-hypoxanthine incorporation assay | [53–56] | | |

laid on data published after 2012. The original data collection, now being included in the African Compounds Database (<http://www.african-compounds.org>), was conducted from querying individual journal websites using the country names in Africa and search terms. The list of journals visited have been included in Additional file 1. The “hit” articles were retrieved, i.e. those for which plant materials were collected from Africa were then hand-picked by reading through the “Materials and methods” section to ensure that the plant materials

were from Africa. Student theses were also randomly collected as made available from university libraries, constituting a small portion of the data. The folder containing the retained articles/theses was further queried with the search terms “malaria”, “malarial”, “plasmodium”, “plasmodial” and a combination of them. Those for which compounds further showed anti-malarial activities published between 2013 and 2019 and for which the measured activities fell within the established cut-off values were selected and included as relevant references. The

Table 2 Summary of alkaloids

| Compound subclass | Isolated metabolites | Plasmodial strain (activities) | Plant species (Family), Taxon ID ^b | Part of the plant studied | Place of harvest (Locality, Country) | Author, references |
|-----------------------|--|---|---|---------------------------|---|---|
| Aporphines | 1 to 2 | K1 (8.24 and 2.90 μ M, respectively) | <i>Annickia kummeriae</i> (Annonaceae), NCBI:txid225831 | Leaves | Amani Nature Reserve, Tanzania | Malebo et al. [58] |
| Furoquinolines | 3^a to 6 | FcB1 (from 162.47 to 298.16 μ M) | <i>Teclea nobilis</i> (Rutaceae), NCBI:txid1220089 | Fruits and leaves | Kamwenge district, Uganda | Lacroix et al. [59] |
| | 7 | Dd2 (IC ₅₀ = 35 μ M) | <i>Melicope madagascariensis</i> (Rutaceae), NCBI:txid1487113 | Stem bark | Antsasaka forest of Moramanga, Madagascar | Rasamison et al. [60] |
| Indoles | 8^a to 13 and 15 | 3D7 (from 0.41 to 110.58 μ M) | <i>Strychnos icaja</i> (Loganiaceae), NCBI:txid1040889 | Stem bark | Bertoua, Cameroon | Tchinda et al. [61] |
| | 14 | FCA20 (0.617 μ M) W2 (0.085 μ M) | <i>Strychnos icaja</i> (Loganiaceae), NCBI:txid1040889 | Roots | Kasongo-Lunda, DR Congo | Beaufay et al. [62], Frédérich et al. [63] |
| Indolosesquiterpenes | 16^a and 17^a | NF54 (7.6 μ M and 29.1 μ M, respectively) | <i>Polyalthia oliveri</i> (Annonaceae), NCBI:txid105756 | Stem bark | Mount Kala, Cameroon | Kouam et al. [64] |
| Naphthylisoquinolines | 18^a and 19^a | NF54 (0.043 and 0.055 μ M, respectively) | <i>Ancistrocladus</i> sp. (Ancistrocladaceae), NCBI:txid 63071 | Leaves | Mbandaka, DR Congo | Lombe et al. [65] |
| | 20^a , 21^a , 22^a and 23 to 25 | NF54 (from 0.090 to 6.54 μ M) K1 (0.228 μ M for compound 19) | <i>Ancistrocladus ileboensis</i> (Ancistrocladaceae), NCBI:txid1367080 | Leaves and root bark | Bambange, DR Congo | Li et al. [66] |
| | 26^a , 27^a , 28^a and 29^a | NF54 (from 0.84 to 22.2 μ M) K1 (from 1.4 to 8.2 μ M) | <i>Ancistrocladus ealaensis</i> (Ancistrocladaceae), NCBI:txid714098 | Twigs and leaves | Mbandaka, DR Congo | Tshitenge et al. [67] |
| Protoberberines | 30 to 33 | K1 (from 0.22 to 0.71 μ M) | <i>Annickia kummeriae</i> (Annonaceae), NCBI:txid225831 | Leaves | Amani Nature Reserve, Tanzania | Malebo et al. [58] |
| | 34 | K1 (IC ₅₀ = 318.66 μ M) | <i>Polyalthia longifolium</i> var. <i>pendula</i> (Annonaceae), NCBI:txid235806 | Stem | Tikrom, near Kumasi, Ghana | Gbedema et al. [68] |
| Pyridinones | 35 | K1 (IC ₅₀ = 81.28 μ M) | <i>Polyalthia longifolium</i> var. <i>pendula</i> (Annonaceae), NCBI:txid235806 | Stem | Tikrom, near Kumasi, Ghana | Gbedema et al. [68] |
| Others | 36 | K1 (IC ₅₀ = 32.12 μ M) | <i>Canthium multiflorum</i> (Rubiaceae), NCBI:txid58501 | Aerial part | Obala, along River Sanaga, Cameroon | Kouam et al. [69] |

^a Compounds identified for the first time in the cited publications

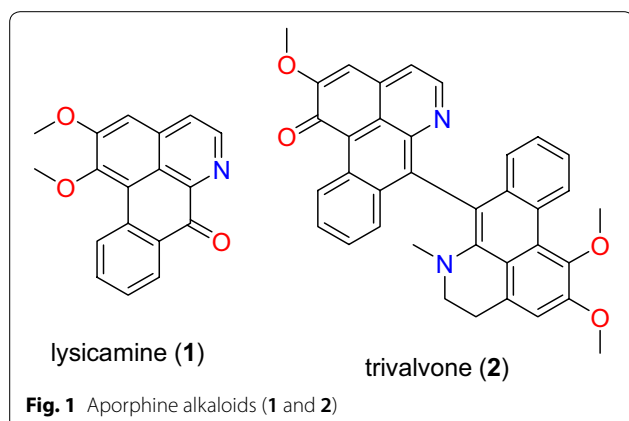
^b Identification number of the source species, derived from the NCBI Taxonomy database

compounds displaying anti-malarial properties were classified according to their NP classes and source species of origin. This represented 187 compounds from 45 species belonging to 23 plant families.

Data analysis

The collected data was arranged into spreadsheets according to plant sources, compound classes, activity cut-offs and plasmodial strains tested. All activity data was converted to IC₅₀ values in μ M.

Throughout the text, the term antiplasmodial is referred to as that which counters the growth of parasites of the genus *Plasmodium*, while anti-malarial is referred to as an agent which prevents or counteracts the progress of the disease caused by the parasite or that which treats the disease (i.e. by killing the parasites in the host). Very often the two terms are used interchangeably in the literature surveyed.



Test methodologies

From the literature collected, a broad range plasmodial strains were tested, including those summarized in Table 1.

Promising anti-malarial compounds derived from the African flora

Alkaloids

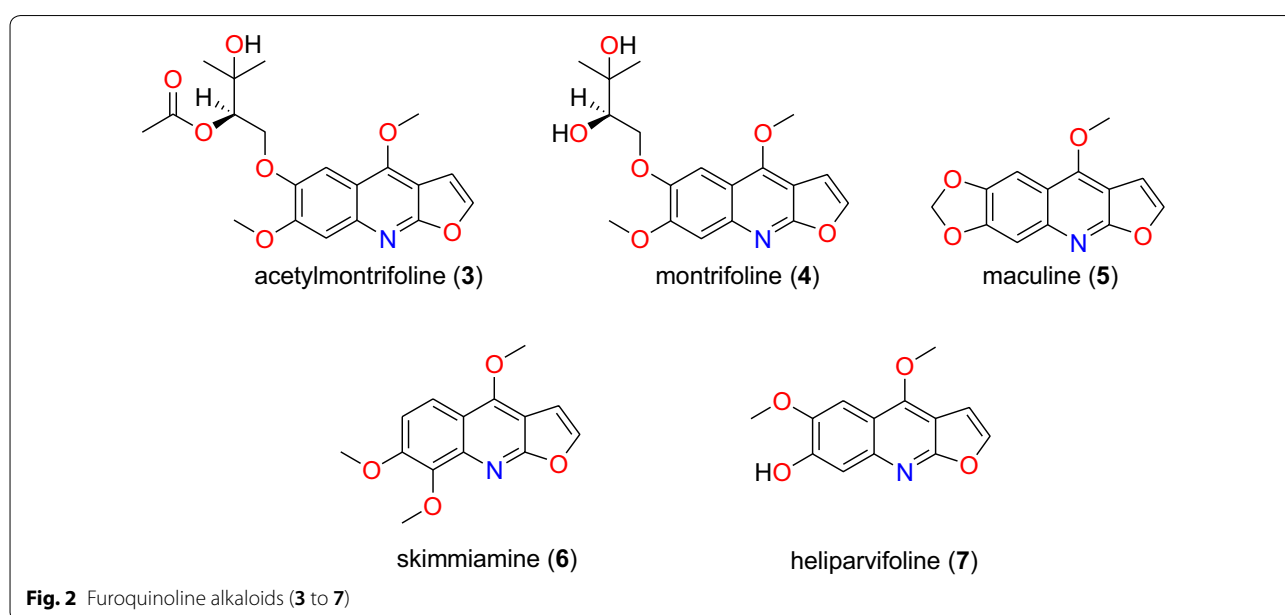
Table 2 summarizes the most promising alkaloids derived from the African flora, published since the earlier review [27], while the chemical structures are shown in Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10, arranged alphabetically according to their respective sub-classes (i.e. aporphines, furoquinolines, indoles, indolosesquiterpenes, Naphthylisoquinolines, protoberberines, pyridinones and others). Several of them had tested

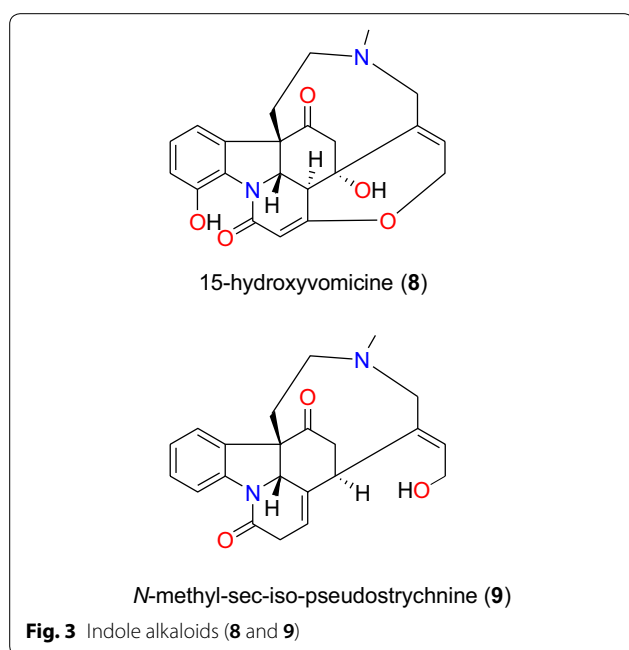
positive against CQ-sensitive and CQ-resistant strains of *P. falciparum* in vitro.

Aporphines The aporphine alkaloids lysicamine (1), trivalvone (2) (Fig. 1) were identified from the leaves of *Annickia kummeriae* (Annonaceae) from Tanzania, along with four other (protoberberine) alkaloids. Plants from the genus *Annickia* (formerly *Enantia*) are popularly known in West and Central Africa for use in the treatment of malaria [70–72]. The study by Maleba et al. [58] showed that compounds 1 and 2 showed respective activities of 8.23 and 2.90 μM against the CQ-resistant K1 strain of *P. falciparum*.

Furoquinolines Four furoquinoline alkaloids (3 to 6) (Fig. 2) were isolated from the fruits and leaves of *Teclea nobilis* (Rutaceae) and tested on the chloroquine (CQ)-resistant FcB1/Colombia strain of *P. falciparum* by Lacroix et al. [59]. This species from Uganda has been used to treat a range of ailments from pain and fever to malaria [73]. The isolated compounds, including the novel acetylmontrifoline (3), and the known montrifoline (4), maculine (5), and skimmiamine (6), were less potent than the reference drug CQ, showing inhibition against the tested parasite strain at $< 300 \mu\text{M}$ [59].

Rasamison et al. also isolated seven furoquinolines from the stem bark of the Madagascari species, *Melicope madagascariensis* (Rutaceae), of which only compound 7 (6-methoxy-7-hydroxydictamnine, commonly called heliparvifoline) exhibited weak anti-malarial activity against the CQ-resistant strain, Dd2, with





$IC_{50} = 35 \mu\text{M}$, the other compounds tested being inactive [60].

Indoles *Strychnos icaia* (Loganiaceae) is found all over Central Africa [74]. In Cameroon, for example, the roots are used by a Pygmy tribe to treat malaria. From their stem bark, six indole alkaloids (8 to 13) (Figs. 3 and 4), were isolated and evaluated against the CQ-sensitive 3D7 strain of *P. falciparum* by Tchinda et al. [61], with IC_{50} values ranging from 0.40 to 110 μM . These include 15-hydroxyvomicine (8), *N*-methyl-sec-iso-pseudostrychnine (9), sungucine (10), isosungucine (11), strychnogucine C (12), bisnordihydrotoxiferine (13), along with the chlorinated indole, N_b -chloromethosungucine (14).

Strychnogucine B (15) (Fig. 4), which was previously isolated from the roots of the same species by Frédéric et al. [63] was further investigated by Beaufoy et al. [62] and the compound now displayed further inhibition against the CQ-sensitive FCA 20/Ghana and CQ-resistant W2/Indochina strains, with IC_{50} values of 0.617 and 0.085 μM , respectively.

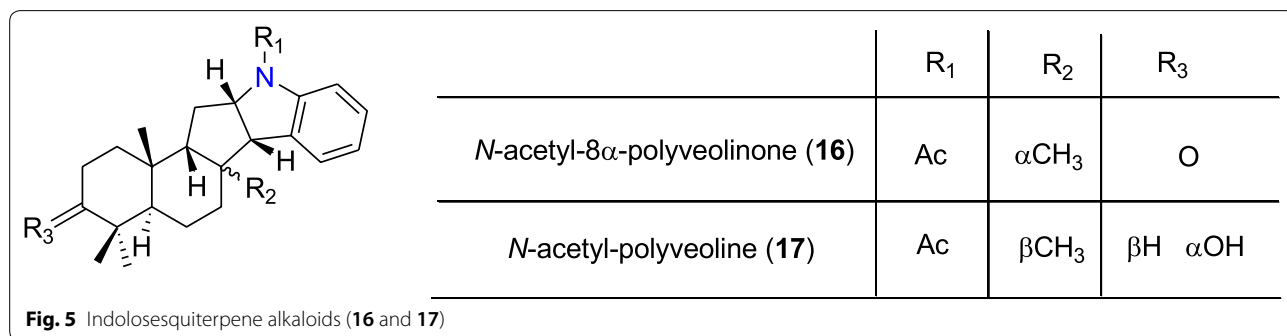
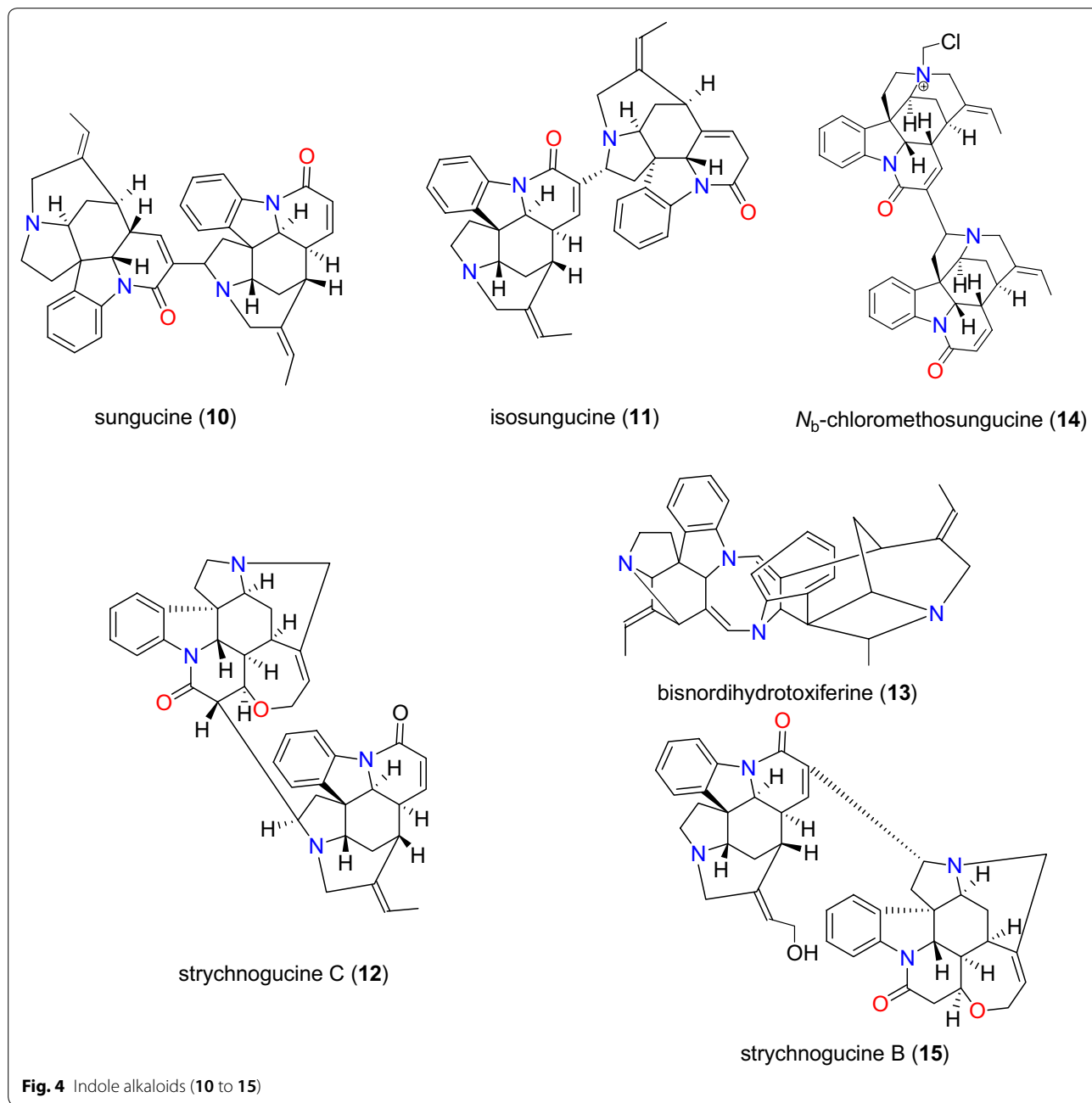
Indolosesquiterpenes The bioactivity-guided screening of the stem bark of *Polyalthia oliveri* (Annonaceae) led Kouam et al. to isolate two indolosesquiterpene alkaloids, named *N*-acetyl-8 α -polyveolinone (16) and *N*-acetyl-polyveoline (17) (Fig. 5) [64]. This species is used in folk medicine for the treatment of malaria [75]. Both compounds were tested against CQ-sensitive NF54 strain and compound 16 showed moderate antiplasmodial activity

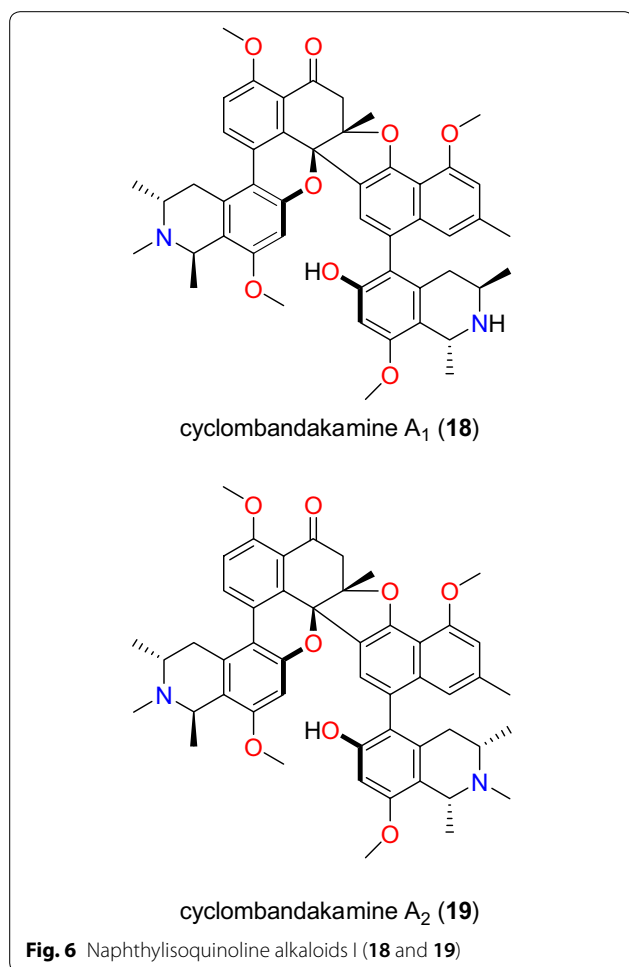
with $IC_{50} = 7.6 \mu\text{M}$, while compound 17 inhibited the strain weakly with an IC_{50} value of 29.1 μM [75].

Naphthylisoquinolines These are compounds characterized by a chiral biarylaxis linkage between the naphthalene and the isoquinoline alkaloids, mainly isolated from plants of the genus *Ancistrocladus* (Ancistrocladaceae), and the closely related genera *Triphyophyllum*, *Dioncophyllum* and *Habropetalum* (Dioncophyllaceae). Cyclombandakamines A₁ (18) and A₂ (19) (Fig. 6) are naphthylisoquinoline alkaloids isolated from the leaves of *Ancistrocladus* sp. (Ancistrocladaceae) by Lombe et al. [65]. These compounds displayed significant inhibitory activities against the NF54 strain of *P. falciparum* with IC_{50} values of 0.043 and 0.055 μM , respectively [65]. Li et al. [66] also investigated the twigs and leaves of *Ancistrocladus ileboensis* (Ancistrocladaceae) from DR Congo. Among the tested compounds with promising anti-malarial activities were dioncophylline F (20) and dioncophylline C₂ (21), in addition to the 7,8'-coupled dioncophylline D₂ (22), ancistrobrevine C (23), 5'-*O*-methyl dioncophylline D (24), and ancistrocladisine A (25) (Fig. 7). Compound 20 showed activities against both the NF54 and K1 strains (with IC_{50} values of 0.090 and 0.045 μM , respectively), compounds 21 to 25 were only tested against the K1 strain, with IC_{50} values ranging from 0.107 to 6.51 μM [66].

Among the compounds identified from Ancistrocladaceae, Tshitenge et al. also isolated four naphthylisoquinolines, named ealamines A–D (26 to 29, Fig. 8) from the twigs and leaves of *Ancistrocladus ealaensis* (Ancistrocladaceae) harvested in Mbandaka, DR Congo [67]. These compounds were tested against CQ-sensitive NF54 and CQ- and pyrimethamine-resistant K1 strains of *P. falciparum*. The activities against the CQ-sensitive NF54 strain showed IC_{50} values of 6.3, 4.9, 0.84 and 22.2 μM , respectively. Meanwhile, compounds 26, 27 and 29 inhibited the CQ- and pyrimethamine-resistant K1 strain with IC_{50} values of 1.6, 1.4, and 8.2 μM , respectively [67].

Protoberberines Maleba et al. [58] showed that against the CQ-resistant K1 strain of *P. falciparum* protoberberine alkaloids are a subclass of promising anti-malarials. The in vitro testing of compounds 30 to 33 (Fig. 9) showed that compound 30 (palmatine) was the most active, with an IC_{50} value of 0.23 μM . Jatrorrhizine (31) exhibited an IC_{50} of 0.71 μM , whereas a mixture of compound 31 and columbamine (32) inhibited the plasmodial strain with an IC_{50} value of 0.14 $\mu\text{g/mL}$, and a mixture of compound 26 and tetrahydro-palmatine (33) inhibited the parasite strain with $IC_{50} = 0.098 \mu\text{g/mL}$, probably explaining the synergistic activity of this plant extract. This justifies its use in African Traditional Medicine for the treatment of malaria [58].

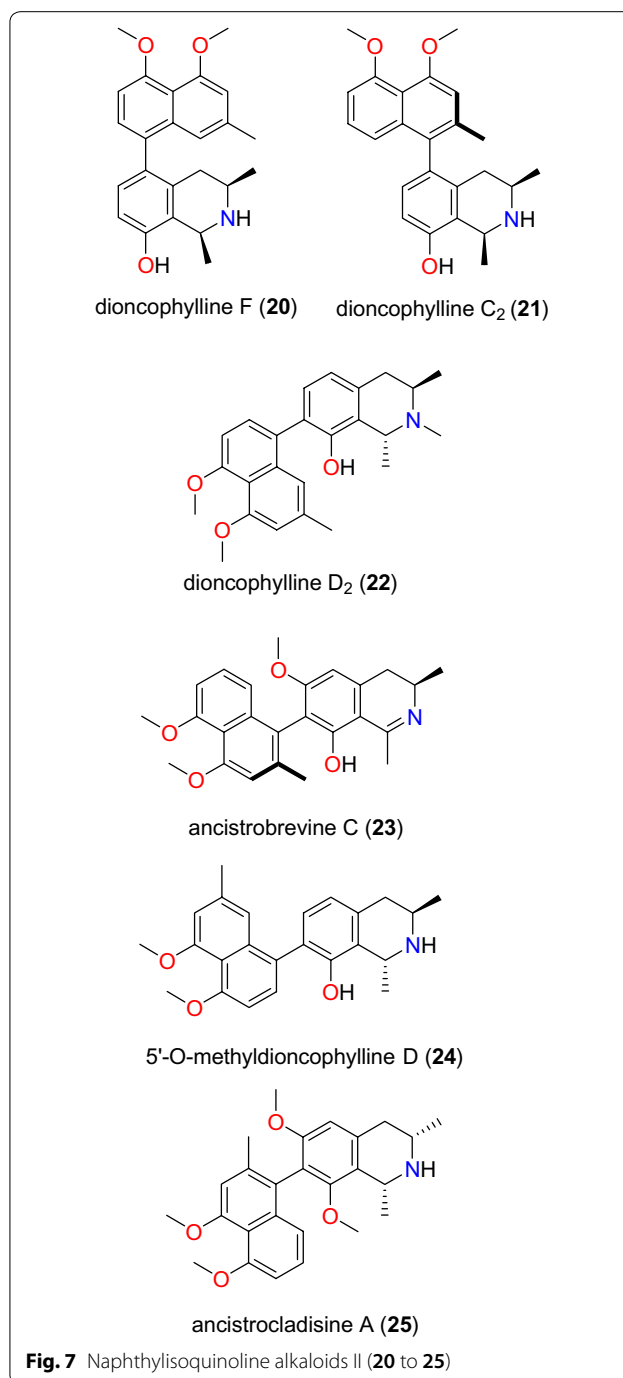




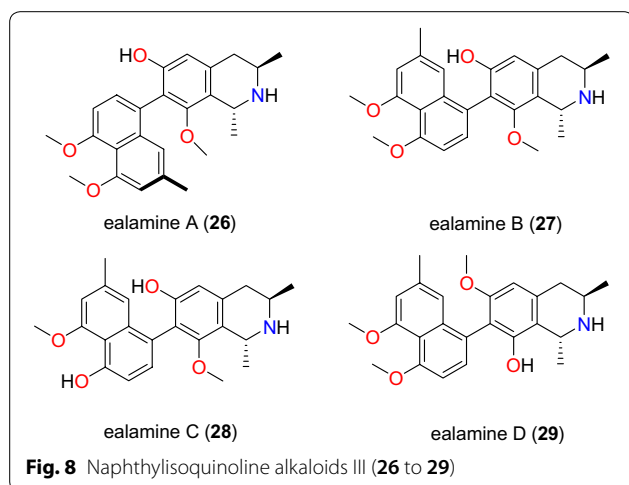
Extracts of *Polyalthia longifolium* (Annonaceae), used in orally consumed preparations in traditional medicine in Ghana, was investigated in order to identify anti-malarial compounds [68]. The protoberberine L-stepholidine (34, Fig. 9) was identified from the stem of species among the isolated compounds [68], but this compound had only a weak antiplasmodial activity against the K1 strain of *P. falciparum*.

Pyridinones Gbedema et al. also isolated darienine (35, Fig. 10), a known alkaloid with anti-malarial activity [68]. This compound exhibited varying degrees of antiplasmodial activity against the K1 strain of *P. falciparum* with an IC₅₀ value of 81.28 μM.

Other alkaloids Gardenine (36, Fig. 10), obtained from the investigation of crude extract of the aerial parts of *Canthium multiflorum* (Rubiaceae), harvested from Cam-



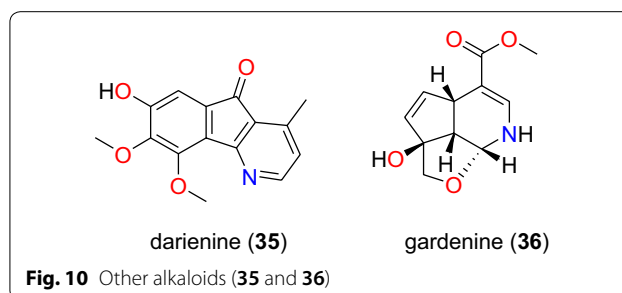
eroon, exhibited antiplasmodial activity against the K1 strain of *P. falciparum*, with an IC₅₀ value of 32.12 μM and weak cytotoxicity against L6 cell lines [69].



Flavonoids

Flavonoids (mainly chalcone, flavanone, isoflavone, and retonoid sub-classes) (Figs. 11, 12, 13 and 14) were previously seen as a promising class of NPs exhibiting anti-malarial and antiplasmodial activities [28].

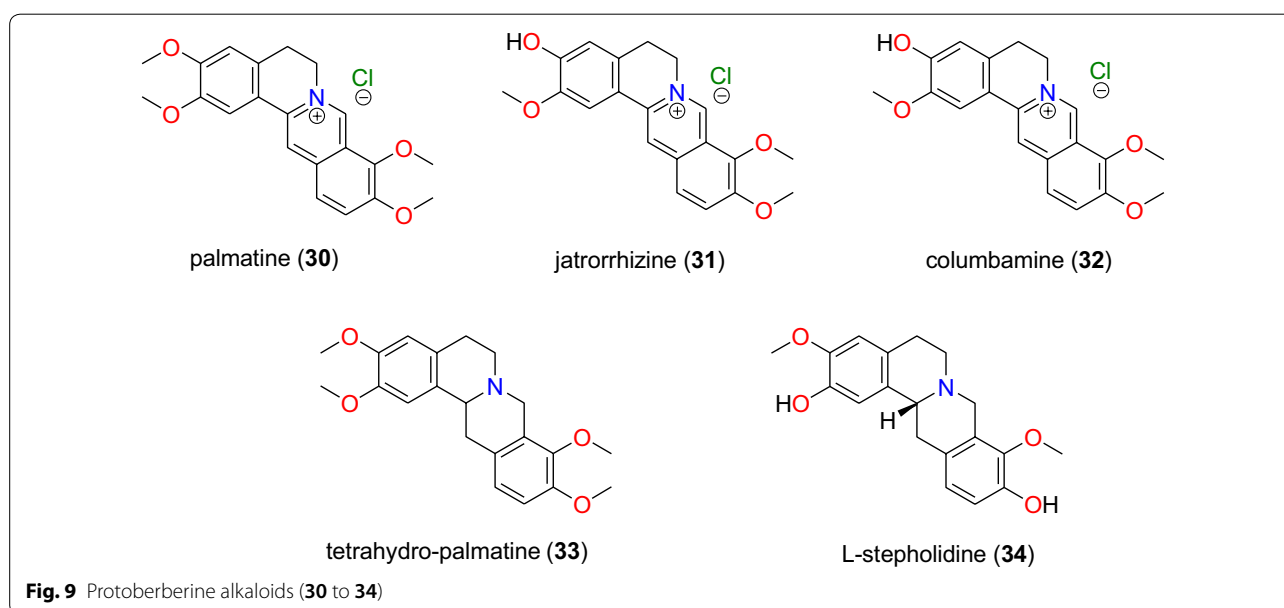
Flavanones and flavones The most promising recently published anti-malarial flavanones and flavones derived from the African flora have been summarized in Table 3. These include 5,4'-dihydroxy-3,6,7-trimethoxyflavone (37), 5,7-dihydroxy-3,4'-dimethoxyflavone (38), quercetin-3,4'-dimethyl ether (39), rhamnazin (40), retusin (41), 5,4'-dihydroxy-3,7,3'-trimethoxyflavone (42), 5,4'-dihydroxy-7-dimethoxyflavanone (43), (+)-tephro-



din (44), tachrosin (45), obovatin methyl ether (46), morloflavone (47), volkensiflavone (48), and 5-demethyl-tangeretin (49), whose chemical structures of are shown in Fig. 11. The compounds were isolated from the species *Senecio roseiflorus* (Compositae-Asteraceae) [76], *Tephrosia villosa* (Leguminosae) [77], *Allanblackia floribunda* (Guttiferae) [78], and *Peperomia vulcanica* (Piperaceae) [79].

Compounds 37 to 43 were derived from leaves of *Senecio roseiflorus* and have shown good to moderate antiplasmodial activities against D6 and W2 strains. The activities in terms of IC_{50} values ranged from 11.25 to 56.31 μM for the D6 strain, while for the W2 strain, this ranged from 15.47 to 87.50 μM [76]. Compounds 44 to 46, were derived from roots of *Tephrosia villosa* and exhibited anti-malarial activities against both the D6 and W2 strains with respective IC_{50} values from 11.30 to 14.00 μM for the D6 strain and from 13.10 to 20.40 μM for the W2 strain [77].

Azebaze et al. investigated the antiplasmodial activities of whole plant extracts of *Allanblackia floribunda* from



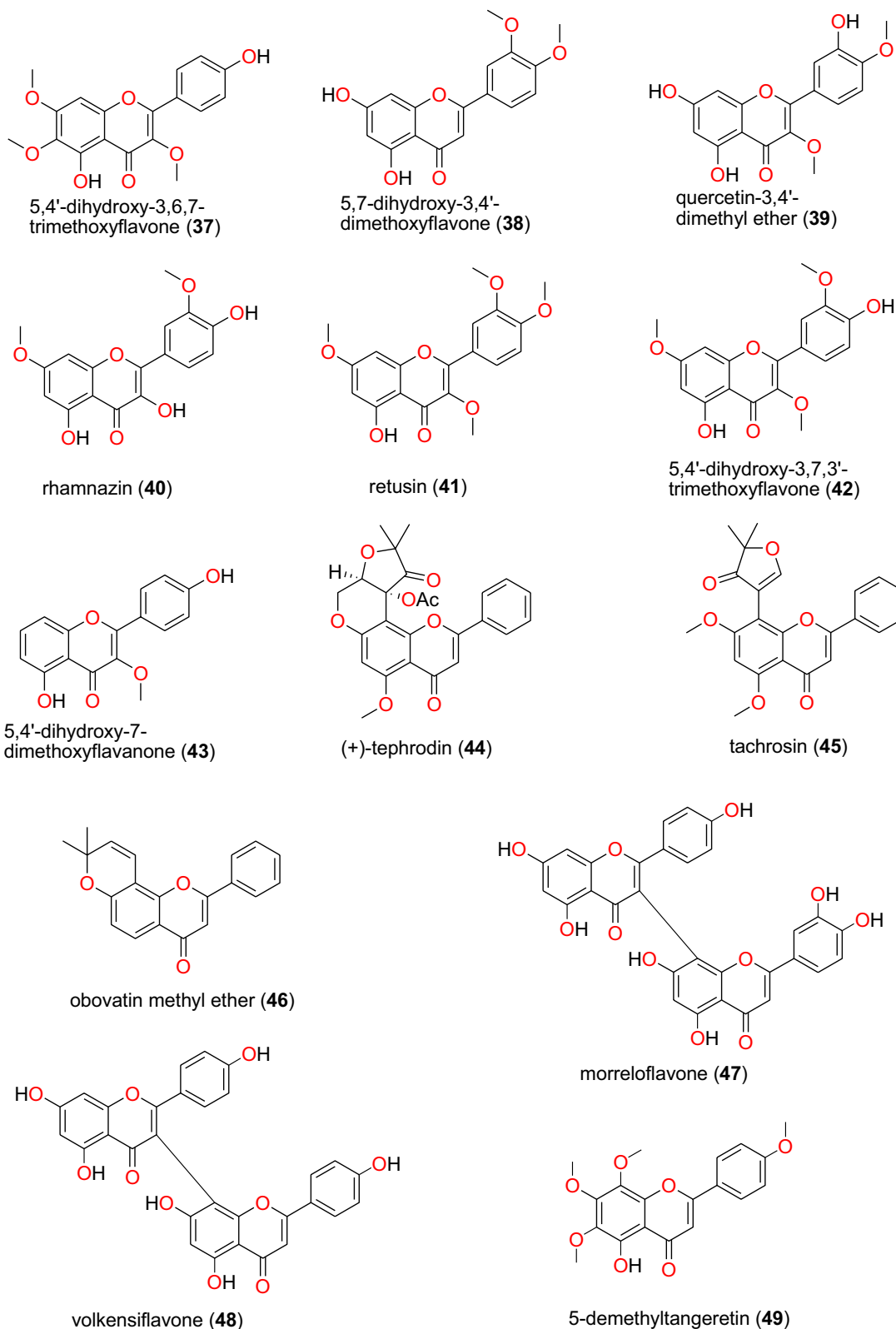
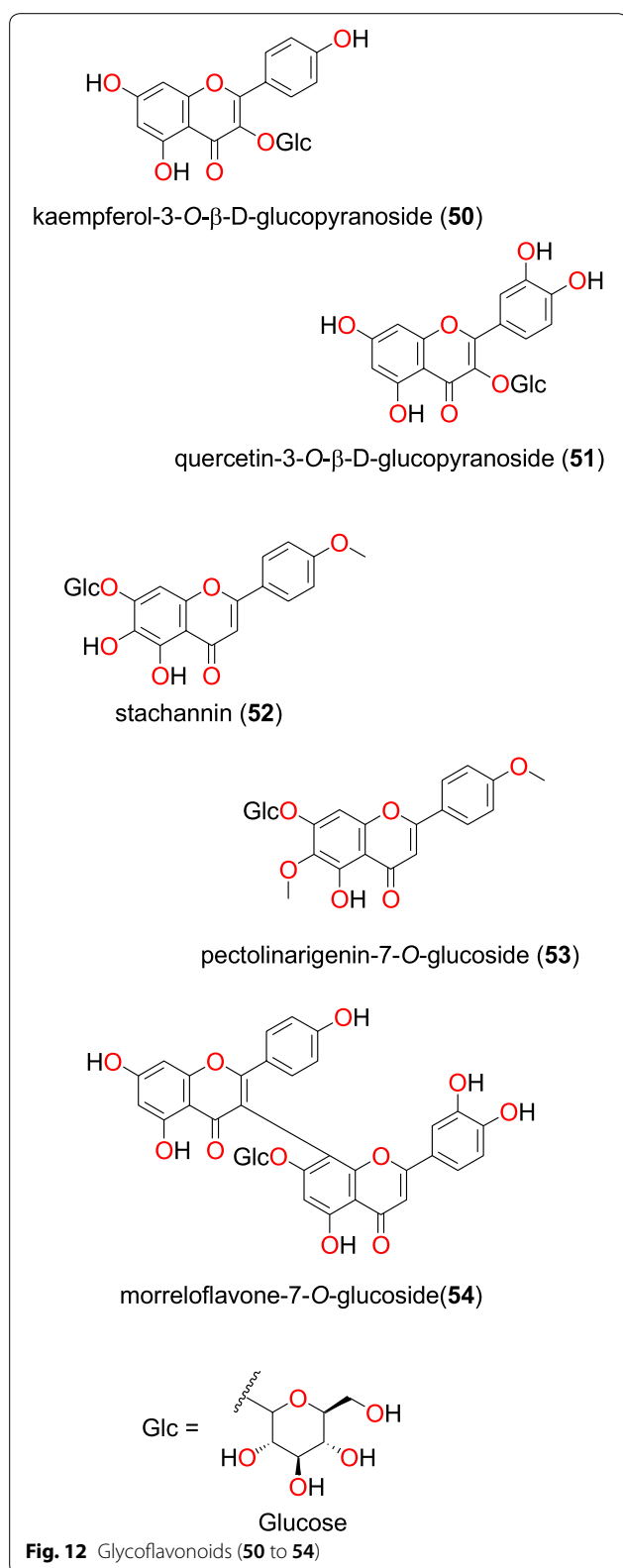


Fig. 11 Flavanones and flavones (37 to 49)

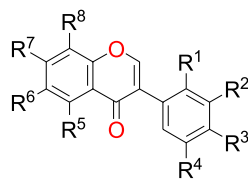


Cameroon [78]. The biflavonoids 47 and 48 were isolated from the plant extract and exhibited in vitro antiplasmodial activities against the F32 and FcM29 strains. The IC_{50} values at 24 h and 72 h against the both strains were 21.13 and 6.03 μ M; 22.59 and 8.61 μ M for compound 47 and 1.83 and 2.18 μ M; 1.44 and 1.75 μ M for compound 48, respectively [78]. Additionally, 5-demethyltangeretin (49) was isolated from the whole plant of *Peperomia vulcanica* by Ngemenya et al. [79]. This compound showed antiplasmodial activity against the multidrug-resistant W2mef and Dd2 strains of *P. falciparum*, with respective IC_{50} values of 19.36 and 3.18 μ M.

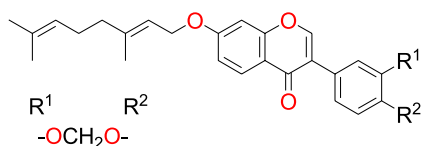
Glycoflavonoids The glycoflavonoids kaempferol-3-O- β -D-glucopyranoside (50), and quercetin-3-O- β -D-glucopyranoside (51), Fig. 12, were isolated from the leaves of *Ekebergia capensis* (Meliaceae) from Kenya by Irungu et al. [80]. Both compounds were observed to possess moderate activities against the D6 and W2 strains of *P. falciparum*. The IC_{50} values of compounds 50 and 51 were, respectively, 97.1 and 42.9 μ M against the D6 strain, while both compounds measured an IC_{50} value of 105.8 μ M against the W2 strain.

Tshitenge et al. investigated the anti-malarial constituents of the medicinal plant-based SIROP KILMA, with constitutive plants composed from *Gardenia ternifolia* (Rubiaceae), *Crossopteryx febrifuga* (Rubiaceae), and *Lantana camara* (Verbenaceae) [26]. The authors identified two flavonoid glycosides; stachannin (52) and pectolarigenin-7-O-glucoside (53) [26]. The flavone glycoside morreloflavone-7-O-glucoside (54) was isolated by Azebaze et al. from *Allanblackia floribunda* (Guttiferae), harvested in Cameroon. This compound presented antiplasmodial activities against the F32 and FcM29 strains with IC_{50} values of 15.98 and 11.69 μ M; 40.36 and 33.24 μ M at 24 h and 72 h, respectively [78].

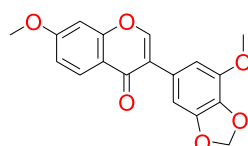
Isoflavones Studies by Derese et al. on the stem bark of *Millettia oblata* (Leguminosae) harvested from Kenya led to the isolation of 7,2'-dimethoxy-4',5'-methylenedioxyisoflavone (55), maximaisoflavone B (56), maximaisoflavone J (57), maximaisoflavone H (58), 7,3'-dimethoxy-4',5'-methylenedioxyisoflavone (59), mildurone (60), wistin (61) nordurlettone (62), 4'-prenyloxyderrone (63), isoerythrin A 4'-(3-methylbut-2-enyl) ether (64), calopogoniumisoflavone A (65), durmillone (66), jamaicin (67), isojamaicin (68), durallone (69), and 6-methoxycalopogonium isoflavone A (70) (Fig. 13) [81]. The plant extracts and isolated compounds were tested in vitro against the *P. falciparum* W2 and D6 strains. All the plant extracts had IC_{50} values ranging from 10.0 to 25.4 μ g/mL. The compounds showed good to moderate antiplasmodial activities with



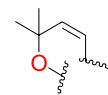
| | R ¹ | R ² | R ³ | R ⁴ | R ⁵ | R ⁶ | R ⁷ | R ⁸ |
|--|------------------|------------------|----------------------|------------------|----------------|------------------|----------------------|----------------|
| 7,2'-dimethoxy-4',5'-methylene dioxylisoflavone (55) | OCH ₃ | H | -OCH ₂ O- | H | H | H | OCH ₃ | H |
| maximaisoflavone B (56) | H | H | -OCH ₂ O- | H | H | H | O-Pry | H |
| maximaisoflavone J (57) | H | H | OCH ₃ | H | H | H | O-Pry | H |
| maximaisoflavone H (58) | H | H | OCH ₃ | H | H | H | -OCH ₂ O- | H |
| 7,3'-dimethoxy-4',5'-methylene dioxylisoflavone (59) | H | OCH ₃ | -OCH ₂ O- | H | H | H | OCH ₃ | H |
| mildurone (60) | OCH ₃ | H | -OCH ₂ O- | H | H | OCH ₃ | OCH ₃ | H |
| wistin (61) | H | H | OCH ₃ | H | H | OCH ₃ | O-β-Glc | H |
| nordurlettone (62) | H | H | O-Pry | H | H | H | OH | H |
| 4'-prenyloxiderrone (63) | H | H | O-Pry | H | OH | H | 2,2-DMC | H |
| isoerythrin A 4'-(3-methylbut-2-enyl) ether (64) | H | H | O-Pry | H | H | H | 2,2-DMC | H |
| calopogoniumisoflavone A (65) | H | H | OCH ₃ | H | H | H | 2,2-DMC | H |
| durmillone (66) | H | H | -OCH ₂ O- | H | H | OCH ₃ | 2,2-DMC | H |
| jamaicin (67) | OCH ₃ | H | -OCH ₂ O- | H | H | H | 2,2-DMC | H |
| isojamaicin (68) | H | OCH ₃ | -OCH ₂ O- | H | H | H | 2,2-DMC | H |
| durallone (69) | H | H | OCH ₃ | OCH ₃ | H | OCH ₃ | 2,2-DMC | H |
| 6-methoxycalopogonium isoflavone A (70) | H | H | OCH ₃ | H | H | OCH ₃ | 2,2-DMC | H |



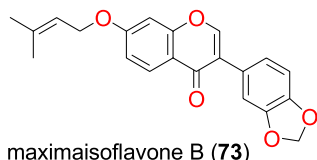
calopogonium isoflavone B (71)



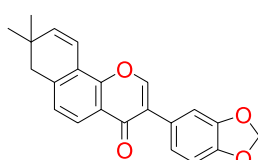
isoerythrin A-4'-(3-methylbut-2-enyl) ether (72)



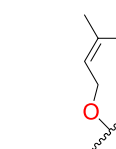
2,2DMC = 2,2-dimethylchromene



maximaisoflavone B (73)

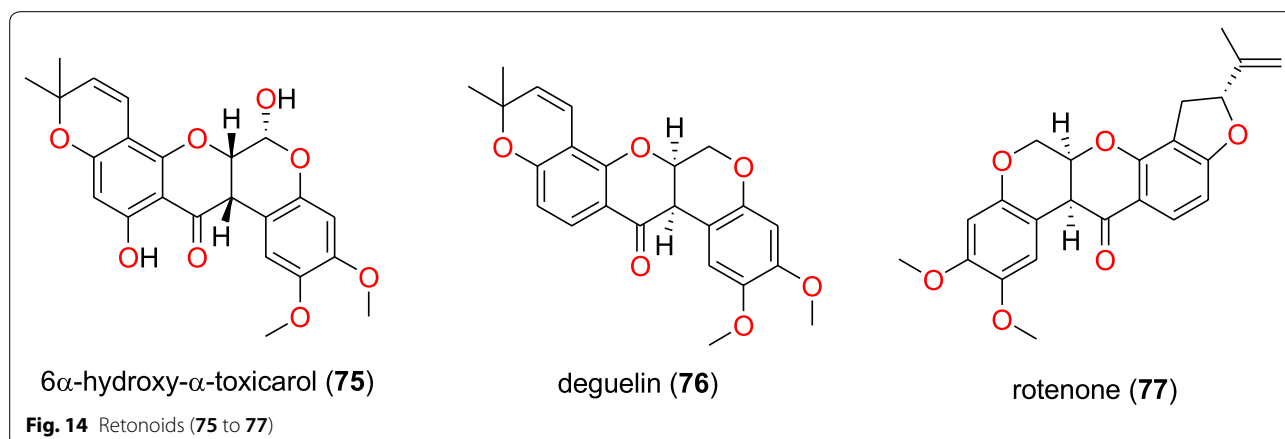


7,2'-dimethoxy-4',5'-methylenedioxyisoflavone (74)



O-Pry = O-Prenyl

Fig. 13 Isoflavones (55 to 74)



the following pairs of IC_{50} values; 45.6 and 47.5; 42.0 and 36.0; 29.7 and 35.7; 38.8 and 45.6; 48.4 and 37.7; 44.1 and 35.9; 23.2 and 22.3; 28.9 and 25.1; 14.9 and 13.3; 21.6 and 19.3; 51.5 and 45.8; 25.1 and 37.3; 38.6 and 41.0; 38.9 and 48.7; 50.0 and 32.7; 53.1 and 34.8 μ M against the W2 and D6 strains, respectively [81].

Marco et al. obtained calopogonium isoflavone B (71) and isoerythrin A-4'-(3-methylbut-2-enyl) ether (72) maximaisoflavone B (73) and 7,2'-dimethoxy-4',5'-methylenedioxyisoflavone (74) from the root bark of *Millettia dura* (Leguminosae) harvested in Tanzania [82]. These compounds showed marginal activities (70 to 90% inhibition at 40 μ M) against the 3D7 and Dd2 strains of *P. falciparum*.

Retonoids Muiva-Mutisya et al. also isolated the retonoids 6 α -hydroxy- α -toxicarol (75), deguelin (76), rotenone (77) (Fig. 14), from the root extract of *Tephrosia villosa* (Leguminosae) [77]. The mixture of compounds 76 and 77 exhibited anti-malarial activities with IC_{50} values of 9.60 and 22.60 μ g/mL against the CQ-sensitive D6 and CQ-resistant W2, respectively. Meanwhile, the activities of compound 75 against the same strains were 18.71 and 28.64 μ M, respectively [77].

Phenolics and quinones

Summaries of the phenolics and quinones with most promising anti-malarial properties have been shown in Table 4 (according to their subclasses), with chemical structures shown in Figs. 15, 16, 17, 18, 19, 20 and 21.

Ellagic acid derivatives The plant *Terminalia brownii* (Combretaceae) is used as a remedy for malaria in Eastern and Central Africa, although the detailed mode of preparation is not fully described in the literature [83]. The phenolic compound, 4-O-(3'',4''-di-O-galloyl- α -L-rhamnopyranosyl) ellagic acid (78) (Fig. 15), obtained from the stem bark of this plant harvested in Kenya was found to be active against chloroquine-sensitive (D6) and chloroquine-resistant (W2) strains of *P. falciparum* [77]. According to Machumi et al. [84], the IC_{50} value obtained against both strains was equal to 8.01 μ M.

Phenolic glycosides In addition to flavonoid glycosides, Tshitenge et al. [26] identified two phenolic glycosides as anti-malarial constituents of the medicinal plant-based SIROP KILMA: acteoside (79) and isoacteoside (80) (Fig. 16).

Anthraquinones The study by Induli et al. of the rhizomes of *Kniphofia foliosa* (Asphodelaceae), growing in Ethiopia, led to the identification of several anthraquinones [85]: the novel 10-acetonylknipholone cyclooxanthrone (81), along with the known knipholone anthrone (82), chryslandicin (83), 10-hydroxy-10-(chrysophanol-7'-yl)-chrysophanol anthrone (84), 10-methoxy-10-(chrysophanol-7'-yl) chrysophanol anthrone (85), asphodelin (86), knipholone (87), isoknipholone (88) knipholone cyclooxanthrone (89), joziknipholone A (90), joziknipholone B (91) and dianellin (92) (Fig. 17). According to the authors, the IC_{50} values obtained for the plant extracts ranged from 3.4 to 8.9 μ g/mL and from 3.4 to 8.9 μ g/mL against the D6 and W2 strains of *P. falciparum*, respectively. The obtained compounds had IC_{50} values ranging from 0.47 to 23.25 μ M and from 0.35 to 18.42 μ M against the respective strains

Table 3 Summary of flavonoids

| Compound subclass | Isolated metabolites | Plasmoidal strain (activities) | Plant species (Family), Taxon ID ^b | Part of the plant studied | Place of harvest (locality, country) | Author, references |
|-------------------------|-----------------------------|---|--|---------------------------|--------------------------------------|---------------------------|
| Flavanones and flavones | 37 to 43 | D6 (from 11.26 to 56.31 μ M) W2 (from 15.48 to 87.50 μ M) | <i>Senecio roseiflorus</i> (Compositae-Asteraceae), NCBI:txid1886451 | Leaves | Mount Kenya Forest, Meru, Kenya | Kerubo et al. [76] |
| | 44 to 46 | D6 (from 11.30 to 14.00 μ M) W2 (from 13.10 to 20.40 μ M) | <i>Tephrosia villosa</i> (Leguminosae-Fabaceae), NCBI:txid62125 | Roots | Manyani, TaitaTaveta County, Kenya | Muiva-Mutisya et al. [77] |
| | 47 and 48 | F32 (from 2.18 to 21.13 μ M) FcM29 (from 1.75 to 22.59 μ M) | <i>Allanblackia floribunda</i> (Guttiferae-Clusiaceae), NCBI:txid469914 | Whole plant | Mount Kala, Cameroon | Azebaze et al. [78] |
| | 49 | W2mef (19.37 μ M) Dd2 (3.18 μ M) | <i>Peperomia vulcanica</i> (Piperaceae), NCBI:txid1719589 | Whole plant | Mount Cameroon, Cameroon | Ngemanya et al. [79] |
| Glycoflavonoids | 50 and 51 | D6 (97.1 and 42.9 μ M, respectively) W2 (105.8 μ M) | <i>Ekebergia capensis</i> (Meliaceae), NCBI:txid124949 | Leaves | Gakoe Forest, Kiambu County, Kenya | Irungu et al. [80] |
| | 52 and 53 | | <i>Gardenia ternifolia</i> (Rubiaceae), NCBI:txid1237590; <i>Crossopteryx febrifuga</i> (Rubiaceae), NCBI:txid170354; and <i>Lantana camara</i> (Verbenaceae), NCBI:txid126435 | Stem barks and leaves | Kinshasa, DR Congo | Tshitenge et al. [26] |
| | 54 | F32 (15.98 and 40.36 μ M, respectively, at 24 h and 72 h) FcM29 (11.69 and 33.24 μ M, respectively, at 24 h and 72 h) | <i>Allanblackia floribunda</i> (Guttiferae-Clusiaceae), NCBI:txid469914 | Whole plant | Mount Kala, Cameroon | Azebaze et al. [78] |
| Isoflavones | 55 to 70 | W2 (from 14.9 to 53.1 μ M) | <i>Millettia oblata</i> ssp. <i>teitensis</i> (Leguminosae-Fabaceae), NCBI:txid53625 | Stem bark | Taita Hill Forest, Kenya | Dereese et al. [81] |
| | 63^a | D6 (from 13.3 to 48.7 μ M) | | | | |
| | 71^a to 74 | 3D7 and Dd2 (70 to 90% inhibition at 40 μ M) | <i>Millettia dura</i> (Leguminosae-Fabaceae), NCBI:txid62119 | Root bark | Kisarawe, Tanzania | Marco et al. [82] |
| Retonoids | 75 to 77 | D6 (18.71 μ M for compound 75 and 9.60 μ g/mL for a mixture of compounds 76 and 77) W2 (28.64 μ M for compound 75 and 22.60 μ g/mL for a mixture of compounds 76 and 77) | <i>Tephrosia villosa</i> (Leguminosae-Fabaceae), NCBI:txid62125 | Roots | Manyani, Taita Taveta County, Kenya | Muiva-Mutisya et al. [77] |

^a Compounds identified for the first time in the cited publications^b Identification number of the source species, derived from the NCBI Taxonomy database

Table 4 Summary of phenolics and quinones

| Compound subclass | Isolated metabolites | Plasmodial strain (activities) | Plant species (Family), Taxon ID ^b | Part of the plant studied | Place of harvest (locality, country) | Author, references |
|-------------------------------------|---|--|---|--|--------------------------------------|-----------------------|
| Ellagic acid derivative (phenolics) | 78 | D6 (8.01 μ M) W2 (8.01 μ M) | <i>Terminalia brownii</i> (Combretaceae), NCBI:txid1548809 | Stem bark | Machakos County, Kenya | Machumi et al. [84] |
| Phenolic glycosides (phenolics) | 79 and 80 | | <i>Gardenia ternifolia</i> (Rubiaceae), NCBI:txid1237590; <i>Crossop-teryx febrifuga</i> (Rubiaceae), NCBI:txid170354; and <i>Lan-tana camara</i> (Verbenaceae), NCBI:txid126435 | Stem barks and leaves | Kinshasa, DR Congo | Tshitenge et al. [26] |
| Anthraquinones (quinones) | 81^a, 82 to 92 | D6 (from 0.47 to 23.25 μ M) | <i>Kniphofia foliosa</i> (Asphodelaceae), NCBI:txid214838 | Rhizomes | Addis Ababa, Ethiopia | Induli et al. [85] |
| | 89 | W2 (from 0.35 to 18.42 μ M) | | | | |
| | 93 | D6 (7.73 μ M) W2 (2.22 μ M) | <i>Kniphofia foliosa</i> (Asphodelaceae), NCBI:txid214838 | Roots | Gedo, Ethiopia | Abdissa et al. [86] |
| | 89^a | D6 (9.40 μ M) W2 (14.58 μ M) | <i>Kniphofia foliosa</i> (Asphodelaceae), NCBI:txid214838 | Roots | Gedo, Ethiopia | Abdissa et al. [86] |
| | 82 | 3D7 (0.7 μ M) | <i>Kniphofia foliosa</i> (Asphodelaceae), NCBI:txid214838 | Leaves | Addis Ababa, Ethiopia | Feilcke et al. [89] |
| | 90^a and 91^a | K1 (0.17 and 0.26 μ M, respectively) | <i>Bulbine frutescens</i> (Asphodelaceae), NCBI:txid210954 | Roots | Chiromo Campus Garden, Kenya | Bringmann et al. [87] |
| | 94 to 96 | D6 (19.66 to 82.80 μ M) W2 (64.46 to 141.95 μ M) | <i>Aloe pulcherrima</i> (Asphodelaceae), NCBI:txid25641 | Roots | Saka Chokorsa, Ethiopia | Abdissa et al. [88] |
| | 86 | 3D7 (1.9 μ M) | <i>Kniphofia foliosa</i> (Asphodelaceae), NCBI:txid214838 | Leaves | Addis Ababa, Ethiopia | Feilcke et al. [89] |
| | 97^a | NF54 (weak activity) | <i>Diospyros canaliculata</i> (Ebenaceae), NCBI:txid13492 | Stem bark | Kribi, Cameroon | Lenta et al. [90] |
| | Anthrones (quinones) | 98 to 101 | Suppression of parasitaemia from 36.8 to 66.8% at doses of 100 to 400 mg/kg/day | <i>Aloe percrassa</i> (Asphodelaceae), NCBI:txid1593100 | Leaf latex | Edagahamus, Ethiopia |
| Naphthohydroquinones (quinones) | 102^a, 103^a, 104^a, 105^a and 106^a | D6 (from 19.59 to 36.03 μ M) W2 (60.08 to 144.43 μ M) | <i>Pentas bussei</i> (syn: <i>Rhodopentas bussei</i> , Rubiaceae), NCBI:txid387051 | Roots | Mombasa, Kenya | Endale et al. [92] |
| Other quinones | 107 | W2mef (52.25 μ M) | <i>Peperomia vulcanica</i> (Piperaceae), NCBI:txid1719589 | Whole plant | Mount Cameroon, Cameroon | Ngemenya et al. [79] |
| | 108^a | D6 (19.28 μ M) W2 (14.17 μ M) | <i>Neoboutonia macrocalyx</i> (Euphorbaceae), NCBI:txid316724 | Stem bark | Kibale National Park, Uganda | Namukobe et al. [93] |

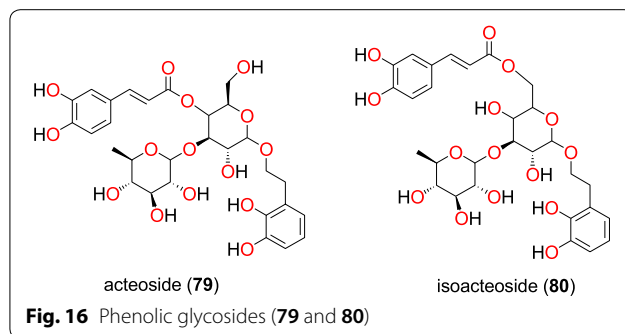
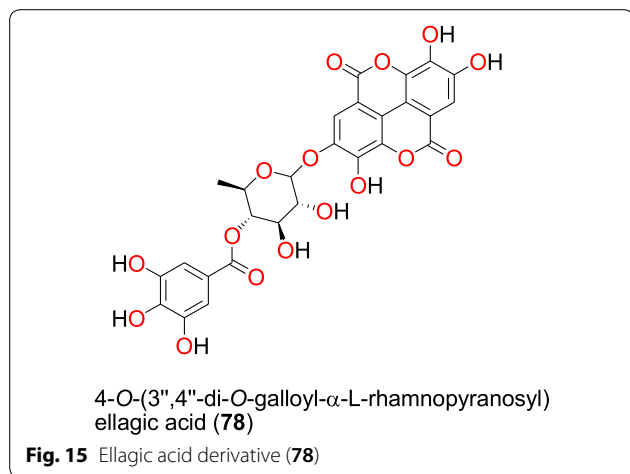
^a Compounds identified for the first time in the cited publications

^b Identification number of the source species, derived from the NCBI Taxonomy database

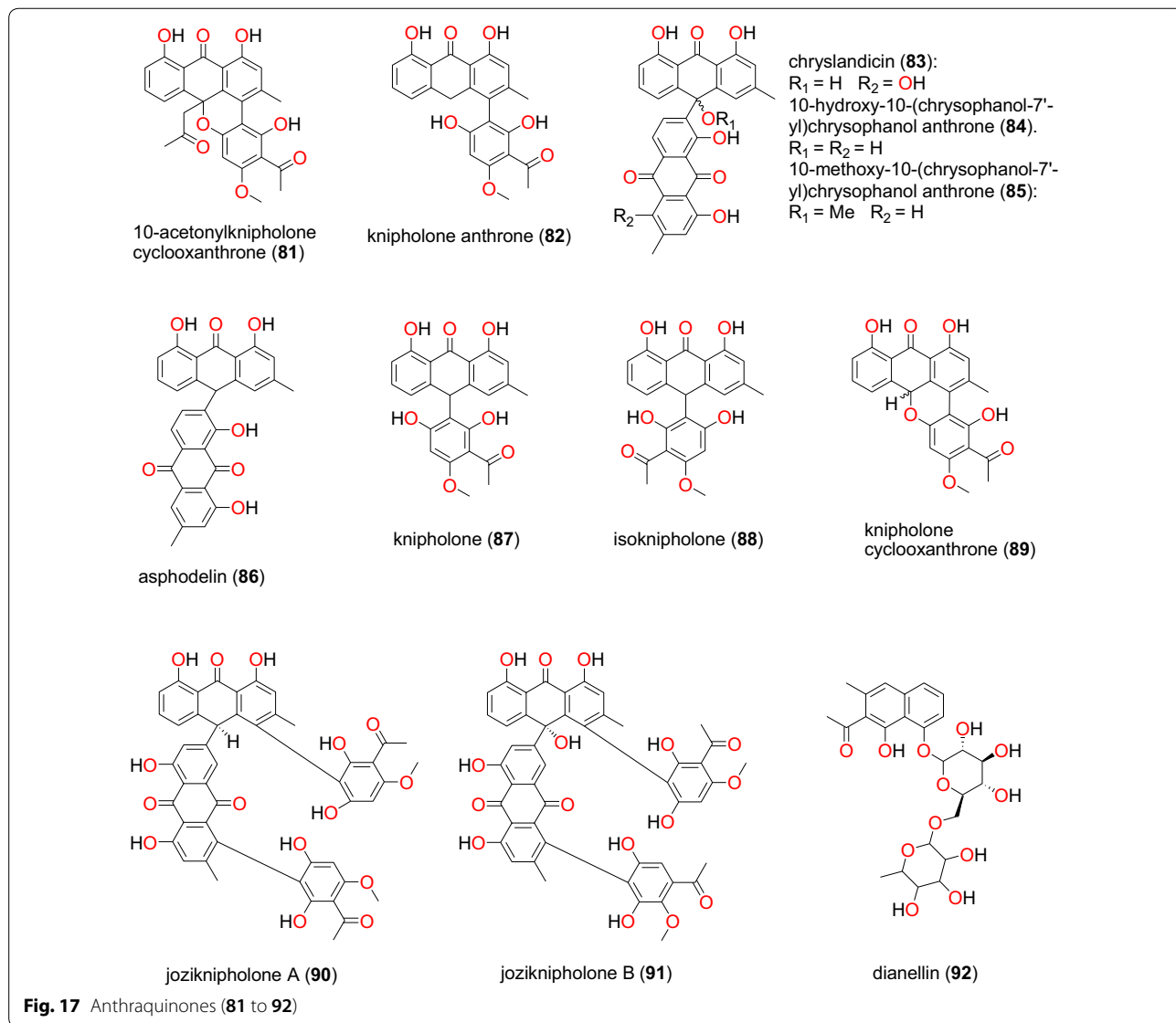
[85]. It is known that knipholone cyclooxanthrone (**89**) was actually isolated for the first time by Abdissa et al. and was also shown to exhibit antiplasmodial activities

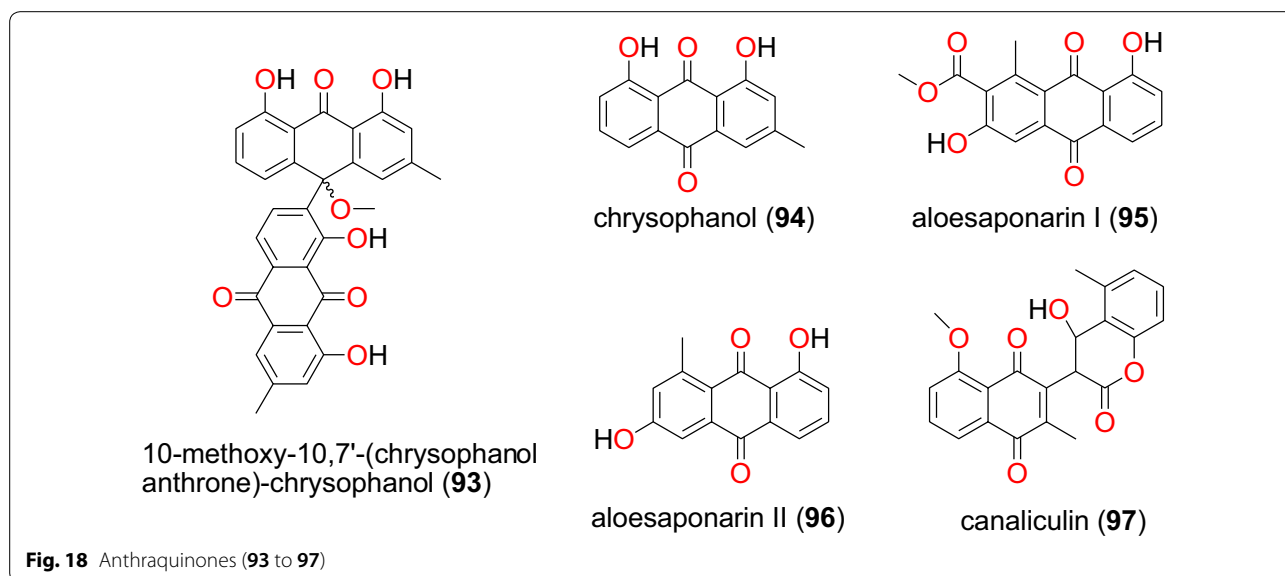
against the W2 and D6 strains with IC₅₀ values of 14.58 and 9.42 μ M, respectively [86].

The roots of the same plant, also harvested in Ethiopia, led Abdissa et al. to isolate a dimeric anthraquinone,



10-methoxy-10,7'-(chrysophanol anthrone)-chrysophanol (**93**) [86]. Compound **93** showed antiplasmodial activities against the W2 and D6 strains with IC₅₀



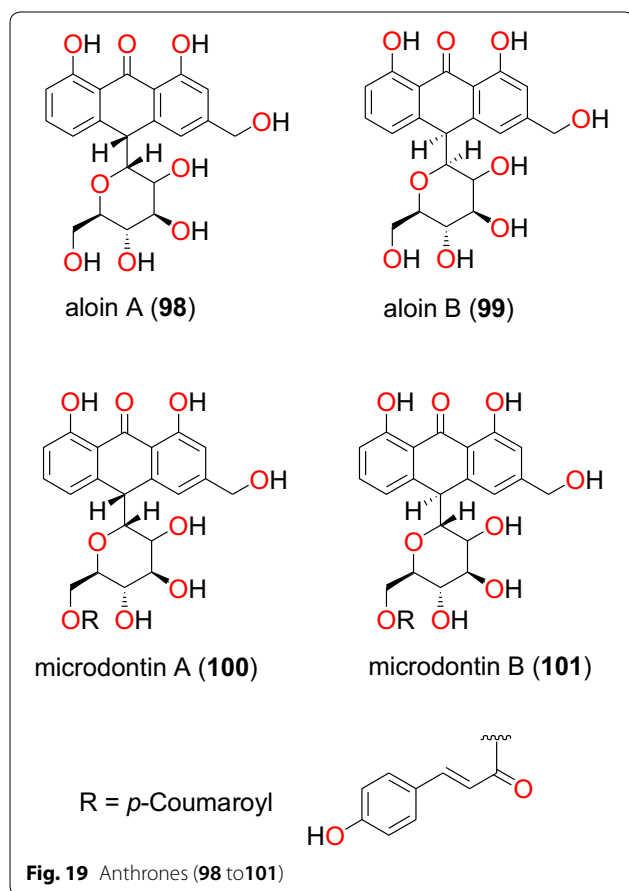


values of 2.22 and 7.73 μM , respectively [86]. The investigations of Bringmann et al. led to the isolation of new dimeric phenylanthraquinones; joziknipholone A (**90**)

and joziknipholone B (**91**) from the roots of *Bulbine frutescens* (Asphodelaceae) harvested in Kenya [87]. The authors also tested the two compounds against the K1 strain of *P. falciparum* and obtained remarkable activities, IC_{50} values of 0.17 and 0.26 μM , respectively [87]. Knipholone anthrone (**82**) was tested again from the leaves of the Ethiopian medicinal plant *Kniphofia foliosa* (Asphodelaceae) by Feilcke et al. [89]. The activity of this compound in several biological assays was described by the authors and showed antiplasmodial activity against 3D7 strain with IC_{50} value 0.7 μM .

The medicinal plant *Aloe pulcherrima* (Asphodelaceae) is one of the endemic *Aloe* species traditionally used for the treatment of malaria and wound healing in Central, Southern, and Northern Ethiopia, although the detailed mode of usage is not properly described in the literature [88]. Three compounds, chrysophanol (**94**), aloesaponarin I (**95**) and aloesaponarin II (**96**) (Fig. 18), were isolated from the acetone root extracts by Abdissa et al. [88]. The evaluation of their in vitro anti-malarial activities revealed moderate activity against D6 and W2 strains with IC_{50} values ranging from 19.66 to 82.80 μM and from 64.46 to 141.95 μM , respectively [88]. Knipholone (**86**) was also tested again from the leaves of *Kniphofia foliosa* (Asphodelaceae) by Feilcke et al. [89], showing significant antiplasmodial activity against the *P. falciparum* 3D7 strain, with an IC_{50} value of 1.9 μM .

Lenta et al. investigated the dichloromethane-methanol (1:1) extract of the stem bark of *Diospyros canaliculata* (Ebenaceae) harvested in Cameroon and obtained a new coumarinyl naphthoquinone, named canaliculin (**97**) [90]. The compound only exhibited weak activity against *P.*



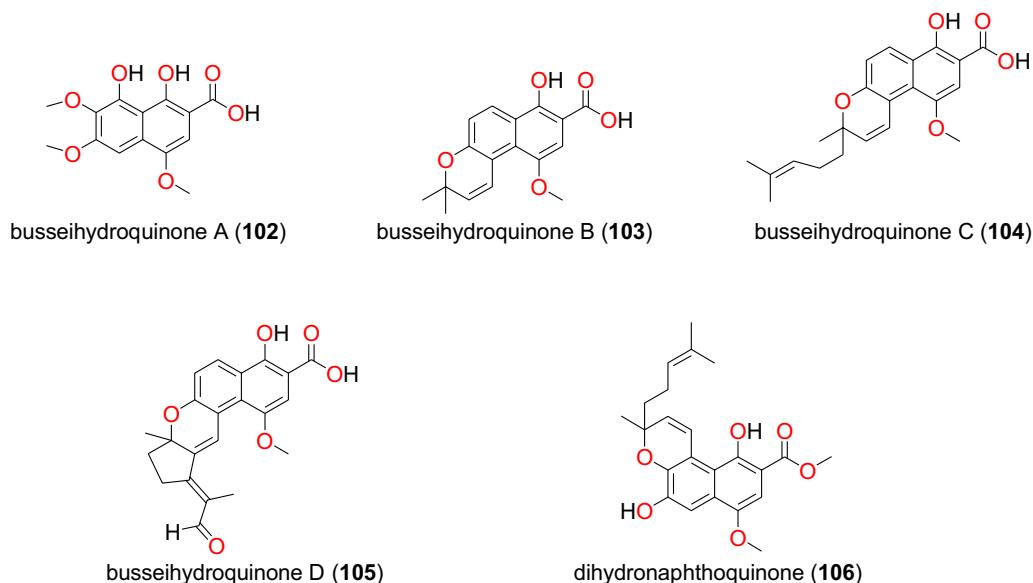


Fig. 20 Naphthoquinones (**102** to **106**)

falciparum NF54 strain, unfortunately, along with pronounced toxicity [90].

Anthrones The plant *Aloe percrassa* (Asphodelaceae), is an indigenous species used in Ethiopian folk medicine to treat malaria, wounds and gastric problems [91]. Aloin A (**98**) Aloin B (**99**) microdontin A (**100**) microdontin B (**101**) (Fig. 19), are four anthrones derived from the leaf latex of *Aloe percrassa* by Gemedhin et al. [91]. The anti-malarial activities of the mixtures of Aloin A/B and microdontin A/B were lower than the latex. The mixtures were shown to have suppressed parasitaemia from 36.8 to 66.8% at doses of 100 to 400 mg/kg/day. This suggested that the compounds within the two mixtures may have acted synergistically.

Naphthoquinones The plant species *Pentas bussei* (Rubiaceae) is frequently used in traditional medicine to treat malaria in Kenya, particularly the boiling of the roots and stems for oral consumption [92]. The roots of this species led Endale et al. to obtain five new

naphthoquinones, called busseihydroquinone A (**102**) busseihydroquinone B (**103**) busseihydroquinone C (**104**) busseihydroquinone D (**105**) and the homoprenylated naphthoquinone named dihydronaphthoquinone (**106**) (Fig. 20). These compounds exhibited marginal activities against the D6 and W2 strains with IC_{50} values ranged from 19.59 to 36.03 μ M and from 60.08 to 144.43 μ M, respectively [92].

Other quinones Peperovulcanone A (**107**), derived from the crude extracts of the whole plant of *Peperomia vulcanica* (Piperaceae), harvested from Cameroon, was shown to be active against the W2mef strain of *P. falciparum* with an IC_{50} value of 52.25 μ M [79]. The new compound named, neoboutomacrin (**108**), was derived from extracts of the stem bark of *Neoboutonia macrocalyx* (Euphorbiaceae) from Uganda by Namukobe et al. [93]. Compound **108** displayed good antiplasmodial activity with IC_{50} values of 19.28 and 14.17 μ M against the D6 and W2 strains, respectively.

Steroids

Ergostane phytosterols A summary of bioactive steroids has been provided in Table 5. The novel steroids; 6 α -methoxy-4,24(28)-ergostadiene-7 α ,20S-diol (**109**), 6 α -methoxy-4,24(28)-ergostadien-7 α -ol (**110**) (Fig. 21), along with the known steroid 7,20S-dihydroxyergosta-4,24(28)-dien-3-one (**111**) (Fig. 22), were isolated from the stem bark of *Antrocaryon klaineanum* (Anacardiaceae) by Douanla et al. [94]. The crude extracts and the isolated compounds were evaluated in vitro against the 3D7 and

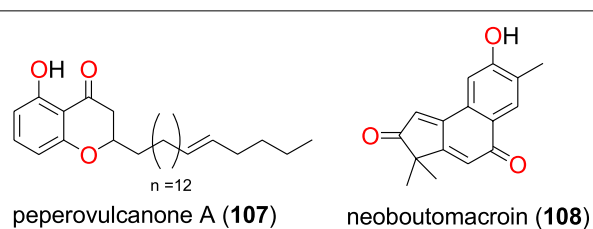


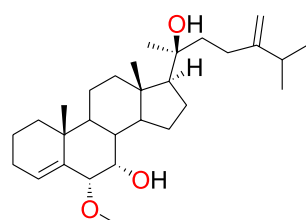
Fig. 21 Other quinones (**107** and **108**)

Table 5 Summary of steroids

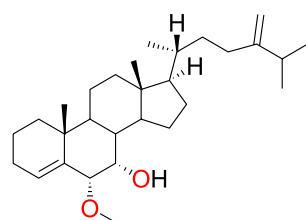
| Compound subclass | Isolated metabolites | Plasmodial strain (activities) | Plant species (Family), Taxon ID ^b | Part of the plant studied | Place of harvest (City, Country) | Author, references |
|------------------------|--|---|---|---------------------------|----------------------------------|----------------------|
| Ergostane phytosterols | 109^a, 110^a, and 111 | 3D7 (IC ₅₀ values range from 11.2 to 22.0 μM) W2 (IC ₅₀ values range from 11.2 to 22.0 μM) | <i>Antrocaryon klaineanum</i> (Anacardiaceae), NCBI:txid289695 | Stem bark | Mount Kala, Cameroon | Douanla et al. [94] |
| | 112 | W2mef (IC ₅₀ value = 53.45 μM) | <i>Peperomia vulcanica</i> (Piperaceae), NCBI:txid1719589 | Whole plant | Mount Cameroon, Cameroon | Ngemanya et al. [79] |
| | 113 | W2 (IC ₅₀ value = 153.79 μM) | <i>Polyalthia longifolium</i> var. <i>pendula</i> (Annonaceae), NCBI:txid235806 | Stem | Tikrom, near Kumasi, Ghana | Gbedema et al. [68] |
| | 113 | W2 (IC ₅₀ value = 172.9 μM) D6 (IC ₅₀ value = 68.3 μM) | <i>Turraea robusta</i> (Meliaceae), NCBI:txid1899148 | Stem bark | Nairobi, Kenya | Irungu et al. [95] |
| Phytosterol glucosides | 114 to 116 | D6 and W2 (from weak to moderate activities) | <i>Turraea nilotica</i> (Meliaceae), NCBI:txid992803 | Stem bark | Nairobi, Kenya | Irungu et al. [95] |

^a Compounds identified for the first time in the cited publications

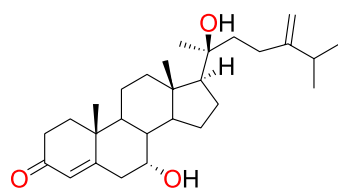
^b Identification number of the source species, derived from the NCBI Taxonomy database



6α-methoxy-4,24(28)-ergostadiene-7α,20S-diol (**109**)

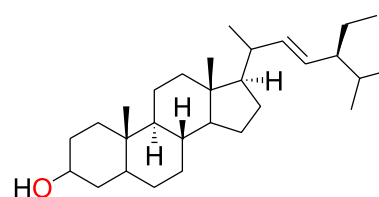


6α-methoxy-4,24(28)-ergostadien-7α-ol (**110**)

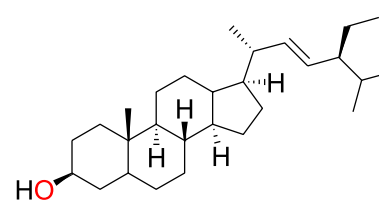


7,20S-dihydroxyergosta-4,24(28)-dien-3-one (**111**)

Fig. 22 Ergostane phytosterols (**109** to **111**)



stigmasterol (**112**)

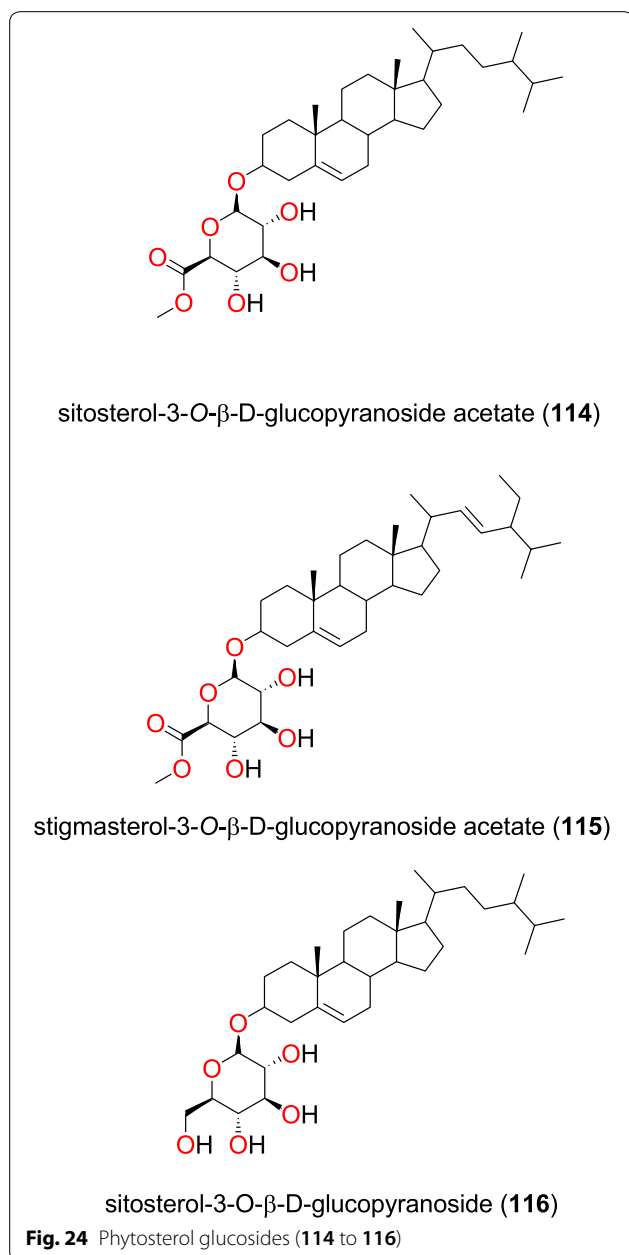


β-stigmasterol (**113**)

Fig. 23 Ergostane phytosterols (**112** and **113**)

W2 strains of *P. falciparum*. While the crude extract showed moderate activity (IC₅₀ = 16.7 μg/mL) against 3D7, the three steroids exhibited potent activity against both strains with IC₅₀ values of 22.0, 11.2 and 21.3 μM, respectively, against the same strain.

The known steroid stigmasterol (**112**) (Fig. 23), obtained from the whole plant of *Peperomia vulcanica* (Piperaceae), also showed antiplasmodial activity against



the W2mef strain with an IC_{50} value of 53.45 μ M [79]. β -stigmasterol (**113**) was also isolated from the stem of the *Polyalthia longifolium* (Annonaceae) harvested in Ghana [68]. This compound exhibited weak antiplasmodial activity against the K1, D6 and W2 strains of *P. falciparum* with IC_{50} values of 153.79, 68.3 and 172.9 μ M, respectively [68, 94].

Phytosterol glucosides The known steroid glycosides; sitosterol-3-O- β -D-glucopyranoside acetate (**114**), stigmasterol-3-O- β -D-glucopyranoside acetate (**115**),

sitosterol-3-O- β -D-glucopyranoside (**116**) (Fig. 24), as well as a mixture of β -sitosterol and stigmasterol (**112**) were identified from the leaves of *Turraea nilotica* (Meliceae) [95]. The glycosides only showed weak to moderate antiplasmodial activities against the D6 and W2 strains.

Terpenoids

The summary of the most promising diterpenoids and sesquiterpenoids has been provided in Table 6, while those of triterpenoids have been shown in Table 7.

Clerodane diterpenes The ethanolic extract of *Polyalthia longifolium* var. *pendula*, which is traditionally used to treat malaria in Ghana (the traditional preparation not properly described in the literature) displayed in vitro antiplasmodial activity against the multidrug-resistant, K1 strain with an IC_{50} value of 22.04 μ g/mL. Spectroscopic analysis of compounds obtained from this extract led to the identification of three known clerodane diterpenes; 16-hydroxycyclo-3,13(14)-dien-16,15-olide (**117**), 16-oxocyclo-3,13(14)*E*-dien-15-oic acid (**118**), and 3,16-dihydroxycyclo-4(18),13(14) *Z*-dien-15,16-olide (**119**) (Fig. 25) [68]. The compounds showed activities with IC_{50} values varying from 9.59 to 18.41 μ M.

Daphnane diterpenes The daphnane diterpenoid melerin B (**120**) was isolated from the stem bark of *Neoboutonia macrocalyx* (Euphorbaceae) and potently inhibited the CQ-resistant FcB1/Colombia strain of *P. falciparum*, with an IC_{50} value 19.02 μ M [96]. Chemical investigation of the stem bark of *Neoboutonia macrocalyx* (Euphorbiaceae) also yielded simplexin (**121**) and montanin (**122**) (Fig. 25), which showed antiplasmodial activities against the D6 and W2 strains, with IC_{50} values of 65.14 and 57.82 μ M, respectively, and 6.96 and 4.10 μ M, respectively [93].

Iridoids, labdanes, and norcassane furanoditerpenes From the aerial part of *Canthium multiflorum* (Rubiaceae) harvested in Cameroon, Kouam et al. also isolated the known iridoid, garjasmine (**123**) (Fig. 26) [69]. This compound only showed weak inhibition against the K1 strain of *P. falciparum*, with an IC_{50} value of 171.68 μ M [69].

The leaves of *Otostegia integrifolia* (Lamiaceae) are used in Ethiopian folk medicine for the treatment of several diseases including malaria [97]. The known labdane diterpenoid, otostegindiol (**124**) (Fig. 26) was isolated from the methanol leaf extract of the species by Endale et al. [97]. The isolated compound 125 displayed a significant ($p < 0.001$) anti-malarial activity at doses of 25, 50 and 100 mg/kg with chemosuppression values of 50.13, 65.58 and 73.16%, respectively. The previously reported

Table 6 Summary of diterpenoids and sesquiterpenoids

| Compound subclass | Isolated metabolites | Plasmodial strain (activities) | Plant species (Family), Taxon ID ^b | Part of the plant studied | Place of harvest (Locality, Country) | Author, references |
|----------------------------|--|--|--|---------------------------|--------------------------------------|------------------------|
| Clerodane diterpenes | 117 to 119 | K1 (IC ₅₀ values range from 9.59 to 18.41 μM) | <i>Polyalthia longifolium</i> var. <i>pendula</i> (Annonaceae), NCBI:txid235806 | Stem | Tikrom, near Kumasi, Ghana | Gbedema et al. [68] |
| Daphnane diterpenoids | 120 | FcB1 (IC ₅₀ value = 19.02 μM) | <i>Neoboutonia macrocalyx</i> (Euphorbaceae), NCBI:txid316724 | Stem bark | Kibale National Park, Uganda | Namukobe et al. [96] |
| | 121 and 122 | D6 (IC ₅₀ values = 65.14 and 6.96 μM, respectively) W2 (IC ₅₀ values = 57.82 and 4.10 μM, respectively) | <i>Neoboutonia macrocalyx</i> (Euphorbaceae), NCBI:txid316724 | Stem bark | Kibale National Park, Uganda | Namukobe et al. [93] |
| Iridoid diterpenoid | 123 | K1 (IC ₅₀ value = 171.68 μM) | <i>Canthium multiflorum</i> (Rubiaceae), NCBI:txid58501 | Aerial part | Obala, along River Sanaga, Cameroon | Kouam et al. [69] |
| Labdane diterpenoids | 124 | Suppression of <i>Plasmodium berghei</i> at doses of 25, 50 and 100 mg/kg with chemosuppression values of 50.13, 65.58 and 73.16%, respectively. | <i>Otostegia integrifolia</i> (syn: <i>Rydingia integrifolia</i> , Lamiaceae), NCBI:txid483857 | Leaves | Chancho, Central Ethiopia | Endale et al. [97] |
| Norcassane furanoditerpene | 125 | 3D7 (IC ₅₀ value = 2.20 μM) Dd2 (IC ₅₀ value = 4.16 μM) | <i>Caesalpinia bonducella</i> (Caesalpiniaceae), NCBI:txid53845 | Roots | Dar es Salaam Region, Tanzania | Nondo et al. [98] |
| Sesquiterpenoids | 126^a, 127^a, 128^a, 129^b, and 130^a | W2 (IC ₅₀ values range from 1.71 to 2.63 μM) | <i>Salacia longipes</i> (Celastraceae), NCBI:txid662028 | Seeds | Mount Kala, Cameroon | Mba'ning et al. [99] |
| | 131^a | NF54 (IC ₅₀ value = 15.69 μM) K1 (IC ₅₀ value = 13.54 μM) | <i>Scleria striatinux</i> (Cyperaceae), NCBI:txid1916803 | Rhizomes | Oku, Cameroon | Nyongbela et al. [100] |

^a Compounds identified for the first time in the cited publications

^b Identification number of the source species, derived from the NCBI Taxonomy database

norcassane furanoditerpene, norcaesalpin D (**125**), was isolated from the roots of *Caesalpinia bonducella* (Caesalpiniaceae) from Tanzania by Nondo et al. [98]. This compound was active with an IC₅₀ value of 2.20 and 4.16 μM against the 3D7 and Dd2 strains, respectively [98].

Sesquiterpenoids The novel sesquiterpenoids salaterpenes A–D (**126 to 129**), and 2β-acetoxy-1α,6β,9β-tribenzoyloxy-4β-hydroxy-dihydro-β-agarofuran (**130**) (Fig. 27), were isolated from the seeds of *Salacia longipes* (Celastraceae), harvested in Cameroon by Mba'ning et al. [99]. The investigation of their potential for anti-malarial drug discovery demonstrated that these compounds inhibited the W2 strain of *P. falciparum* with IC₅₀ values varying from 1.71 to 2.63 μM [99].

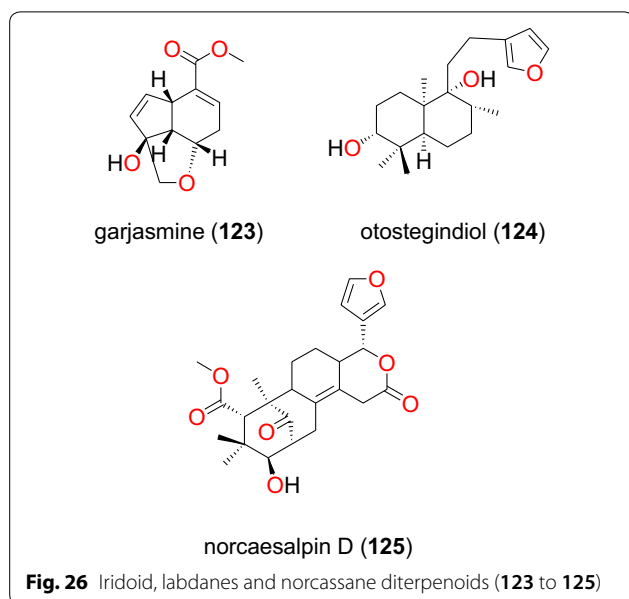
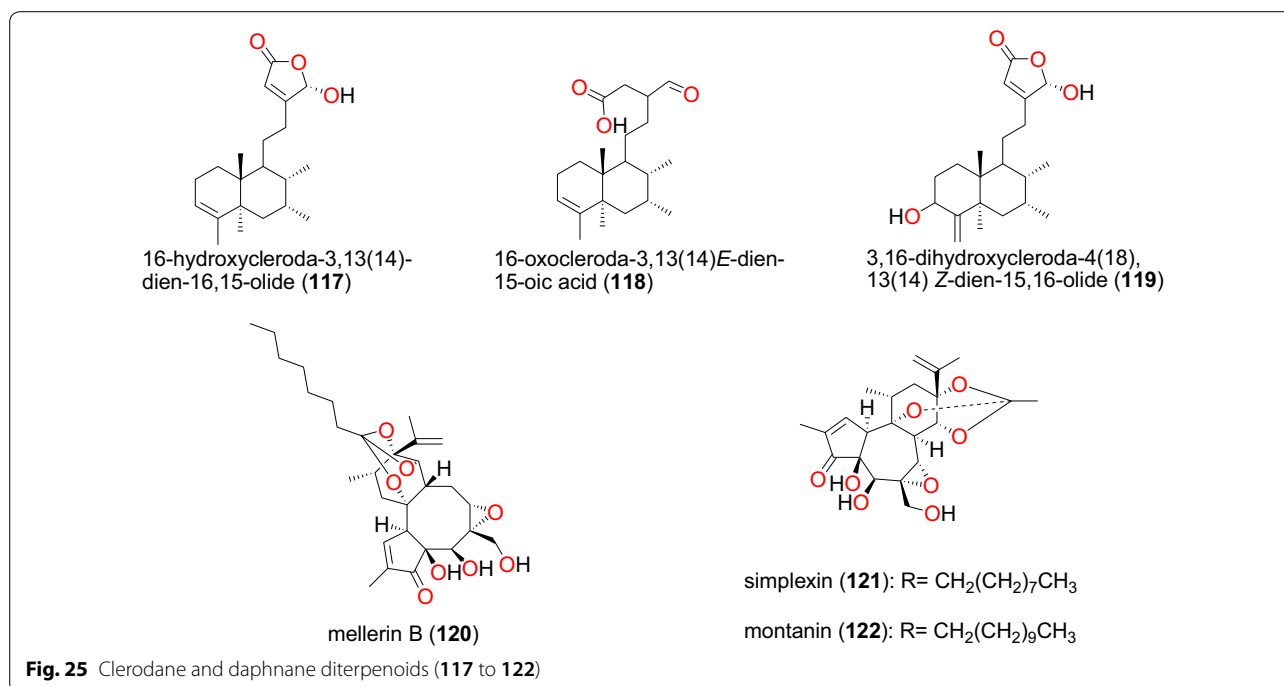
Nyongbela et al. [100] isolated the new sesquiterpene sclerienone C (**131**) from the rhizomes of *Scleria striatonux* (Cyperaceae), harvested from Cameroon. According to the authors, this compound exhibited antimicrobial and antiplasmodial activities with IC₅₀ values against the NF54 and K1 strains of 15.69 and 13.54 μM, respectively [100].

Acyclic triterpenes The previously reported acyclic triterpenes; 2-hydroxymethyl-2,3,22,23-tetrahydroxy-6,10,15,19,23-pentamethyl-6,10,14,18-tetracosatetraene (**132**) and 2,3,22,23-tetrahydroxy-2,6,10,15,19,23-hexamethyl-6,10,14,18-tetracosatetraene (**133**) were isolated from the leaves of the *Ekebergia capensis* (Meliaceae) har-

Table 7 Summary of triterpenoids

| Compound subclass | Isolated metabolites | Plasmodial strain (activities) | Plant species (Family), Taxon ID ^b | Part of the plant studied | Place of harvest (Locality, Country) | Author, references |
|--|--|--|--|---------------------------|---|--------------------------|
| Acyclic triterpenes | 132 and 133 | D6 (IC ₅₀ values = 27.1 and 56.1 μM, respectively) W2 (IC ₅₀ values = 66.9 and 64.3 μM, respectively) | <i>Ekebergia capensis</i> (Meliaceae), NCBI:txid124949 | Leaves | Gakoe forest, Kiambu County, Kenya | Irungu et al. [80] |
| Apotirucallane triterpenoids | 134 ^a , 135 ^a , 136 ^a , 137 ^a , 138 ^a , 139 ^a , and 140 to 142 | NF54 (IC ₅₀ values range from 0.67 to 19.3 μM) | <i>Entandrophragma congoense</i> (Meliaceae), NCBI:txid2590899 | Bark | Nkomokui, Cameroon | Happi et al. [101] |
| Cycloartane triterpenes | 143 to 150 ^a All new | FcB1 (all IC ₅₀ values < 11 μM, the lowest value being 1.48 μM) | <i>Neoboutonia macrocalyx</i> (Euphorbaceae), NCBI:txid316724 | Stem bark | Kibale National Park, Uganda | Namukobe et al. [96] |
| Lanostane triterpene | 151 ^a | D6 (IC ₅₀ value = 257.8 nM) W2 (IC ₅₀ value = 2000.0 nM) | <i>Ganoderma</i> sp. (Ganodermataceae), NCBI:txid5314 | Whole organism | Egypt | Wahba et al. [102] |
| Limonoids | 152 | D6 (IC ₅₀ value = 84.7 μM) W2 (IC ₅₀ value = 150.2 μM) | <i>Ekebergia capensis</i> (Meliaceae), NCBI:txid124949 | Leaves | Gakoe forest, Kiambu County, Kenya | Irungu et al. [80] |
| | 153 to 157 | D6 (IC ₅₀ values range from 2.4 to 36.6 μM) W2 (from 1.1 to 40.5 μM) | <i>Turraea robusta</i> (Meliaceae), NCBI:txid1899148 | Root bark | Nairobi, Kenya | Irungu et al. [95] |
| Oleanane triterpenes | 158 to 161 | D6 (IC ₅₀ values range from 38.8 to 205.0 μM) W2 (IC ₅₀ values range from 76.7 to 179.4 μM) | <i>Ekebergia capensis</i> (Meliaceae), NCBI:txid124949 | Leaves | Gakoe forest, Kiambu County, Kenya | Irungu et al. [80] |
| | 160 and 162 | 3D7 (IC ₅₀ values = 59.4 and 32.4 μM, respectively) | <i>Keetia leucantha</i> (Rubiaceae), NCBI:txid43504 | Twigs | Adjarrá-Ouémé, Benin Republic | Bero et al. [103] |
| | 162 , 163 and 164 | D10 (IC ₅₀ values range from 3.81 to 15.54 μM) | <i>Mimusops caffra</i> (Sapotaceae), NCBI:txid362720 | Leaves | Durban, Kwa-Zulu-Natal Province, South Africa | Simelane et al. [104] |
| Tirucallane-type triterpenoids | 165 ^a , 166 ^a and 167 | NF54 (IC ₅₀ values range from 2.4 to 6.1 μM) | <i>Entandrophragma congoense</i> (Meliaceae), NCBI:txid2590899 | Bark | Nkomokui, Cameroon | Happi et al. [105] |
| Protolimonoids | 168 to 170 | D6 (IC ₅₀ values range from 36.8 to 48.2 μM) W2 (IC ₅₀ values range from 37.2 to 77.0 μM) | <i>Turraea nilotica</i> (Meliaceae), NCBI:txid992803 | Stem bark | Nairobi, Kenya | Irungu et al. [95] |
| Other triterpenoids (hopane-type and cycloartane-type) | 171 | NF54 (IC ₅₀ value = 112.94 μM) | <i>Diospyros canaliculata</i> (Ebenaceae), NCBI:txid13492 | Stem bark | Kribi, Cameroon | Lenta et al. [90] |
| | 172 | NF54 (IC ₅₀ value = 97.73 μM) | <i>Erythrina caffra</i> (Papilionaceae), NCBI:txid3842 | Stem bark | Pietermaritzburg, South Africa | Chukwujekwu et al. [106] |
| | 173 | FcB1 (IC ₅₀ value = 2.15 μM) | <i>Neoboutonia macrocalyx</i> (Euphorbaceae), NCBI:txid316724 | Stem bark | Kibale National Park, Uganda | Namukobe et al. [96] |

^a Compounds identified for the first time in the cited publications^b Identification number of the source species, derived from the NCBI Taxonomy database

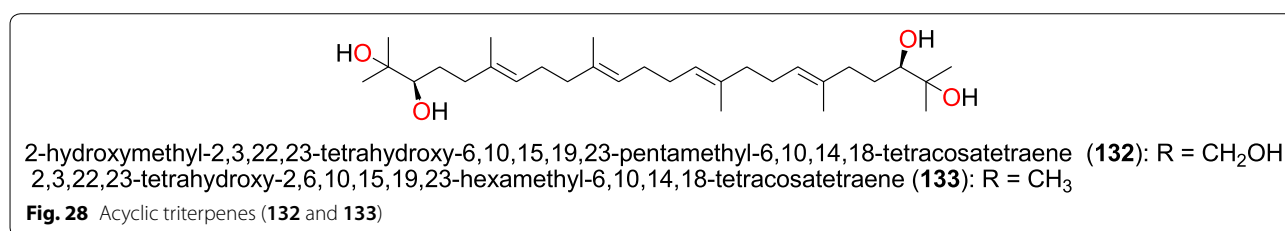
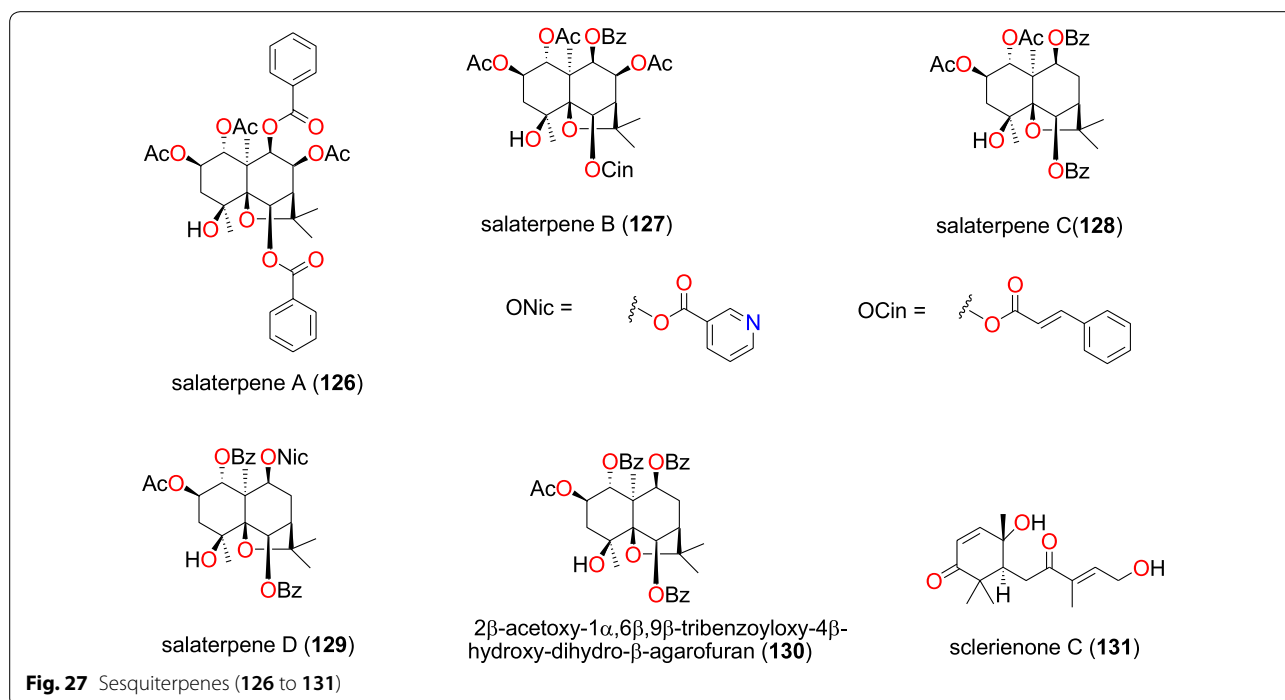


vested in Kenya [80]. The compounds (Fig. 28) exhibited selective antiparasitic activity against the W2 strain, with IC₅₀ values of 27.1 and 56.1 μM, and against the D6 66.9 and 64.3 μM, respectively [80].

Apotirucallane triterpenoids Phytochemical investigations of the root barks of *Entandrophragma congoense* (Meliaceae) harvested from Kenya, led Happi et al. to

the isolation of the novel apotirucallane triterpenoids with antiparasitic activities; prototiamins A–F (**134–139**) (Fig. 29), as well as the known lupeone (**140**), prototiamin G (**141**) and seco-tiaminic acid A (**142**) [101]. The obtained compounds (**134–142**) were also evaluated against the CQ-sensitive strain NF54. Compound **134** displayed strong selectivity for the NF54 strain against rat skeletal myoblast L6 cells (with a selectivity index of 104.7), while **136** and **138** had selective indices of 12 and 13, respectively. Compounds **135**, **137**, **139**, and **140** were active against *P. falciparum*, with IC₅₀ values ranging from 1.3 to 2.0 μM, and were less selective, while compound **142** inhibited the strain with an IC₅₀ value of 19.3 μM.

Cycloartane triterpenes The plant species *Neoboutonia macrocalyx* (Euphorbiaceae) is traditionally used to treat malaria in Southwestern Uganda around Kibale National Park, where the stem bark is widely used [96]. The investigation of the stem bark of this plant by Namukobe et al. led to the isolation nine new cycloartane triterpenes, among which eight; neomacrolactone (**143**), 22α-acetoxynemacrolactone (**144**), 6-hydroxynemacrolactone (**145**), 22α-acetoxy-6-hydroxynemacrolactone (**146**), 6,7-epoxynemacrolactone (**147**), 22α-acetoxy-6,7-epoxynemacrolactone (**148**), 4-methylenemacrolactone (**149**), and neomacroin (**150**), Fig. 30, displayed anti-malarial properties [96]. The obtained compounds were also evaluated for antiparasitic activity against



the FcB1/Colombia strain and for cytotoxicity against the KB (nasopharyngeal epidermoid carcinoma) and MRC-5 (human diploid embryonic lung) cells. Compounds (143–147, 149,150) exhibited antiplasmodial activities with IC₅₀ of < 11 μ M [96].

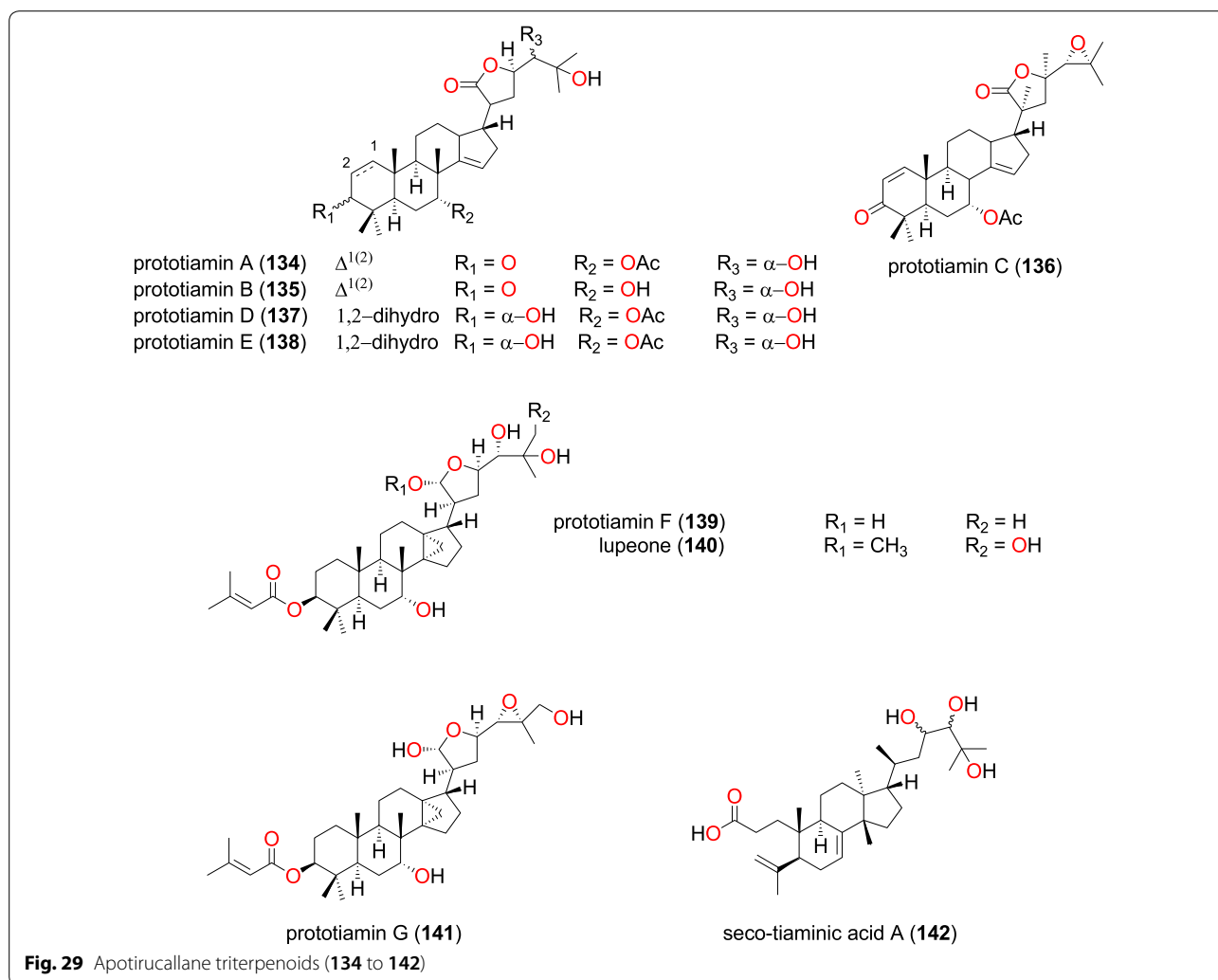
Lanostane triterpene and limonoids Ganoderic acid AW1 (151) (Fig. 31), a new lanostane triterpene, was isolated from the whole organism of the *Ganoderma* sp. (Ganodermataceae) collected from Egypt [102]. This compound exhibited good anti-malarial activity against the D6 strain of *P. falciparum* with an IC₅₀ value of 257.8 nM with no cytotoxicity up to the concentration of 9 μ M. The compound also tested positive against the W2 strain with an IC₅₀ value of 2000 nM [102].

The known limonoid, proceranolide (152) (Fig. 32) was found in the leaves of *Ekebergia capensis* (Meliaceae) by Irungu et al. [80]. The isolated compound was then

evaluated in vitro against the D6 and W2 strains of *P. falciparum*. This compound exhibited weak antiplasmodial activity against the D6 and W2 strains with IC₅₀ values of 84.7 and 150.2 μ M, respectively [80].

Additionally, three known limonoids; azadirone (153), 12 α -acetoxy-7-deacetylazadirone (154), mzikonone (155), 11-*epi*-toonacilin (156) and azadironolide (157), which were isolated from the stem bark of *Turraea nilotica* (Meliaceae), all showed potent antiplasmodial activity against the D6 and W2 strains with IC₅₀ values ranged from 2.4 to 36.6 μ M and from 1.1 to 40.5 μ M, respectively [95].

Oleanane triterpenes The known oleanonic acid (158), 3-*epi*-oleanolic acid (159), oleanolic acid (160) and ekeberin A (161) (Fig. 33) were also isolated from the leaves of *Ekebergia capensis* by Irungu et al. [80]. The four oleanane triterpenes potentially inhibited the D6 and W2 strains



of *P. falciparum* with IC_{50} values ranging from 38.8 to 205.0 μM and from 76.7 to 179.4 μM , respectively, against both strains.

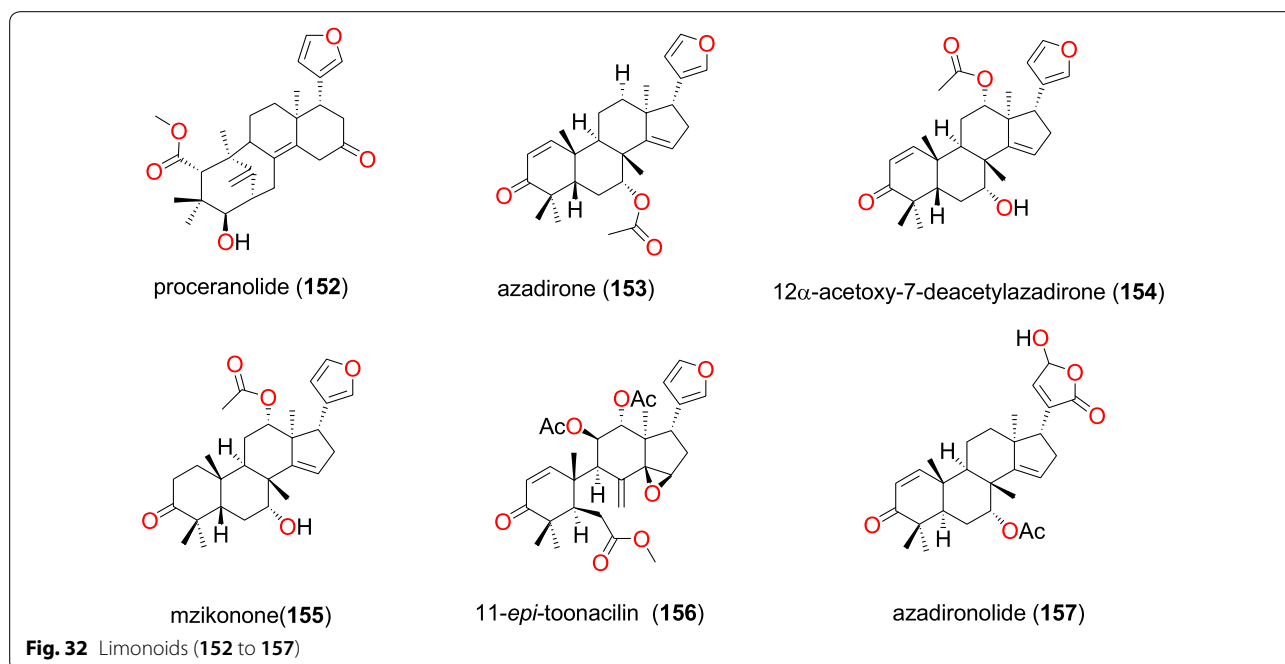
Bero et al. also isolated the known ursolic acid (162) and oleanolic acid (160) from the twigs of *Keetia leucantha* (Rubiaceae). The authors re-tested the compounds, showing them to have in vitro activities on the 3D7 strain of *P. falciparum* with IC_{50} values of 32.4 and 59.4 μM , respectively [104]. From the leaves of *Mimusops caffra* (Sapotaceae) growing in South Africa, ursolic acid acetate (163) and 3-oxo-ursolic acid (164), as well as the known compound 162 were isolated by Simelane et al. [104]. These three compounds showed promising in vitro activities against the D10 strain with IC_{50} values ranging from 3.81 to 15.54 μM [104].

Tirucallane-type triterpenoids Two new tirucallane-type triterpenoids, namely congoensin A (165) and congoensin B (166), along with the known tirucallane-type

triterpenoid gladoral A (167) (Fig. 34) were isolated from the bark of *Entandrophragma congouense* (Meliaceae) harvested from Cameroon by Happi et al. [105]. These compounds exhibited activities against the NF54 strain with IC_{50} values ranging from 2.4 to 6.1 μM [105].

Protolimonoids Irungu et al. [95] also examined the stem bark of *Turraea nilotica* (Meliaceae) growing in Kenya. Three known potent anti-malarial protolimonoids; nilotycin (168), hispidol B (169) and piscidinol A (170) were isolated (Fig. 35). These compounds exhibited activities against the D6 strain with IC_{50} values ranging from 36.8 to 48.2 μM and against the W2 strain, with IC_{50} values ranging from 37.2 to 77.0 μM [95].

Other triterpenoids The known hopane type triterpenoids; betulin (171) and lupeol (172) (Fig. 36) were isolated from the stem bark of *Diospyros canaliculata* (Ebenaceae) and *Erythrina caffra* (Papilionaceae), respectively



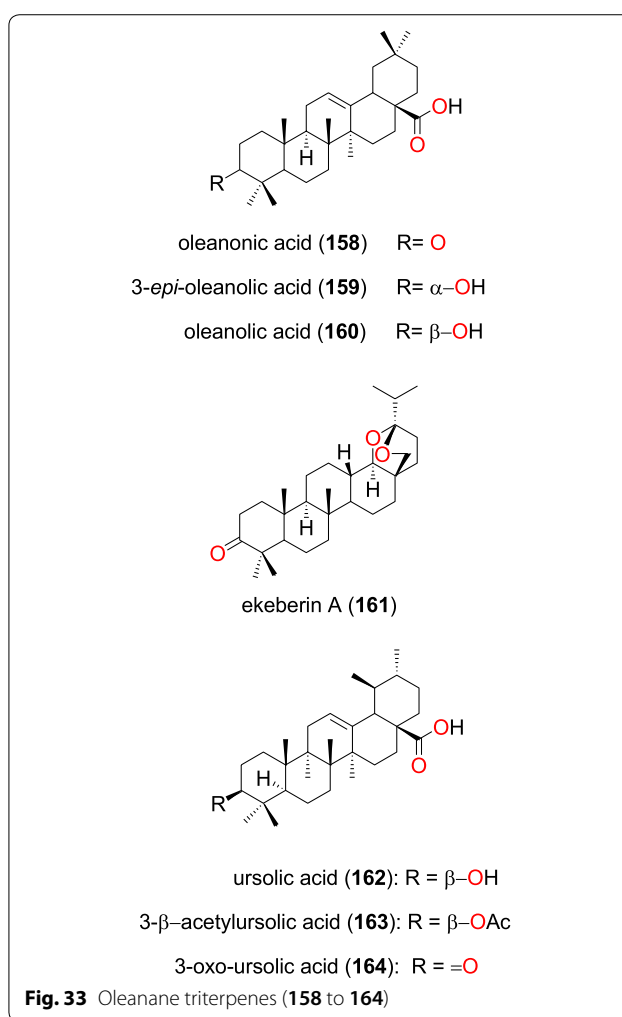
were also evaluated [109, 110]. Compound **180** only inhibited the NF54 strain weakly, with an IC_{50} value of 109.99 μ M, while compound **181** inhibited the D10 strain moderately with an IC_{50} value of 24.70 μ M [109].

Two naphthalene derivatives; dianellin (**182**) and 2-acetyl-1-hydroxy-8-methoxy-3-methylnaphthalene (**183**) isolated from the rhizomes of *Kniphofia foliosa* (Asphodelaceae) harvested in Ethiopia both inhibited the D6, W2, and 3D7 strains of *P. falciparum* with IC_{50} ranging from 6.32 to 67.32 μ M [85, 86]. The spirobisnaphthalene bipendensin (**184**) was isolated from the bark of *Entandrophragma congoense* (Meliaceae) collected in Cameroon by Happi et al. [101], this naphthalene derivative inhibiting the NF54 strain with an IC_{50} value of 73.28 μ M.

Three xanthenes; 1,7-dihydroxyxanthone (**185**), macluraxanthone (**186**) and allaxanthone B (**187**) were obtained from *Allanblackia floribunda* (Guttiferae) by Azebaze et al. [78]. The three compounds exhibited anti-plasmodial activities against the F₃₂ and FCM₂₉ strains with IC_{50} values ranging from 0.91 to 70.33 μ M for the first strain and from 0.68 to 67.22 μ M against the second [78].

Novel compounds identified and principal compound classes

It was observed that 53 out of the 187 compounds (about 28%) were described in the literature for the very first time. Besides, from Fig. 38, the majority of the NPs were terpenoids (30%), followed by flavonoids (22%), alkaloids (19%) and quinones (15%), the rest of the compound



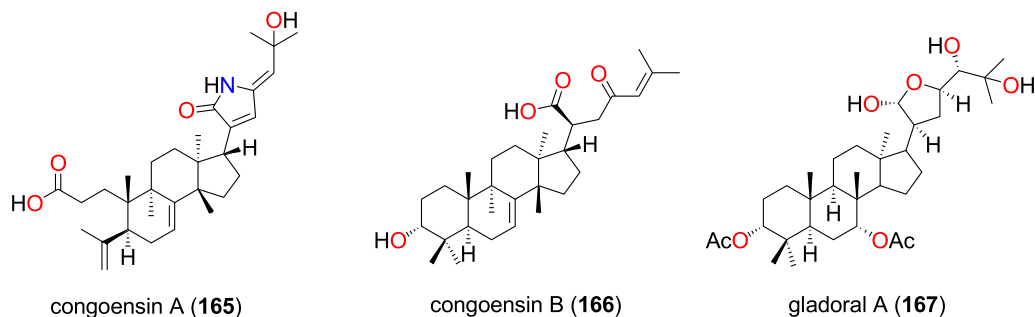


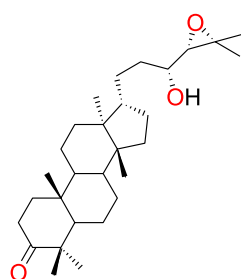
Fig. 34 Tirucallane-type triterpenoids (165 to 167)

classes, each representing only less than 5% of the entire compound collection. It was also observed that most of the plant species from which the compounds were identified were of the families Rubiaceae, Meliaceae, and Asphodelaceae (Fig. 39).

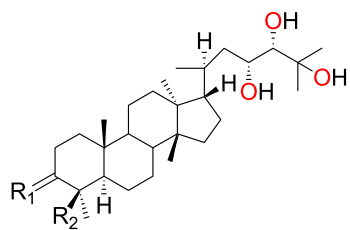
Compound distribution by plant families

A classification of the compounds by class into the plant families showed that most of the plant families represented their typical (chemotaxonomic) compound

classes, often seen in the literature for species harvested from the African continent [34, 111–114]. As an example, for the collected data (Fig. 40), all the 26 compounds from the Leguminosae-Fabaceae were flavonoids, while 23 out of the 25 anti-malarial NPs from the Asphodelaceae were quinones. It was also noted that 27 out of the 34 compounds from the Meliaceous species were terpenoids, just like the Euphorbiaceous species that included



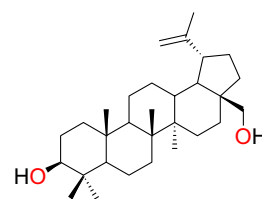
Niloticin (168)



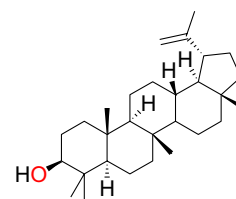
Hispidol B (169) $R_1 = \beta\text{-OH}$ $R_2 = \text{CH}_3$

Piscidinol A (170) $R_1 = =\text{O}$ $R_2 = \text{CH}_3$

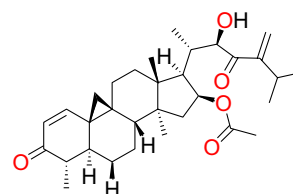
Fig. 35 Protolimonoids (168 to 170)



Betulin (171)



Lupeol (172)



22-de-O-acetyl-26-deoxyneoboutomellerone (173)

Fig. 36 Other triterpenes (171 and 173)

Table 8 Summary of other compound classes

| Compound subclass | Isolated metabolites | Plasmodial strain (activities) | Plant species (Family), Taxon ID ^a | Part of the plant studied | Place of harvest (Locality, Country) | Author, references |
|-------------------------|-------------------------|---|---|---------------------------|--|--------------------------|
| Amide | 174 | 3D7 (IC ₅₀ value = 39.04 μM) | <i>Zanthoxylum heterophyllum</i> (Rutaceae), NCBI:txid1908418 | Leaves | Langevin, Reunion Island | Ledoux et al. [107] |
| Coumarins | 175 | 3D7 (viability percentage = 101.15) | <i>Diospyros conocarpa</i> (Ebenaceae), NCBI:txid13492 | Leaves, trunk, and roots | Ntouessong and Nkoemvone, Cameroon | Fouokeng et al. [108] |
| | 176, 177 and 178 | NF54 (IC ₅₀ values vary from 2.17 to 60.09 μM) | <i>Diospyros canaliculata</i> (Ebenaceae), NCBI:txid13492 | Stem bark | Kribi, Cameroon | Lenta et al. [90] |
| Ester | 179 | NF54 (IC ₅₀ value = 42.59 μM) | <i>Erythrina caffra</i> (Papilionaceae), NCBI:txid3842 | Stem bark | Pietermaritzburg, South Africa | Chukwujekwu et al. [106] |
| Lactones | 180 | NF54 (IC ₅₀ value = 109.99 μM) | <i>Vangueria infausta</i> spp. <i>infausta</i> (Rubiaceae), NCBI:txid164485 | Roots | Mutale Municipality, Limpopo Province, South Africa | Bapela [109] |
| | 181 | D10 (IC ₅₀ value = 24.70 μM) | <i>Lippia javanica</i> (Verbenaceae), NCBI:txid925357 | Leaves | Thathe Vondo village, Limpopo Province, South Africa | Ludere et al. [110] |
| Naphthalene derivatives | 182 and 183 | D6 (IC ₅₀ value = 10.52 μM for compound 182) W2 (IC ₅₀ value = 6.32 μM for compound 182) 3D7 (IC ₅₀ value = 67.32 μM for compound 183) | <i>Kniphofia foliosa</i> (Asphodelaceae), NCBI:txid214838 | Rhizomes | Addis Ababa, Ethiopia | Induli et al. [85] |
| | 182 | D6 (IC ₅₀ value = 10.48 μM) W2 (IC ₅₀ value = 6.28 μM) | <i>Kniphofia foliosa</i> (Asphodelaceae), NCBI:txid214838 | Roots | Gedo, Ethiopia | Abdissa n [86] |
| Spirobisnaphthalene | 184 | NF54 (IC ₅₀ value = 73.28 μM) | <i>Entandrophragma congoense</i> (Meliaceae), NCBI:txid2590899 | Bark | Nkomokui, Cameroon | Happi et al. [101] |
| Xanthones | 185 to 187 | F32/24h (IC ₅₀ values range from 1.16 to 70.33 μM) | <i>Allanblackia floribunda</i> (Guttiferae- Clusiaceae), NCBI:txid469914 | Whole plant | Mount Kala, Cameroon | Azebaze et al. [78] |
| | | F32/72h (from 0.91 to 50.23 μM) | | | | |
| | | FCM29/24h (from 0.83 to 17.93 μM) | | | | |
| | | FCM29/24h (from 0.68 to 67.22 μM) | | | | |

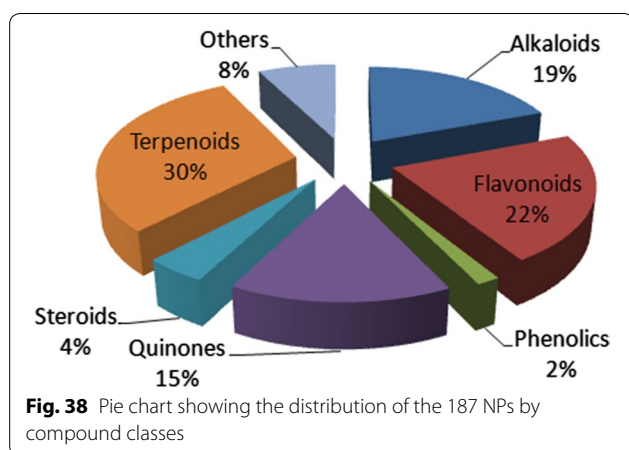
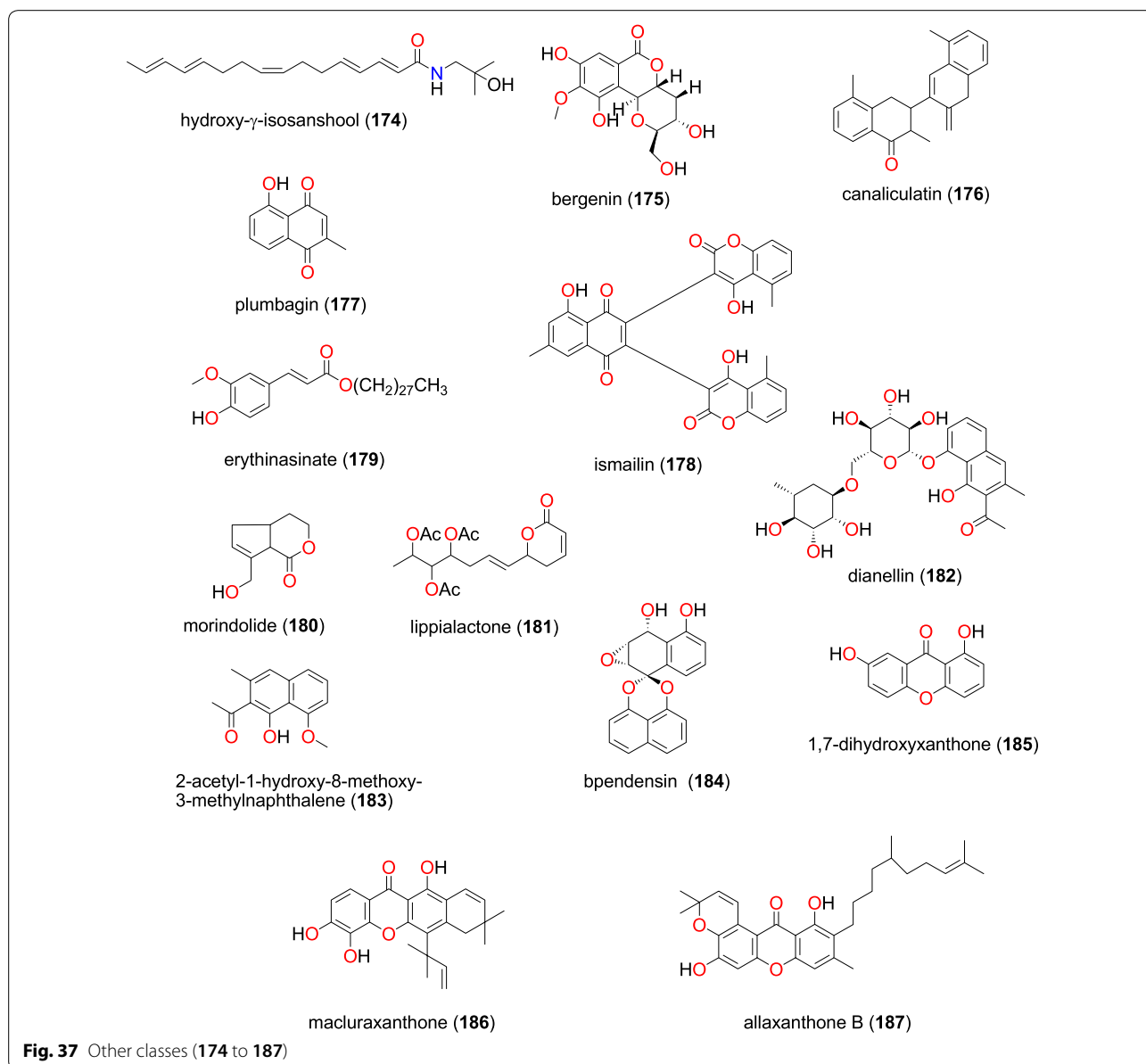
^a Identification number of the source species, derived from the NCBI Taxonomy database

12 terpenoids out of 13 compounds identified within the family. Meanwhile, all the 12 compounds from the Anacardiaceae were alkaloids, just like the Loganiaceae and Annonaceae for which all 8 compounds and 9 out of the 12 identified compounds were, respectively, alkaloids. On the contrary, the compounds from the Rubiaceae

were distributed among different classes, the majority being phenolics and quinones.

The most active compounds

Raw data retrieved from the literature showed activities reported in diverse units. A classification of the compounds by potencies (after all measured IC₅₀ values were



converted to μM), and taking a cut off of $10 \mu\text{M}$ for the most promising secondary metabolites most likely to be lead compounds. The most active compounds within this range for at least one plasmodial strain, i.e. 25 out of 66 NPs were alkaloids \sim (38%), while 23 of them were terpenoids \sim (35%) and 11 were quinones \sim (17%). Taking a cut off IC_{50} value of at most $1 \mu\text{M}$ left us with 19 compounds, 14 of them being alkaloids. Besides, the majority of the 187 NPs were terpenoids (30%), followed by flavonoids (22%), alkaloids (19%) and quinones (15%), the rest of the compound classes only represent a negligible part of the current collection.

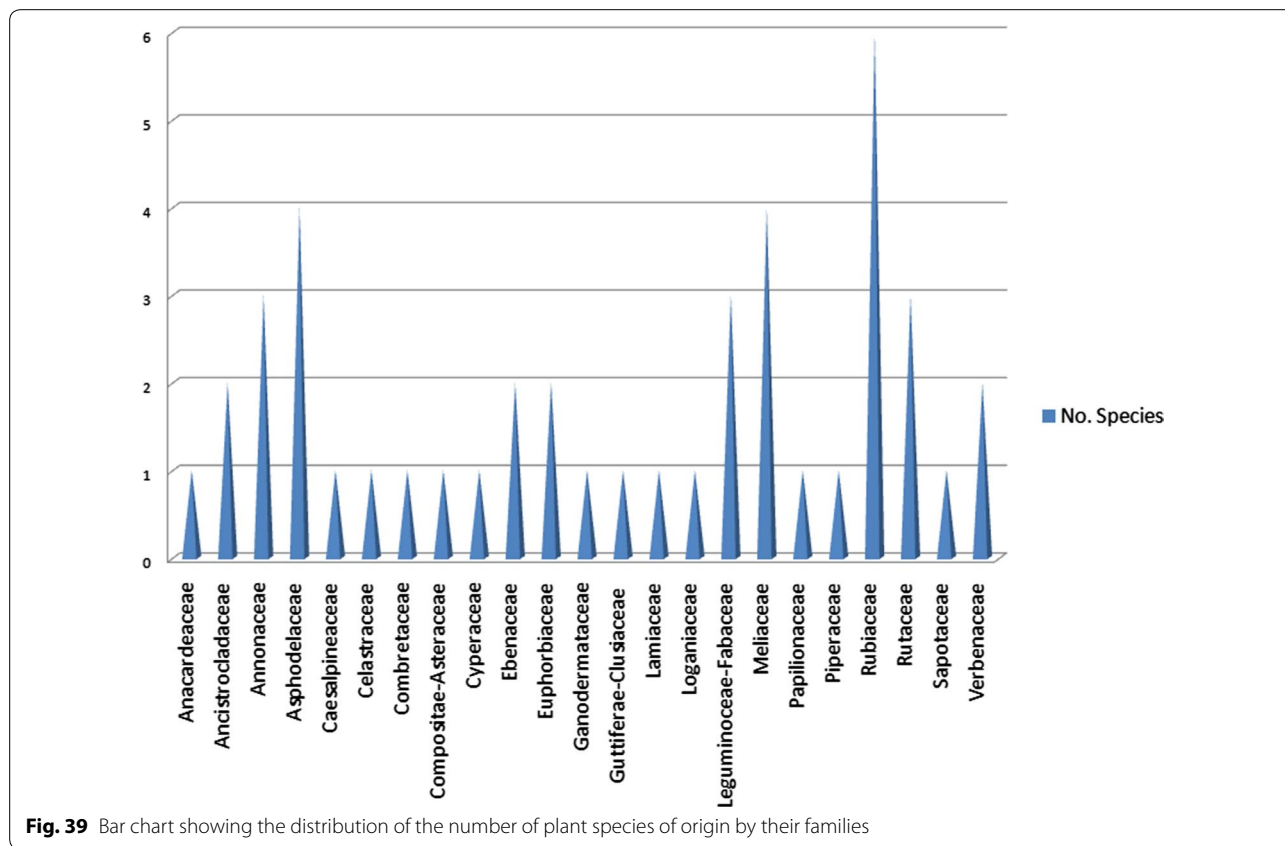


Fig. 39 Bar chart showing the distribution of the number of plant species of origin by their families

Conclusions

In this review, an attempt has been made to document the anti-malarial/antiplasmodial activities of NPs derived from African medicinal plants in their various compound classes and source species, published between 2013 and 2019. A description of the in vitro and available in vivo activities for 187 compounds is shown, as well as their classification into the various known NP compound classes and plant families of origin. From the collected data, the most active compounds belong to the same compound classes as the malarial drugs of natural origin, e.g. the alkaloid class for quinine and the terpenoid class for artemisinin. A previous report from Titanji et al. [115] had shown that plant-derived alkaloids from African medicinal plants have a great potential for anti-malarial drug development.

Although recently published reviews have described the activities of anti-malarial secondary metabolites of

terrestrial and marine origins, input data from African sources has not been the focal point. Tajuddeen and van Heerden recently published a review of 1524 natural compounds from around the world, which have been assayed against at least one strain of *Plasmodium*, out of which 39% were described as new NPs, with 29% having IC₅₀ values ≤ 3.0 μM against at least one of the tested plasmodial strains [116]. However, the study was limited to the period between 2010 and 2017 and did not include data from 2018 to 2019. Although the ability of NPs to block the transmission of malaria is still in the early stage, the current review, along with the previous studies that covers data for antiplasmodial compounds from African flora [27, 28], could serve as the baseline data for the discovery of new anti-malarial compounds from Africa.

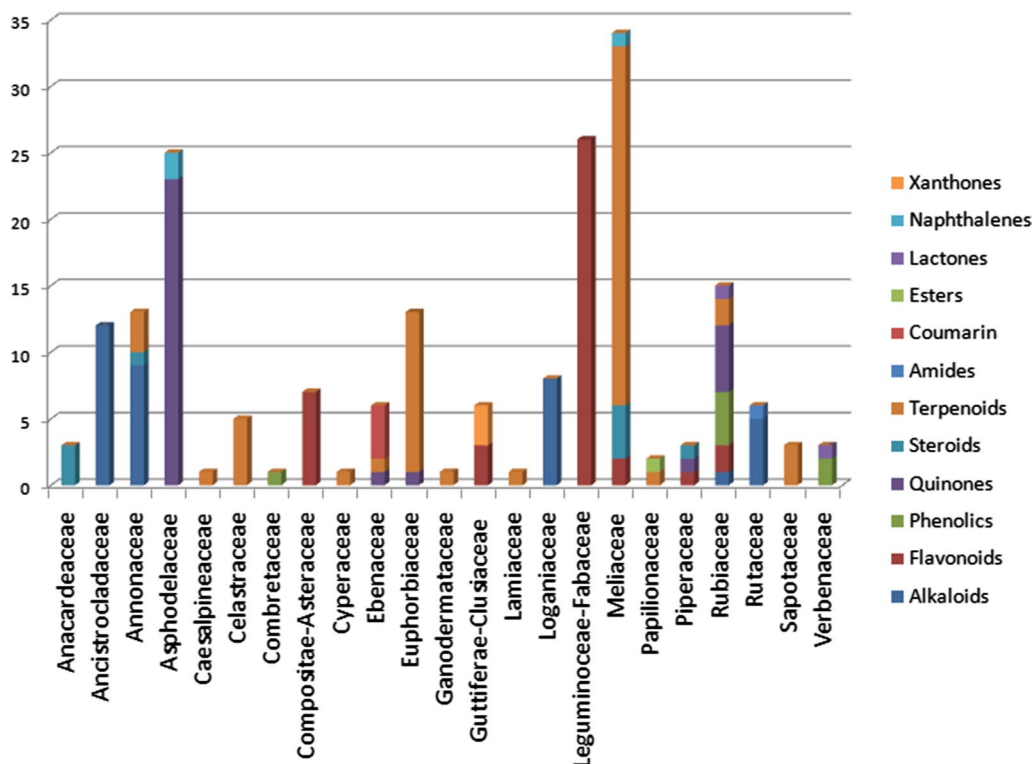


Fig. 40 An attempted chemotaxonomic distribution of compound by their classes sorted by the families of the plant species from which they were identified

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12936-020-03231-7>.

Additional file 1. List of journals consulted in building the initial data collection.

Abbreviations

ATM: African Traditional Medicine; AT: Artemisinin; CQ: Chloroquine; NP: Natural product; WHO: World Health Organization.

Acknowledgements

The authors heartily thank Dr. David J. Newman for proofreading the original manuscript.

Authors' contributions

FNK conceived the idea. BDB, FNK, and PAO participated in the data collection. BDB, FNK and PAO contributed to the data analysis, the discussion of results and the conception of the paper under the supervision of LCOO, WS, KF and LLL. BDB and FNK wrote the first draft of the paper. All authors read and approved the final manuscript.

Funding

FNK acknowledges an equipment donation by the Alexander von Humboldt Foundation and a return fellowship to Germany, accompanied by BDB. Funding by Bundesministerium für Forschung und Entwicklung (BmBF, Germany) and German Academic Exchange Service (DAAD) through the PhytoSustain/Trisustain project.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Physics, Ecole Normale Supérieure, University of Yaoundé I, P. O. Box 47, Yaoundé, Cameroon. ² Department of Chemistry, Faculty of Science, University of Buea, P. O. Box 63, Buea, Cameroon. ³ Department of Pharmaceutical Chemistry, Martin-Luther University of Halle-Wittenberg, Kurt-Mothes Str. 3, 06120 Halle (Saale), Germany. ⁴ Institut für Botanik, Technische Universität Dresden, Zellescher Weg 20b, 01062 Dresden, Germany. ⁵ Department of Chemistry, University Institute of Wood Technology Mbalmayo, University of Yaoundé I, BP 50, Mbalmayo, Cameroon. ⁶ Faculty of Natural and Environmental Sciences, Zittau/Görlitz University of Applied Sciences, Theodor-Körner-Allee 16, 02763 Zittau, Germany.

Received: 27 January 2020 Accepted: 8 April 2020

Published online: 18 May 2020

References

- WHO. World malaria report 2018. Geneva: World Health Organization; 2018. https://www.who.int/malaria/publications/world_malaria_report/en/. Accessed 27 Nov 2019.
- Mosnier E, Roux E, Cropet C, Lazrek Y, Moriceau O, Gaillet M, et al. Prevalence of *Plasmodium* spp. in the Amazonian border context (French

- Guiana-Brazil): associated factors and spatial distribution. *Am J Trop Med Hyg.* 2019;102:130–41.
3. Haeseleer C, Martiny D, Van Laethem Y, Cantinieaux B, Martin C. Reactivation of *Plasmodium* infection during a treatment with infliximab: a case report. *Int J Infect Dis.* 2019;91:101–3.
 4. Musyoka KB, Kiiru JN, Aluvaala E, Omondi P, Chege WK, Judah T, et al. Prevalence of mutations in *Plasmodium falciparum* genes associated with resistance to different antimalarial drugs in Nyando, Kisumu County in Kenya. *Infect Genet Evol.* 2019;78:104121.
 5. Sumarnrote A, Overgaard HJ, Marasri N, Fustec B, Thanispong K, Chareonviriyaphap T, et al. Status of insecticide resistance in *Anopheles* mosquitoes in Ubon Ratchathani province, Northeastern Thailand. *Malar J.* 2017;16:299.
 6. Chaumeau V, Cerqueira D, Zadrozny J, Kittiphanakun P, Andolina C, Chareonviriyaphap T, et al. Insecticide resistance in malaria vectors along the Thailand–Myanmar border. *Parasit Vectors.* 2017;10:165.
 7. Rakotoson JD, Fornadel CM, Belemvire A, Norris LC, George K, Caranci A, et al. Insecticide resistance status of three malaria vectors, *Anopheles gambiae* (s.l.), *An. funestus* and *An. mascarensis*, from the south, central and east coasts of Madagascar. *Parasit Vectors.* 2017;10:396.
 8. Hanscheid T, Schlagenhaut P, Grobusch MP. Atovaquone/proguanil for malaria chemoprophylaxis—could a difference in susceptibility during hepatic development explain the need to continue drug intake for 7 days post-exposure? *Travel Med Infect Dis.* 2019;20:101527.
 9. Schlagenhaut P, Grobusch MP, Leder K, Toovey S, Patel D. Complex choices: which malaria chemoprophylaxis can be recommended for the pregnant traveller? *Travel Med Infect Dis.* 2019;20:101525.
 10. Haston JC, Hwang J, Tan KR. Guidance for using tafenoquine for prevention and antirelapse therapy for malaria—United States, 2019. *MMWR.* 2019;68:1062–8.
 11. Hassett MR, Riegel BE, Callaghan PS, Roepe PD. Analysis of *Plasmodium vivax* chloroquine resistance transporter mutant isoforms. *Biochemistry.* 2017;56:5615–22.
 12. Zhang M, Wang C, Otto TD, Oberstaller J, Liao X, Adapa SR, et al. Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science.* 2018;360:eaap7847.
 13. Parobek CM, Parr JB, Brazeau NF, Lon C, Chaorattanakawee S, Gosi P, et al. Partner-drug resistance and population substructuring of artemisinin-resistant *Plasmodium falciparum* in Cambodia. *Genome Biol Evol.* 2017;9:1673–86.
 14. Pinheiro LCS, Feitosa LM, Gandi MO, Silveira FF, Boechat N. The development of novel compounds against malaria: quinolines, triazolopyridines, pyrazolopyridines and pyrazolopyrimidines. *Molecules.* 2019;24:E4095.
 15. Okello D, Kang Y. Exploring antimalarial herbal plants across communities in Uganda based on electronic data. *Evid Based Complement Altern Med.* 2019;2019:3057180.
 16. Ekor M. The growing use of herbal medicines: issues relating to adverse reactions and challenges in monitoring safety. *Front Pharmacol.* 2013;4:177.
 17. Bodeker C, Bodeker G, Ong CK, Grundy CK, Burford G, Shein K. WHO global atlas of traditional, complementary and alternative medicine. Geneva: World Health Organization; 2005.
 18. Alebie G, Urga B, Worku A. Systematic review on traditional medicinal plants used for the treatment of malaria in Ethiopia: trends and perspectives. *Malar J.* 2017;16:307.
 19. Esmaili S, Ghiaee A, Naghibi F, Mosaddegh M. Antiplasmodial activity and cytotoxicity of plants used in traditional medicine of Iran for the treatment of fever. *Iran J Pharm Res.* 2015;14:103–7.
 20. Mukungu N, Abuga K, Okalebo F, Ingwela R, Mwangi J. Medicinal plants used for management of malaria among the Luhya community of Kakamega East sub-County, Kenya. *J Ethnopharmacol.* 2016;194:98–107.
 21. Appiah KS, Oppong CP, Mardani HK, Omari RA, Kpabitey S, Amoatey CA, et al. Medicinal plants used in the Ejisu-Juaben Municipality, Southern Ghana: an ethnobotanical study. *Medicines (Basel).* 2019;6:1.
 22. Salehi B, Zakaria ZA, Gyawali R, Ibrahim SA, Rajkovic J, Shinwari ZK, et al. *Piper* species: a comprehensive review on their phytochemistry, biological activities and applications. *Molecules.* 2019;24:1364.
 23. Pan W-H, Xu X-Y, Shi N, Tsang SW, Zhang H-J. Antimalarial activity of plant metabolites. *Int J Mol Sci.* 2018;19:1382.
 24. Youmsi RDF, Fokou PVT, Menkem EZ, Bakarnga-Via I, Keumoe R, Nana V, et al. Ethnobotanical survey of medicinal plants used as insects repellents in six malaria endemic localities of Cameroon. *J Ethnobiol Ethnomed.* 2017;13:33.
 25. Karar MGE, Kuhnert N. Herbal drugs from Sudan: traditional uses and phytoconstituents. *Pharmacogn Rev.* 2017;11:83–103.
 26. Tshitenge DT, Ioset KN, Lami JN, Ndelo-di-Phanzu J, Mufusama J-PKS, Bringmann G. Rational quality assessment procedure for less-investigated herbal medicines: case of a Congolese antimalarial drug with an analytical report. *Fitoterapia.* 2016;110:189–95.
 27. Onguéné PA, Ntie-Kang F, Lifongo LL, Ndom JC, Sippl W, Mbaze LM. The potential of anti-malarial compounds derived from African medicinal plants, part I: a pharmacological evaluation of alkaloids and terpenoids. *Malar J.* 2013;12:449.
 28. Ntie-Kang F, Onguéné PA, Lifongo LL, Ndom JC, Sippl W, Mbaze LM. The potential of anti-malarial compounds derived from African medicinal plants, part II: a pharmacological evaluation of non-alkaloids and non-terpenoids. *Malar J.* 2014;13:81.
 29. Mahmoudi N, de Julian-Ortiz JV, Cicerone L, Galvez J, Mazier D, Danism M, et al. Identification of new antimalarial drugs by linear discriminant analysis and topological virtual screening. *J Antimicrob Chemother.* 2006;57:489–97.
 30. Willcox M, Bodeker G, Rasanaiwo P. Traditional medicinal plants and malaria. Boca Raton: CRC Press; 2004.
 31. Rasoanaivo P, Oketch-Rabah H. Preclinical considerations on anti-malarial phytomedicines. Part II, Efficacy evaluation. Antananarivo: Institut Malgache de Recherches Appliquées; 1998.
 32. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 2001;46:3–26.
 33. Onguéné PA, Ntie-Kang F, Mbah JA, Lifongo LL, Ndom JC, Sippl W, et al. The potential of anti-malarial compounds derived from African medicinal plants, part III: an in silico evaluation of drug metabolism and pharmacokinetics profiling. *Org Med Chem Lett.* 2014;4:6.
 34. Ntie-Kang F, Mbah JA, Mbaze LM, Lifongo LL, Scharfe M, Ngo Hanna J, et al. CamMedNP: building the Cameroonian 3D structural natural products database for virtual screening. *BMC Complement Altern Med.* 2013;13:88.
 35. Ntie-Kang F, Onguéné PA, Scharfe M, Owono LCO, Megnassan E, Mbaze LM, et al. ConMedNP: a natural product library from Central African medicinal plants for drug discovery. *RSC Adv.* 2014;4:409–19.
 36. Ntie-Kang F, Zofou D, Babiaka SB, Meudom R, Scharfe M, Lifongo LL, et al. AfroDb: a select highly potent and diverse natural product library from African medicinal plants. *PLoS ONE.* 2013;8:e78085.
 37. Ntie-Kang F, Telukunta KK, Döring K, Simoben CV, Moumbock AFA, Malange YI, et al. NANPDB: a resource for natural products from Northern African sources. *J Nat Prod.* 2017;80:2067–76.
 38. Organization for Economic Growth and Development (OECD). OECD guidelines for the testing of chemicals: acute oral toxicity up and down-procedure (UDP). 2008; 1–27.
 39. Dikasso D, Makonnen E, Debella A, Abebe D, Urga K, Makonnen W, et al. *In vivo* antimalarial activity of hydroalcoholic extracts from *Asparagus africanus* Lam. in mice infected with *Plasmodium berghei*. *Ethiop J Health Dev.* 2006;280:112–8.
 40. Akuodor GC, Idris-Uzman M, Anyalewechi N, Eucheria O, Ugwu CT, Akpan JL, et al. *In vivo* antimalarial activity of ethanolic leaf extract of *Verbena hastata* against *Plasmodium berghei* in mice. *J Herb Med Toxicol.* 2010;4:17–23.
 41. Desjardins RE, Canfield CJ, Haynes JD, Chulay JD. Quantitative assessment of antimalarial activity in vitro by a semiautomated microdilution technique. *Antimicrob Agents Chemother.* 1979;16:710–8.
 42. Trager W, Jensen JB. Human malaria parasites in continuous culture. *Science.* 1976;193:673–5.
 43. Duffy S, Avery VM. Development and optimization of a novel 384-well antimalarial imaging assay validated for high-throughput screening. *Am J Trop Med Hyg.* 2012;86:84–92.
 44. Novac O, Guenier AS, Pelletier J. Inhibitors of protein synthesis identified by a high throughput multiplexed translation screen. *Nucleic Acids Res.* 2004;32:902–5.

45. Kigundu EV, Rukunga GM, Keriko JM, Tonui WK, Gathirwa JW, Kirira PG, et al. Antiparasitic activity and cytotoxicity of selected medicinal plants from Kenya. *J Ethnopharmacol.* 2009;123:504–9.
46. Gathirwa JW, Rukunga GM, Njagi ENM, Omar SA, Mwitari PG, Guantai AN, et al. The in vitro antiparasitic and in vivo antimalarial efficacy of combinations of some medicinal plants used traditionally for treatment of malaria by the Meru community in Kenya. *J Ethnopharmacol.* 2008;115:223–31.
47. Samoylenko V, Jacob MR, Khan SI, Zhao J, Tekwani BL, Midiwo JO, et al. Antimicrobial, antiparasitic and cytotoxic spermine alkaloids from *Albizia schimperiana*. *Nat Prod Commun.* 2009;4:791–6.
48. Makler MT, Ries JM, Williams JA, Bancroft JE, Piper RC, Gibbins BL, et al. Parasite lactate dehydrogenase as an assay for *Plasmodium falciparum* drug sensitivity. *Am J Trop Med Hyg.* 1993;48:739–41.
49. Smilkstein M, Sriwilajjaron N, Kelly JX, Wilairat P, Riscoe M. Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening. *Antimicrob Agents Chemother.* 2004;48:1803–6.
50. Juma WP, Akala HM, Eyase FL, Muiva LM, Heydenreich M, Okalebo FA, et al. Terpurinflavone: antiparasitoid flavones from the stem of *Tephrosia purpurea*. *Phytochem Lett.* 2011;4:176–8.
51. Johnson JA, Denuil RA, Gerena L, Lopez-Sanchez M, Roncal NE, Waters NC. Assessment and continued validation of the malaria SYBR Green I-based fluorescence assay for use in malaria drug screening. *Antimicrob Agents Chemother.* 2007;51:1926–33.
52. Makler MT, Hinrichs DJ. Measurement of the lactate dehydrogenase activity of *Plasmodium falciparum* as an assessment of parasitemia. *Am J Trop Med Hyg.* 1993;48:205–10.
53. Matile H, Pink JRL. Chapter 15: *Plasmodium falciparum* malaria parasite cultures and their use in immunology. In: Lefkowitz I, Pernis B, editors. *Immunological methods*, vol. 4. San Diego: Academic Press; 1990. p. 221–34.
54. Thaithong S, Beale GH, Chutmongkonkul M. Susceptibility of *Plasmodium falciparum* to five drugs: an in vitro study of isolates mainly from Thailand. *Trans R Soc Trop Med Hyg.* 1983;77:228–31.
55. Ahmed SA, Gogal RM, Walsh JE. A new rapid simple non-radioactive assay to monitor and determine the proliferation of lymphocytes: an alternative to [³H]thymidine incorporation assay. *J Immunol Methods.* 1994;170:211–24.
56. Ponnudurai T, Leeuwenberg AD, Meuwissen JH. Chloroquine sensitivity of isolates of *Plasmodium falciparum* adapted to in vitro culture. *Trop Geogr Med.* 1981;33:50–4.
57. Malik S, Khan S, Das A, Samantaray JC. *Plasmodium* lactate dehydrogenase assay to detect malarial parasites. *Natl Med J India.* 2004;17:237–9.
58. Malebo HM, Wenzler T, Cal M, Swaleh SM, Omolo MO, Hassanali A, et al. Anti-protozoal activity of aporphine and protoberberine alkaloids from *Annickia kummeriae* (Engl. & Diels) Setten & Maas (Annonaceae). *BMC Complement Altern Med.* 2013;13:48.
59. Lacroix D, Prado S, Kamoga D, Kasenene J, Bodo B. Absolute configuration of 2'(*R*)-acetylmontrifoline and 2'(*R*)-montrifoline, furoquinolines from the fruits of *Teclea nobilis*. *Phytochem Lett.* 2012;5:22–5.
60. Rasamison VE, Brodie PJ, Merino EF, Cassera MB, Ratsimbason MA, Rakotonandrasana S, et al. Furoquinoline alkaloids and methoxyflavones from the stem bark of *Melicope madagascariensis* (Baker) T.G. Hartley. *Nat Prod Bioprospect.* 2016;6:261–5.
61. Tchinda AT, Tamze V, Ngono ARN, Ayimele GA, Cao M, Angenot L, et al. Alkaloids from the stem bark of *Strychnos icaja*. *Phytochem Lett.* 2012;5:108–13.
62. Beaufay C, Ledoux A, Jansen O, Bordignon A, Zhao S, Tejjaro CN, et al. In vivo antimalarial and antitrypanosomal activity of strychnogucine B, a bisindole alkaloid from *Strychnos icaja*. *Planta Med.* 2018;84:881–5.
63. Frédéric M, De Pauw M-C, Prosperi C, Tits M, Brandt V, Penelle J, et al. Strychnogucines A and B, two new antiparasitoid bisindole alkaloids from *Strychnos icaja*. *J Nat Prod.* 2001;64:12–6.
64. Kouam SF, Ngouonpe AW, Lamshöft M, Talontsi FM, Bauer JO, Strohmman C, et al. Indolosesquiterpene alkaloids from the Cameroonian medicinal plant *Polyalthia oliveri* (Annonaceae). *Phytochemistry.* 2014;105:52–9.
65. Lombe BK, Bruhn T, Feineis D, Mudogo V, Brun R, Bringmann G. Cyclom-bandakamines A1 and A2, oxygen-bridged naphthylisoquinoline dimers from a Congolese *Ancistrocladus* liana. *Org Lett.* 2017;19:1342–5.
66. Li J, Seupel R, Feineis D, Mudogo V, Kaiser M, Brun R, et al. Dioncophyllines C2, D2, and F and related naphthylisoquinoline alkaloids from the Congolese liana *Ancistrocladus ileboensis* with potent activities against *Plasmodium falciparum* and against multiple myeloma and leukemia cell lines. *J Nat Prod.* 2017;80:443–58.
67. Tshitenge DT, Bruhn T, Feineis D, Schmidt D, Mudogo V, Kaiser M, et al. Ealamines A-H, a series of naphthylisoquinolines with the rare 7,8'-coupling site, from the Congolese liana *Ancistrocladus ealaensis*, targeting pancreatic cancer cells. *J Nat Prod.* 2019;82:3150–64.
68. Gbedema SY, Bayor MT, Annan K, Wright CW. Clorodane diterpenes from *Polyalthia longifolia* (Sonn) Thw. var. *pendula*: Potential antimalarial agents for drug resistant *Plasmodium falciparum* infection. *J Ethnopharmacol.* 2015;169:176–82.
69. Kouam SF, Ngouonpe AW, Bullach A, Lamshöft M, Kuigoua GM, Spittler M. Monoterpenes with antibacterial activities from a Cameroonian medicinal plant *Canthium multiflorum* (Rubiaceae). *Fitoterapia.* 2013;91:199–204.
70. Betti JL. Medicinal plants sold in Yaoundé markets, Cameroon. *Afr Study Monogr.* 2002;23:47–64.
71. Bouquet A, Debray M. Plantes médicinales de la Côte d'Ivoire, vol. 32. Paris: Mémoires Office de la Recherche Scientifique et Technique d'Outre-Mer (O.R.S.T.O.M.); 1974. p. 232.
72. Wafo P, Nyasse B, Fontaine C, Sondengam BL. Aporphine alkaloids from *Enantia chlorantha*. *Fitoterapia.* 1999;70:157–60.
73. Lacroix D, Prado S, Kamoga D, Kasenene J, Namukobe J, Krief S, et al. Antiparasitoid and cytotoxic activities of medicinal plants traditionally used in the village of Kiohima, Uganda. *J Ethnopharmacol.* 2011;133:850–5.
74. Neuwinger HD. African ethnobotany: poisons and drugs: chemistry, pharmacology, toxicology. Boca Raton: CRC Press; 1996.
75. Boyom FF, Kemgne EM, Tepongning R, Ngouana V, Mbacham WF, Tsamo E, et al. Antiparasitoid activity of extracts from seven medicinal plants used in malaria treatment in Cameroon. *J Ethnopharmacol.* 2009;123:483–8.
76. Kerubo LO, Midiwo JO, Derese S, Langat MK, Akala HM, Waters NC, et al. Antiparasitoid activity of compounds from the surface exudates of *Senecio roseiflorus*. *Nat Prod Commun.* 2013;8:175–6.
77. Muiva-Mutisya L, Macharia B, Heydenreich M, Koch A, Akala HM, Derese S, et al. 6a-Hydroxy- α -toxicarol and (+)-tephrodin with antiparasitoid activities from *Tephrosia* species. *Phytochem Lett.* 2014;10:179–83.
78. Azebaze AGB, Teinkela JEM, Nguemfo EL, Valentin A, Dongmo AB, Vardamides JC. Antiparasitoid activity of some phenolic compounds from Cameroonians *Allanblackia*. *Afr Health Sci.* 2015;15:835–40.
79. Ngemenya MN, Metuge HM, Mbah JA, Zofou D, Babiaka SB, Titanji VPK. Isolation of natural product hits from *Peperomia* species with synergistic activity against resistant *Plasmodium falciparum* strains. *Eur J Med Plants.* 2015;5:77–87.
80. Irungu BN, Orwa JA, Gruhonjic A, Fitzpatrick PA, Landberg G, Kimani F, et al. Constituents of the roots and leaves of *Ekebergia capensis* and their potential antiparasitoid and cytotoxic activities. *Molecules.* 2014;19:14235–46.
81. Derese S, Barasa L, Akala HM, Yusuf AO, Kamau E, Heydenreich M, et al. 4'-Prenyloxyderrone from the stem bark of *Millettia oblata* ssp. *teitensis* and the antiparasitoid activities of isoflavones from some *Millettia* species. *Phytochem Lett.* 2014;31:69–72.
82. Marco M, Deyou T, Gruhonjic A, Holleran JP, Duffy S, Heydenreich M, et al. Pterocarpans and isoflavones from the root bark of *Millettia micans* and of *Millettia dura*. *Phytochem Lett.* 2017;21:216–20.
83. Mbwambo ZH, Moshi MJ, Masimba MJ, Kapingu MC, Nondo RSO. Antimicrobial activity and brine shrimp toxicity of extracts of *Terminalia brownii* roots and stem. *BMC Complement Altern Med.* 2007;7:9.
84. Machumi F, Midiwo JO, Jacob MR, Khan SI, Tekwani BL, Zhang J, et al. Phytochemical, antimicrobial and antiparasitoid investigations of *Terminalia brownii*. *Nat Prod Commun.* 2013;8:761–4.
85. Induli M, Geburu M, Abdissa N, Akala H, Wekesa I, Byamukama R, et al. Antiparasitoid quinones from the rhizomes of *Kniphofia foliosa*. *Nat Prod Commun.* 2013;8:1261–4.
86. Abdissa N, Induli M, Akala HM, Heydenreich M, Midiwo JO, Ndakala A, et al. Knipholone cyclooxanthrone and an anthraquinone dimer with

- antiplasmodial activities from the roots of *Kniphofia foliosa*. *Phytochem Lett.* 2013;6:241–5.
87. Bringmann G, Mutanyatta-Comar J, Maksimenka K, Wanjohi JM, Heydenreich M, Brun R, et al. Joziknipholones A and B: the first dimeric phenylanthraquinones, from the roots of *Bulbine frutescens*. *Chemistry.* 2008;14:1420–9.
 88. Abdissa D, Geleta G, Bacha K, Abdissa N. Phytochemical investigation of *Aloe pulcherrima* roots and evaluation for its antibacterial and antiplasmodial activities. *PLoS ONE.* 2017;12:e0173882.
 89. Feilcke R, Arnouk G, Raphane B, Richard K, Tietjen I, Andrae-Marobela K, et al. Biological activity and stability analyses of knipholone anthrone, a phenyl anthraquinone derivative isolated from *Kniphofia foliosa* Hochst. *J Pharm Biomed Anal.* 2019;174:277–85.
 90. Lenta BN, Ngamgwe RF, Kamdem LM, Ngatchou J, Tantangmo F, Antheaume C, et al. Compounds from *Diospyros canaliculata* (Ebenaceae) and their antiparasitic activities. *Int Res J Pure Appl Chem.* 2015;6:56.
 91. Geremedhin G, Bisrat D, Asres K. Isolation, characterization and in vivo antimalarial evaluation of anthrones from the leaf latex of *Aloe percrassa* Todaro. *J Nat Remedies.* 2014;14:1–7.
 92. Endale M, Ekberg A, Akala HM, Alao JP, Sunnerhagen P, Yenesew A, et al. Bussei hydroquinones A–D from the roots of *Pentas bussei*. *J Nat Prod.* 2012;75:1299–304.
 93. Namukobe J, Kiremire BT, Byamukama R, Kasenene JM, Akala HM, Kamau E, et al. Antiplasmodial compounds from the stem bark of *Neoboutonia macrocalyx* Pax. *J Ethnopharmacol.* 2015;162:317–22.
 94. Douanla PD, Tabopda TK, Tchinda AT, Ciekiewicz E, Frédéric M, Boyom FF, et al. Antrocarines A–F, antiplasmodial ergostane steroids from the stem bark of *Antrocaryon klaineianum*. *Phytochemistry.* 2015;117:521–6.
 95. Irungu BN, Adipo N, Orwa JA, Kimani F, Heydenreich M, Midiwo JO, et al. Antiplasmodial and cytotoxic activities of the constituents of *Turraea robusta* and *Turraea nilotica*. *J Ethnopharmacol.* 2015;174:419–25.
 96. Namukobe J, Kiremire BT, Byamukama R, Kasenene JM, Dumontet V, Guérite F, et al. Cycloartane triterpenes from the leaves of *Neoboutonia macrocalyx* L. *Phytochemistry.* 2014;102:189–96.
 97. Endale A, Bisrat D, Animut A, Bucar F, Asres K. In vivo antimalarial activity of a labdane diterpenoid from the leaves of *Otostegia integrifolia* Benth. *Phytother Res.* 2013;27:1805–9.
 98. Nondo RS, Moshi MJ, Erasto P, Masimba PJ, Machumi F, Kidukuli AW, et al. Anti-plasmodial activity of Norcaesalpin D and extracts of four medicinal plants used traditionally for treatment of malaria. *BMC Complement Altern Med.* 2017;17:167.
 99. Mba'ning BM, Lenta BN, Nougoué DT, Antheaume C, Fongang YF, Ngouela SA, et al. Antiplasmodial sesquiterpenes from the seeds of *Salacia longipes* var. *camerunensis*. *Phytochemistry.* 2013;96:347–52.
 100. Nyongbela KD, Makolo FL, Hoyer TR, Efang SMN. Isolation and characterization of Sclerinenone C from *Scleria striatinux*. *Nat Prod Commun.* 2016;11:5–6.
 101. Happi GM, Kouam SF, Talontsi FM, Lamshöft M, Zühlke S, Bauer JO, et al. Antiplasmodial and cytotoxic triterpenoids from the bark of the Cameroonian medicinal plant *Entandrophragma congoense*. *J Nat Prod.* 2015;78:604–14.
 102. Wahba AE, El-Sayed AKA, El-Falal AA, Soliman EM. New antimalarial lanostane triterpenes from a new isolate of Egyptian *Ganoderma* species. *Med Chem Res.* 2019;28:2246–51.
 103. Bero J, Hérent MF, Schmeda-Hirschmann G, Frédéric M, Quetin-Leclercq J. In vivo antimalarial activity of *Keetia leucantha* twigs extracts and in vitro antiplasmodial effect of their constituents. *J Ethnopharmacol.* 2013;149:176–83.
 104. Simelane MBC, Shonhai A, Shode FO, Smith P, Singh M, Opoku AR. Antiplasmodial activity of some Zulu medicinal plants and of some triterpenes isolated from them. *Molecules.* 2013;18:12313–23.
 105. Happi GM, Kouam SF, Talontsi FM, Zühlke S, Lamshöft M, Spitteller M. Minor secondary metabolites from the bark of *Entandrophragma congoense* (Meliaceae). *Fitoterapia.* 2015;102:35–40.
 106. Chukwujekwu JC, de Kock CA, Smith PJ, van Heerden FR, van Staden J. Antiplasmodial activity of compounds isolated from *Erythrina caffra*. *S Afr J Bot.* 2016;106:101–3.
 107. Ledoux A, Maraetefau H, Jansen O, Etienne D, Quetin-Leclercq J, Clerc P, et al. Phytochemical profile and biological activity evaluation of *Zanthoxylum heterophyllum* leaves against malaria. *Planta Med Lett.* 2015;2:e10–1.
 108. Fouokeng Y, Feusso HMF, Teinkela JEM, Noundou XS, Wintjens R, Isaacs M, et al. In vitro antimalarial, antitrypanosomal and HIV-1 integrase inhibitory activities of two Cameroonian medicinal plants: *Antrocaryon klaineianum* (Anacardiaceae) and *Diospyros conocarpa* (Ebenaceae). *S Afr J Bot.* 2019;122:510–7.
 109. Bapela MJ. NMR-based metabolomic study of medicinal plants used against malaria and the isolated bioactive alkaloids. Ph.D thesis, University of Pretoria, South Africa. 2016.
 110. Ludere MT, van Ree T, Vleggaar R. Isolation and relative stereochemistry of lip-pialactone, a new antimalarial compound from *Lippia javanica*. *Fitoterapia.* 2013;86:188–92.
 111. Zofou D, Ntie-Kang F, Sippl W, Efang SMN. Bioactive natural products derived from the Central African flora against neglected tropical diseases and HIV. *Nat Prod Rep.* 2013;30:1098–200.
 112. Ntie-Kang F, Lifongo LL, Mbaze LM, Ekwelle N, Owono LCO, Megnassan E, et al. Cameroonian medicinal plants: a bioactivity versus ethnobotanical survey and chemotaxonomic classification. *BMC Complement Altern Med.* 2013;13:147.
 113. Lifongo LL, Simoben CV, Ntie-Kang F, Babiaka SB, Judson PN. A bioactivity versus ethnobotanical survey of medicinal plants from Nigeria, West Africa. *Nat Prod Bioprospect.* 2014;4:1–19.
 114. Ntie-Kang F, Lifongo LL, Simoben CV, Babiaka SB, Sippl W, Mbaze LM. The uniqueness and therapeutic value of natural products from West African medicinal plants, part I: uniqueness and chemotaxonomy. *RSC Adv.* 2014;4:28728–55.
 115. Titanji VPK, Zofou D, Ngemenya MN. The antimalarial potential of medicinal plants used for the treatment of malaria in Cameroonian folk medicine. *Afr J Trad Cam.* 2008;5:302–21.
 116. Tajuddeen N, van Heerden FR. Antiplasmodial natural products: an update. *Malar J.* 2019;18:404.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Computer-aided design of peptidomimetic inhibitors of falcipain-3: QSAR and pharmacophore models

Boris D. Bekono^a, Akori E. Esmel^b, Brice Dali^b, Fidele Ntie-Kang^{c,d,*}, Melalie Keita^b, Luc C.O. Owono^{a,e}, Eugene Megnassan^{b,*}

^a Department of Physics, Ecole Normale Supérieure, University of Yaoundé I, P.O. Box 47, Yaoundé, Cameroon

^b Laboratory of Fundamental and Applied Physics, University of Abobo-Adjame (now Nangui Abrogoua), Abidjan 02 BP 801, Cote d'Ivoire

^c Department of Chemistry, Faculty of Science, University of Buea, P.O. Box 63, Buea, Cameroon

^d Department of Informatics and Chemistry, University of Chemistry and Technology Prague, Technická 5 166 28 Prague 6, Dejvice, Czech Republic

^e CEPAMOQ, Faculty of Science, University of Douala, P.O. Box 8580, Douala, Cameroon

* Corresponding author: ntiékfidele@gmail.com (FNK); megnase@yahoo.com (EM)

Received: date; Accepted: date; Published: date

Abstract: In this work antiparasitic peptidomimetics inhibitors (PEP) of falcipain-3 (FP3) of *Plasmodium falciparum* (Pf) have been proposed using structure-based and computer-aided molecular design. Beginning with the crystal structure of PfFP3-K11017 complex (PDB ID: 3BWK), three-dimensional (3D) models of FP3-PEPx complexes with known activities (IC_{50}^{exp}) were prepared by *in situ* modification, based on molecular mechanics and implicit solvation to compute Gibbs free energies (GFE) of inhibitor-FP3 complex formation. This resulted in a quantitative structure-activity relationships (QSAR) model based on a linear correlation between computed GFE ($\Delta\Delta G_{comp}$) and the experimentally measured IC_{50}^{exp} : ($pIC_{50} = -\log_{10} IC_{50}^{exp}/10^9 = -0.8037 \times \Delta\Delta G_{comp} + 4.0009$; $R^2 = 0.95$). Apart from the structure-based relationship, a ligand-based quantitative pharmacophore model (PH4) of novel PEP analogs where substitutions were directed by comparative analysis of the active site interactions was derived using the bound conformation of the PEPx. This provided structural information useful for the design of virtual combinatorial libraries (VL), which was focused by means of computationally predicted absorption, distribution, metabolism, excretion and toxicity (ADMET). The end results were predictive inhibitory activities falling within the low nanomolar concentration range.

Keywords: drug design; falcipain; malaria; molecular dynamics; peptidomimetics; *Plasmodium falciparum*; virtual screening.

1. Introduction

Malaria is a widespread disease, with causative agent *Plasmodium falciparum* (Pf), transmitted mainly by female Anopheles mosquito bites. The disease has been declared a public health concern by the World Health Organization (WHO) in many developing countries (Dye 2006; WHO 2015). Besides, since the implementation of artemisinin-combined therapy (ACT) in 2006, resistance cases have been recorded (Rosenthal et al. 2002; White

et al. 2005; Yeh and Altman 2006; Zhang et al. 2004). Meanwhile, the treatment of malaria mainly depends on ACT, despite resistance to this combination. This suggests the need for industry-academia partnerships for the search of new antimalarials which act *via* alternative modes of action. Two strategic approaches have been suggested in the search for new remedies against malarial; one focused on eliminating the parasite or preventing its contact with potential human hosts and a second aimed at developing efficacious drugs to treat infected patients. The latter is often aimed at the inhibition of a therapeutic target, often a vital enzyme involved in the parasite's lifecycle. This often requires the search for or the design of new molecules capable of binding in a specific manner to known parasite vital enzymes.

During the last two decades, the identification of drug targets against *Pf* has increased tremendously (Rosenthal et al. 2002; White et al. 2005; Yeh and Altman 2006; Zhang et al. 2004), thus favouring the second approach. This is known as "rational drug design and discovery". As an example, the parasite breaks down a large amount of hemoglobin (Hb) from human red blood cells in order to obtain the required nutrients for its growth during the blood stage (Elliott et al. 2008). This involves several proteases, known as validated drug targets in *Pf* (Francis et al. 1997; Miller and Su 2011; Nethavhani and van Ree 2014; Liu et al. 2004). These drug targets could be divided into two major groups:

- i) those which are directly involved in the invasion and rupture of the red blood cells, and
- ii) those dedicated to the breakdown of Hb (Blackman 2000).

Two protease families are involved in Hb breakdown by hydrolysis. These include aspartic proteases (plasmepsins) and cysteine proteases (falcipains, FPs) (Blackman 2000). One metalloprotease called falcilysin (Eggleston et al. 1999), and one dipeptidyl aminopeptidase (Klemba et al. 2004; Bekono et al. 2018) are also involved.

Previous studies have focused on the search for inhibitors of falcipains 2 and 3 (FP-2 and FP-3), respectively (Coterón et al. 2010), even though FP-3, shown to be expressed later in the parasite life cycle, appeared to be a more efficient haemoglobinase than FP-2 (Sijwali et al. 2006). This indicates that FP-3 inhibition is lethal to the parasite and, therefore, constitutes an attractive target in *Pf* drug discovery. Several FP-3 inhibitors have been identified and described in the literature, which are capable of blocking the enzyme's activity by forming reversible or irreversible covalent bonds within the enzyme active site (Klemba et al. 2004; Bekono et al. 2018). These inhibitors could be sub-classified into three categories: peptide-based, non-peptidic, and peptidomimetic inhibitors (Ang et al. 2011; Ramjee et al. 2006), although preference has been given to those known to be reversible and, hence, considered to be potentially more effective than irreversible ones (Chibale and Musonda 2003; Desai et al. 2006). The most promising inhibitors so far are those discovered by chemical synthesis (Olson 1999; Rosenthal et al. 1993; Rosenthal et al. 1996; Rosenthal et al. 1991; Verissimo et al. 2008; Weldon et al. 2014), by molecular docking (Potshangbam et al. 2011) and virtual screening studies (Desai et al. 2006; Desai et al. 2004; Ghasemi et al. 2015; Shah et al. 2010; Shah et al. 2011), particularly from the compound class of peptidomimetics.

Weldon et al. recently designed, synthesised and evaluated a series of peptidomimetic pseudo-prolyl-homophenylalanyl ketones for their inhibition of the *Pf* cysteine proteases FP-2 and FP-3 (Weldon et al. 2014). One of these compounds showed nanomolar range activities against both enzymes (i.e. 80 nM against FP-2 and 60 nM against FP-3 (Weldon et al. 2014). These interesting results have been compounded by the presence of the crystal structures of the FP-3 apostructure co-crystallised with the inhibitor within the protein data bank (Berman

et al. 2000; Kerr et al. 2009a; Kerr et al. 2009b). These have constituted the foundation of this work, which involves the design of PEP2 peptidomimetic analogues with the goal of identifying even more potent candidates *via* quantitative structure-activity relationship (QSAR) with FP3 inhibition pharmacophore. This is intended to further orientate the design of more potent non-peptidic FP3 inhibitors.

Table 1 Training and validation sets of PEP inhibitors obtained from the literature (Weldon et al. 2014).

| Training Set ^[a] | M _w ^[b] (g. mol ⁻¹) | IC ₅₀ ^{exp} ^[c] (nM) |
|-------------------------------|--|--|
| PEP1 (ref) | 482 | 36360 |
| PEP2 | 440 | 60 |
| PEP3 | 574 | 520 |
| PEP4 | 498 | 3560 |
| PEP5 | 450 | 20180 |
| PEP6 | 438 | 23900 |
| PEP7 | 466 | 47230 |
| PEP8 | 488 | 11910 |
| Validation Set ^[a] | M _w ^[b] (g.mol ⁻¹) | IC ₅₀ ^{exp} ^[c,d] (nM) |
| VEP1 | 452 | 540 |
| VEP2 | 470 | 8220 |

[a] See Fig. 1 for chemical structures of training and validation.

[b] Molar weight of PEP analogues.

[c] This is IC₅₀^{exp} expressed in nanomolar concentration (Weldon et al. 2014).

In the present work, a Hansch-type 'complexation' FP-3 inhibition QSAR models based on *in vitro* activities of eight (8) selected PEP derivatives against FP-3 we have built and validated, beginning from the experimental (X-ray crystal) structure of the protein-ligand complex of the enzyme and the potent inhibitor K11017 (PDB ID: 3BWK). This consisted in computing the Gibbs free energies for the formation of the ligand-receptor complexes ($\Delta\Delta G_{\text{comp}}$) based on Molecular Mechanics Poisson-Boltzmann (MM-PB) approach for the training set molecules, followed by the correlation with the experimentally tested biological activities (pIC_{50}). The established QSAR equation was then used to predict the activities of newly designed analogues based on the initial compound scaffold. Additionally, a FP-3 inhibition pharmacophore model (PH4) from the bound conformation of the training set of PEPs was used to screen the virtual library of proposed PEP analogs to identify best

candidates, which have predicted ADMET profiles within the acceptable range for 95% of known drugs.

2. Results and discussion

2.1. Selection of training and validation (or test) data sets

A data set of ten (10) FP-3 inhibitors with a broad range of *in vitro* activities (IC_{50}^{exp}), obtained from the same laboratory, with a sufficiently broad range of activities (60 – 47230 nM) (Weldon et al. 2014) were used to generate a 3D-QSAR model. This was divided into a training set of eight (08) analogues (PEP₁₋₈) used to build the QSAR model and a validation (or test) set of two (02) analogues for evaluating the model (Table 1 and Fig. 1).

2.2. Obtained QSAR model

The relative Gibbs free energy of the non-covalent enzyme-inhibitor (E:I) complex formation from free enzyme (E) and free inhibitor (I), shown in the Experimental Section, were computed for each FP3-PEP_x prepared complex. This was done by modifying *in situ* of the inhibitor K11017 within the binding site of FP-3 of the refined crystal structure, with PDB ID: 3BWK (Kerr et al. 2009a; Kerr et al. 2009b). Table 2 provides the computed values of complex formation GFE ($\Delta\Delta G_{comp}$) and its components (see Experimental Section). Since the $\Delta\Delta G_{comp}$ values were computed in an approximate way, the relevance of the binding model is evaluated by correlating it with the experimental activity data (IC_{50}^{exp}) using linear regression, equation (1).

For this training set, a plot of the linear correlation is shown in Fig. 2 and the statistical data of the regression are provided in Table 3. For the correlation involving $\Delta\Delta G_{comp}$, relatively high regression coefficient on the values, together with the statistical significance Fischer F-test, suggest that there is no chance correlation between the binding mode and the observed inhibitory potencies of the training set.

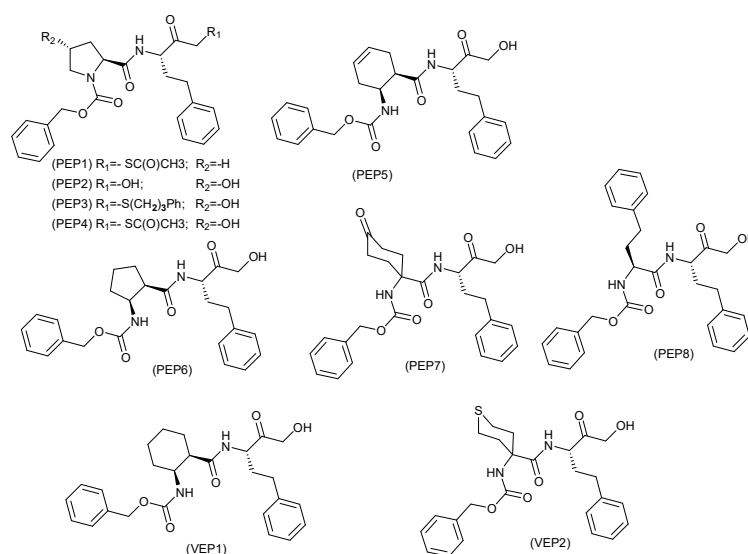


Figure 1. Chemical structures of training (PEP_x) and validation (VEP_x) sets of inhibitors obtained from literature (Weldon et al. 2014)

Table 2 Energy contributions towards $\Delta\Delta G_{\text{comp}}$ for a dataset of PEP analogs against FP-3 complexation Gibbs free energy and its components for the training set of FP3 inhibitors PEP1-8 and validation set VEP1-2.

| Training Set [a] | $\Delta\Delta H_{\text{MM}}^{[b]}$ (kcal. mol ⁻¹) | $\Delta\Delta G_{\text{sol}}^{[c]}$ (kcal. mol ⁻¹) | $\Delta\Delta TS_{\text{vib}}^{[d]}$ (kcal. mol ⁻¹) | $\Delta\Delta G_{\text{comp}}^{[e]}$ (kcal. mol ⁻¹) | $\text{pIC}_{50}^{\text{exp}[f]}$ |
|--------------------|--|---|--|--|-----------------------------------|
| PEP1 | 0.00 | 0.00 | 0.00 | 0.00 | 4.44 |
| PEP2 | -1.92 | -0.08 | 1.78 | -3.78 | 7.22 |
| PEP3 | -1.16 | -1.54 | 0.07 | -2.77 | 6.28 |
| PEP4 | -1.26 | -0.07 | 0.74 | -2.07 | 5.45 |
| PEP5 | -1.01 | -0.22 | -0.13 | -1.10 | 4.70 |
| PEP6 | 0.94 | -2.75 | -0.81 | -1.00 | 4.62 |
| PEP7 | -1.45 | -0.51 | -1.33 | -0.63 | 4.33 |
| PEP8 | -0.56 | 3.25 | 3.72 | -1.03 | 4.92 |
| Validation Set [a] | $\Delta\Delta H_{\text{MM}}^{[b]}$ (kcal. mol ⁻¹) | $\Delta\Delta G_{\text{sol}}^{[c]}$ (kcal. mol ⁻¹) | $\Delta\Delta TS_{\text{vib}}^{[d]}$ (kcal. mol ⁻¹) | $\Delta\Delta G_{\text{comp}}^{[e]}$ (kcal. mol ⁻¹) | Ratio ^[g] |
| VEP1 | -2.10 | -0.23 | 0.18 | -2.51 | 0.96 |
| VEP2 | -0.28 | 0.61 | 1.61 | -1.28 | 0.98 |

[a] For the chemical structures of the training/validation set of inhibitors see Fig. 1.

[b] $\Delta\Delta H_{\text{MM}}$ represents the relative enthalpic contribution to the Gibbs free energy change related to the intermolecular interactions in the enzyme–inhibitor complex derived by molecular mechanics (I_{ref} is the reference inhibitor PEP1):

$$\Delta\Delta H_{\text{MM}} = [E_{\text{MM}}\{E: I_x\} - E_{\text{MM}}\{I_x\}] - [E_{\text{MM}}\{E: I_{\text{ref}}\} - E_{\text{MM}}\{I_{\text{ref}}\}],$$

[c] $\Delta\Delta G_{\text{sol}}$ represents the relative solvation GFE contribution to the GFE of EI complex formation:

$$\Delta\Delta G_{\text{sol}} = [G_{\text{sol}}\{E: I_x\} - G_{\text{sol}}\{I_x\}] - [G_{\text{sol}}\{E: I_{\text{ref}}\} - G_{\text{sol}}\{I_{\text{ref}}\}].$$

[d] $-\Delta\Delta TS_{\text{vib}}$ represents the relative entropic contribution of the inhibitor to the GFE related to the EI complex:

$$-\Delta\Delta TS_{\text{vib}} = [\Delta\Delta TS_{\text{vib}}\{I_x\}_E - \Delta\Delta TS_{\text{vib}}\{I_x\}] - [\Delta\Delta TS_{\text{vib}}\{I_{\text{ref}}\}_E - \Delta\Delta TS_{\text{vib}}\{I_{\text{ref}}\}]$$

[e] $\Delta\Delta G_{\text{comp}}$ represents the relative GFE change related to the enzyme-inhibitor complex formation:

$$\Delta\Delta G_{\text{comp}} = \Delta\Delta H_{\text{MM}} - \Delta\Delta TS_{\text{vib}} + \Delta\Delta G_{\text{sol}}.$$

[f] $\text{IC}_{50}^{\text{exp}}$ (Weldon et al. 2014) represents the inhibitor concentration that causes 50% decrease in the rate of substrate conversion by FP-3 measured in the enzyme assay:

$$\text{IC}_{50}^{\text{exp}} = K_i + [S] \cdot (K_i/K_M) + [\text{FP} - 3]/2$$

where $[S]$ and $[\text{FP} - 3]$ are the substrate and active enzyme concentrations respectively. K_M is the Michaelis

constant determined in FP-3 inhibition assay (Copeland et al. 1995); $\text{pIC}_{50}^{\text{exp}} = -\log_{10}\left(\frac{\text{IC}_{50}^{\text{exp}}}{10^9}\right)$

[g] This is the ratio of the predicted activity on the experimental activity, $\text{pIC}_{50}^{\text{pre}}/\text{pIC}_{50}^{\text{exp}}$. This ratio is close to 1, indicating the predictivity of the QSAR model.

The ratio of the predicted and observed inhibition constants ($pIC_{50}^{pred}/pIC_{50}^{exp}$) for the validation set of two PEPs (not included in the training set) were ~ 1 . This proves the predictive power of the QSAR model, suggesting that the regression equation (1) (Table 3), and the computed $\Delta\Delta G_{comp}$ quantities of the newly designed PEP analogs can be used to predict their inhibitory potencies (IC_{50}^{pred}) against FP-3, on condition that the binding modes of the designed analogues and those of the training set compounds are the same relative to the receptor site. Such an approach could reduce the required number of molecules to be synthesized in a rational drug development project quite considerably. The above procedure has been previously applied by our group in several drug design projects (Dali et al. 2012; Esmel et al. 2017; Frecer et al. 2008; Frecer et al. 2004; Keita et al. 2014; Kily Herve Fagnidi et al. 2018; Kouassi et al. 2015; Megnassan et al. 2012; N'Guessan et al. 2017; N'Guessan and Megnassan 2017; Owono Owono et al. 2013; Owono Owono et al. 2015).

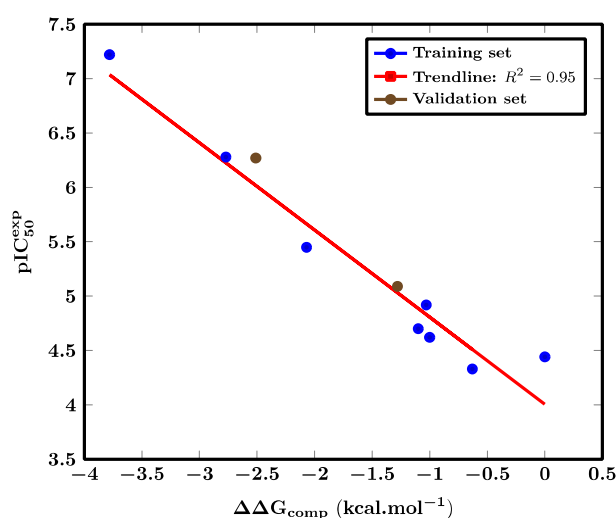


Figure 2. Correlation plot between pIC_{50} and relative complexation Gibbs free energies of the 8 training set of FP-3 inhibitors $\Delta\Delta G_{comp}$

2.3. Inhibitor binding modes

The predicted binding mode of the best active PEP2 coming from the complexation model are illustrated in 3D depiction in Fig. 3. The main interactions with the active site residues namely the H-Bond with the catalytic residue Cys51 are in line with docking study and WaterMap calculations (Weldon et al. 2014) which, unfortunately, did not provide any statistical correlation between binding affinity and activity (results not shown). The bound conformation of PEP shed light on the structural features for binding affinity, which are vital for the design of novel potent non-peptidic FP-3 inhibitors by exploiting the S1' to S2 pockets. In order to get to verify whether other interesting interactions not displayed have to be taken into account in the description of PEP binding mode at FP-3 active site for the rational design of new analogs, the interaction energy (IE) between each active site residue and PEPx was computed. The breakdown of interaction energy diagram into each S1'-S2 subsite residues contribution of FP-3 for PEPs, displayed in Fig. 4, indicates the particular behavior of P1' and the contribution of His183 to IE. This contribution of His183 to IE at the P2 position of PEPs can be correlated with the observed activity of the training set inhibitors as shown in Figure 5.

Moreover, the breakdown of IE diagram into each S1' pocket residues contribution of FP-3 for PEPs, as displayed in Fig. 6 shows that Trp215 is mostly involved in the contribution to IE for the two most active compounds PEP2 and PEP3, and that Asn213 residue is most involved in IE with the most active PEP2 than the others residues in S1' pocket. We have focused on the S1' and S2 subsites because substitutions were made in P1' and P2 during the original SAR described in the literature (Weldon et al. 2014).

It was observed that the IE diagrams analysis could not significantly guide the choice of the R-groups in S1' and S2 subsites, when compared with the case for the design of thymine-like inhibitors of thymidine monophosphate kinase (Keita et al. 2014). It would rather be suggested that a large and diverse combinatorial virtual library (VL) of PEPs be built and screened with our FP-3 inhibition 3D PH4, based on the complexation one descriptor QSAR model. A successful case study was in the design pyrrolidine carboxamide inhibitors of *Mycobacterium tuberculosis* InhA (Kouassi et al. 2015).

Table 3. Statistical data of correlation between computed $\Delta\Delta G_{\text{comp}}$ and experimental activity IC_{50}^{exp} of training set PEPs.

| | |
|--|----------|
| $pIC_{50}^{\text{exp}} = -\log_{10}(IC_{50}^{\text{exp}}/10^9) = -0.8037 \times \Delta\Delta G_{\text{comp}} + 4.0009$ (1) | |
| Statistical data of linear regression: | |
| Number of compounds n | 8 |
| Squared correlation coefficient of regression R^2 | 0.95 |
| Leave-one-out cross-validated squared correlation coefficient R_{cv}^2 | 0.88 |
| Standard error of the regression σ | 0.24 |
| Statistical significance of regression, Fisher F-test | 120.7 |
| Level of statistical significance α | > 95% |
| Range of activities of IC_{50}^{exp} (nM) | 60–47230 |

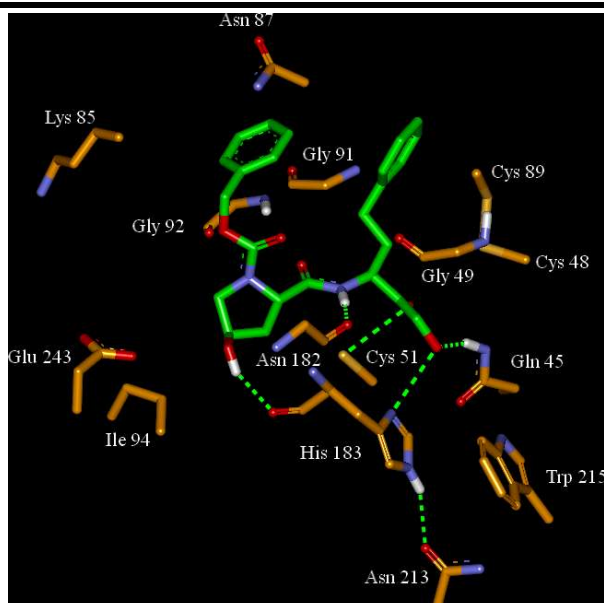


Figure 3. Enzyme-ligand interactions at the active site of FP-3 depicted in 3D for the most active PEP2

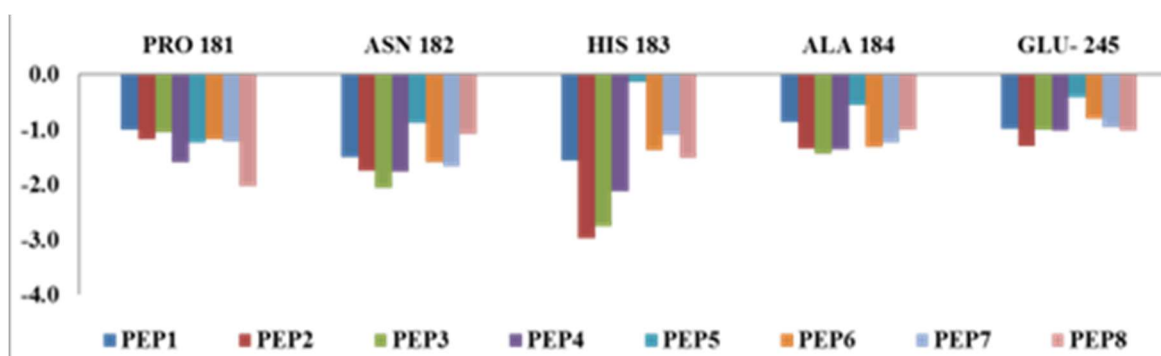


Figure 4. Breakdown of FP3-PEP interaction energy into some active site residue contribution at S1'-S2 pockets

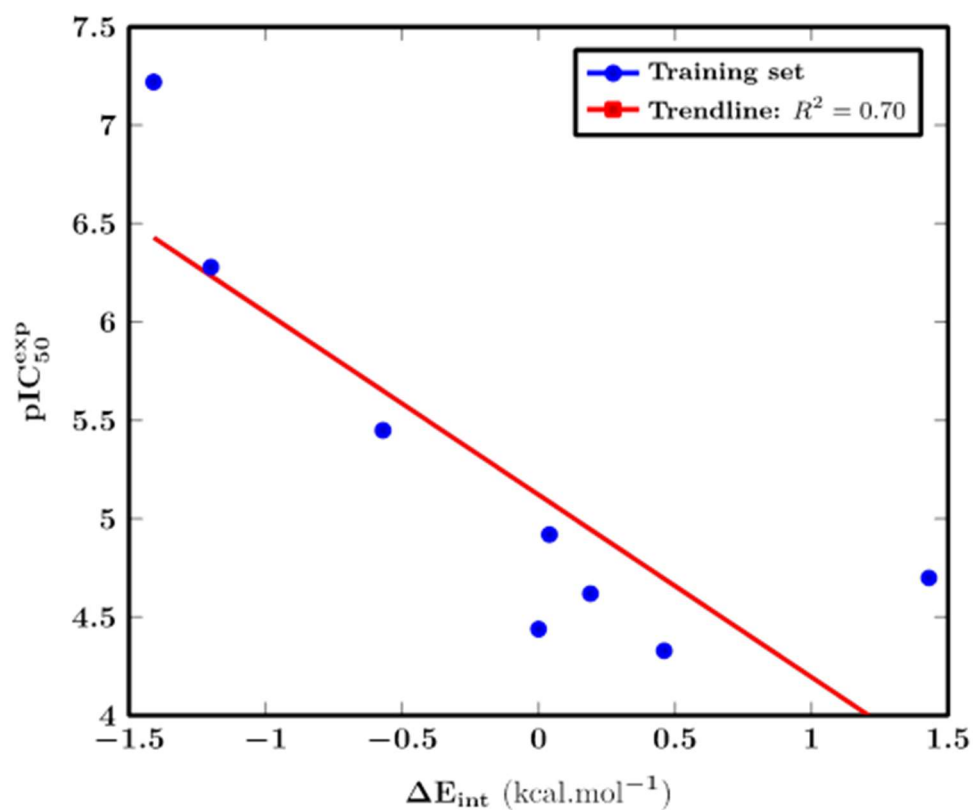


Figure 5. Plot of correlation between pIC_{50}^{exp} and His183 contribution to IE.

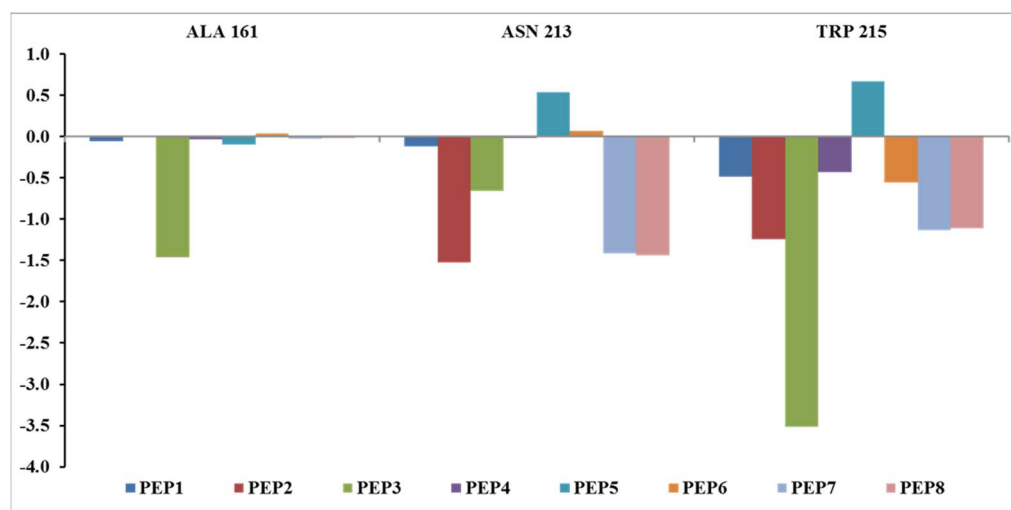


Figure 6: Breakdown of FP3-PEP interaction energy into active site residue contribution at S1' pocket.

2.4. Ligand-based 3D-QSAR PH4 model of FP3 inhibition

The 3D-QSAR PH4 pharmacophore generation process follows three main steps namely: the constructive, the subtractive and the optimisation steps (Accelrys 2009). The constructive phase of Hypo-Gen has automatically selected the most active compounds for which $IC_{50}^{exp} \leq 1.6 \times 60$ nM as leads. Thus, only the most active compound PEP2 ($IC_{50}^{exp} = 60$ nM) was used to generate the starting PH4 features. Only those features were retained which matched this lead. In the subtractive phase, which is normally used to remove pharmacophoric features present in poorly active molecules, none of the training set compounds were found inactive ($IC_{50}^{exp} > 60 \times 10^{3.5} = 189736$ nM). During the optimisation phase, the score of the pharmacophoric hypothesis is improved. Hypotheses are scored according to errors in activity estimates from regression and complexity via a simulated annealing approach. At the end, the top scoring 10 unique pharmacophoric hypotheses (Table 4) were kept, all displaying four features. The generated pharmacophore models were then assessed for their reliability based on the calculated cost parameters. The overall costs ranged from 37.97 (Hypo1) to 38.94 (Hypo10). The relatively small gap between the highest and lowest cost parameter corresponds well with the homogeneity of the generated hypotheses and the consistency of the training set. For this PH4 model, the fixed cost (33.76) is lower than the null cost (107.92) by a difference $\Delta = 74.17$. This difference is a major quality indicator of the PH4 predictability ($\Delta > 70$ correspond to an excellent chance or a probability higher than 90% that the model represents a true correlation (Accelrys 2009)). To be statistically significant the hypotheses have to be as close as possible to the fixed cost and as far as possible from the null cost. The difference $\Delta \geq 70.03$ for the set of 10 hypotheses confirm the high quality of the pharmacophore model.

The standard indicators such as the RMSDs between the hypotheses ranged from 1.024 to 1.098 and the squared correlation coefficient (R^2) falls to an interval from 0.978 to 0.971. The first PH4 hypothesis with the 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 geometry of the Hypo1 pharmacophore of FP-3 inhibition is displayed in Fig. 7. Table 5 list the regression equation (equation 2) for pIC_{50}^{exp} vs

pIC_{50}^{pred} estimated from Hypo1 with related indicators such as R^2 , R_{xv}^2 , Fisher F-test, σ and α , while Figure 8 display a plot of regression equation for pIC_{50}^{exp} vs pIC_{50}^{pre} . To check the consistency of the generated pharmacophore model we have computed the ratio of predicted and observed activities ($pIC_{50}^{pred}/(pIC_{50}^{exp})$) for the validation set. The computed ratios are as follows: VEP1 0.95, VEP2 1.10 all of them relatively close to one, which documents substantial predictive power of the regression for the best PH4 model.

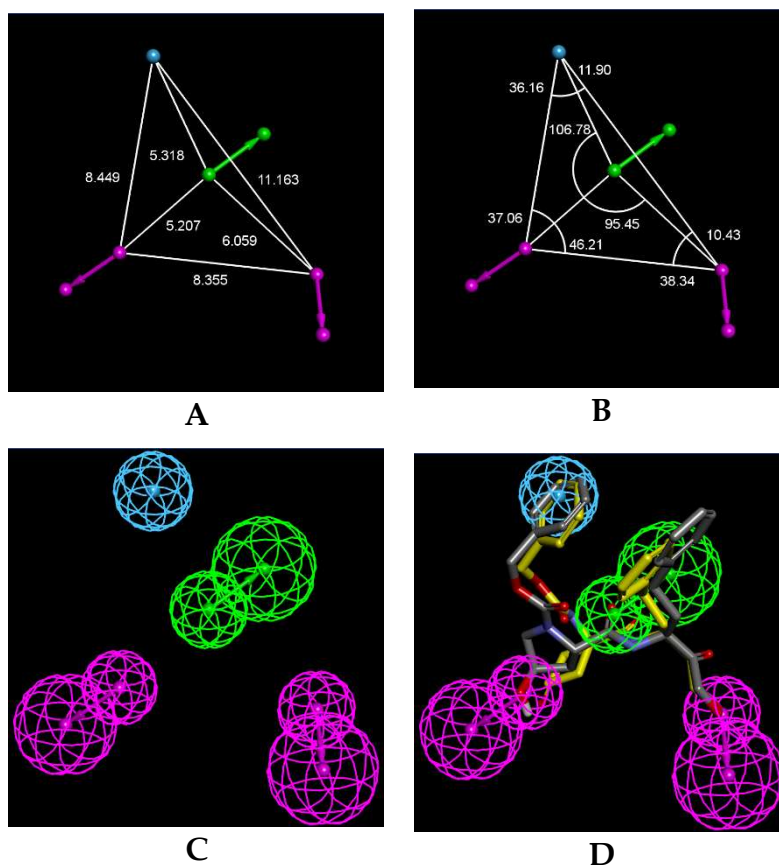


Fig. 7. (A,B) Coordinates of the Hypo1 pharmacophore of FP-3 inhibition, (C) features of the pharmacophore of FP-3 inhibition, (D) pharmacophore mapping with PEP2 (with blue carbons) and PEP7 (with yellow carbons). Color legend of features: hydrophobic (blue), acceptor (green), donor (purple)

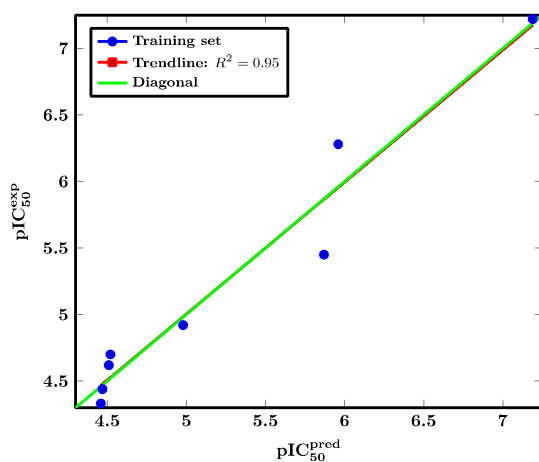


Fig. 8. plot of estimated and experimental activity for PH4

Table 4 Output parameters of the 10 generated PH4 pharmacophoric hypotheses for FP-3 inhibitors after CatScramble validation procedure.

| Hypothes | RMSD ^[a] | R ² ^[b] | Total costs ^[c] |
|-----------|---------------------|-------------------------------|----------------------------|
| Hypo 1 | 1.024 | 0.978 | 37.97 |
| Hypo 2 | 1.050 | 0.977 | 38.15 |
| Hypo 3 | 1.033 | 0.975 | 38.21 |
| Hypo 4 | 0.994 | 0.977 | 38.32 |
| Hypo 5 | 1.053 | 0.974 | 38.54 |
| Hypo 6 | 1.065 | 0.973 | 38.56 |
| Hypo 7 | 1.075 | 0.973 | 38.63 |
| Hypo 8 | 1.013 | 0.976 | 38.67 |
| Hypo 9 | 1.102 | 0.972 | 38.92 |
| Hypo 10 | 1.098 | 0.971 | 38.94 |
| Fixed | 0.000 | 1.000 | 33.76 |
| Null Cost | 4.685 | 0.000 | 107.93 |

^[a] root mean square deviation (RMSD);

^[b] correlation coefficient;

^[c] overall cost parameter of the PH4 pharmacophore

2.5. Library design and ADME focusing

In order to identify more potent orally bioavailable *Pf*FP-3 inhibitors, we have built a virtual library of new analogs inhibitors of *Pf*FP-3 based on substitutions at two positions (P1', and P2) of a scaffold of a lead compound PEP2 in order to better fit S1' and S1 pockets. This virtual library was built to identifying more potent orally bioavailable *Pf*FP-3 inhibitors. The 35 R-groups listed in Table 6 have been attached in positions R₁ to R₅ of the appropriate scaffold to form a combinatorial library of the size:

$$R_1 \times R_2 \times R_3 \times R_4 \times R_5 = 35^5 = 52,521,875 \text{ PEPAs.}$$

It should be noted that one of the important criteria for the design of new anti-malarials, is their oral bioavailability. Thus in the goal to design a more focused library of a reduced size and increased content of drug-like and orally bioavailable molecules, we have introduced a set of filters and penalties, which can help to select smaller number of suitable PEPs which can undergo *in silico* screening. The initial virtual library has been then filtered in an ADME-based focusing step to remove compounds with expected poor oral bioavailability and low drug likeness. Only compounds not violating any rule of Lipinski's (Lipinski et al. 2001) computed for the entire virtual library using

QikProp software (Schrödinger 2014), were kept. This focusing has reduced the size of the initial library to 125,887 PEPA less than 5% its original number size.

2.6. Screening PEPs virtual library using the obtained in silico model

The library of PEP analogs has been further screened for molecular structures matching to the 3D-QSAR PH4 pharmacophore model Hypo1 of FP-3 inhibition. From the set of 125,887 analogs, few thousands of PEPAs mapped to at least 2 features, 592 of which mapped to 4 features of the pharmacophore. Out of them, only 80 best fitting analogs (PH4 hits) have been retained and submitted to screening with help of the complexation QSAR model. Their Gibbs free energy (GFE) upon complex formation with PffFP-3 has been computed along with its component and their predicted half-maximal inhibitory concentration IC_{50}^{pred} has been estimated with the correlation equation (1) (Table 3). The results obtain are given in Table 7. Of the 80 analogs whose inhibitory activities were predicted in Table 7, 30 showed better activities than the most active compound of the training set among them three showed even more activity **PEP-2-1-1-16-2** $IC_{50}^{pred} = 0.004$ nM, **PEP-2-1-12-21-30** $IC_{50}^{pred} = 0.04$ nM and **PEP-2-1-30-7-2** $IC_{50}^{pred} = 0.003$ nM.

Table 5. Statistical data on regression analysis of correlation for the training set between PH4 predicted activity (pIC_{50}^{pred}) and experimental one (pIC_{50}^{exp}) against FP-3.

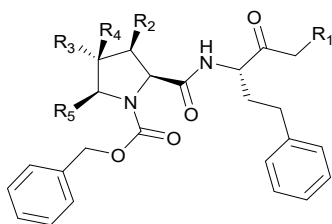
| | |
|---|----------|
| $pIC_{50}^{exp} = -\log_{10}(IC_{50}^{exp} * 10^{-9}) = 0.9958 \times pIC_{50}^{pred} + 0.0219 \quad (2)$ | |
| Statistical data of linear regression: | |
| Number of compounds n | 8 |
| Squared correlation coefficient of regression R^2 | 0.95 |
| Leave-one-out cross-validated squared correlation coefficient R_{cv}^2 | 0.91 |
| Standard error of the regression σ | 0.25 |
| Statistical significance of regression, Fisher F-test | 108.41 |
| Level of statistical significance α | > 95% |
| Range of activities of IC_{50}^{exp} (nM) | 60–47230 |

2.7. Analysis of new inhibitors

In order to identify the substituents that make the analogs predicted to be active, we have analysed the frequency of occurrence of certain substituents chosen from Table 7, on the predicted active analogs. From the 7 best analogs proposed (seen chemical structure in Figure 9) the following R-groups are present 1, 2, 3, 6, 7, 12, 13, 16, 21, 24, 30 and 31. And Figure 10 display the best virtual hit, analog **PEP-2-1-30-7-2** and the least active PEP7 mapped a PH4. Figure 11 display 2D schematic

interaction diagram of the most potent inhibitor PEP2 and the most potent analog design at the active site of *Pf*FP3 as well as Connolly surface of the active site of *Pf*FP-3.

Table 6 R-groups (fragments, building blocks, substituents) used in the design of the initial diversity library of PEP analogs. Dashed bonds indicate the attachment points of the fragments



| | | | |
|----|----------------------------------|----|-----------------------------------|
| 1 | -H | 2 | -CH ₃ |
| 3 | -Cl | 4 | -Br |
| 5 | -I | 6 | |
| 7 | -CH ₂ F | 8 | -CH ₂ Cl |
| 9 | -CH ₂ Br | 10 | -CH ₂ I |
| 11 | -CH ₂ CH ₃ | 12 | -OH |
| 13 | -OCH ₃ | 14 | -OCH ₂ CH ₃ |
| 15 | | 16 | -SH |
| 17 | -SCH ₃ | 18 | -SCH ₂ CH ₃ |
| 19 | | 20 | |
| 21 | | 22 | |
| 23 | | 24 | -N≡ |
| 25 | | 26 | |
| 27 | | 28 | -≡ |
| 29 | | 30 | -NH ₂ |
| 31 | | 32 | |
| 33 | | 34 | -CH=C=CH ₂ |
| 35 | | | |

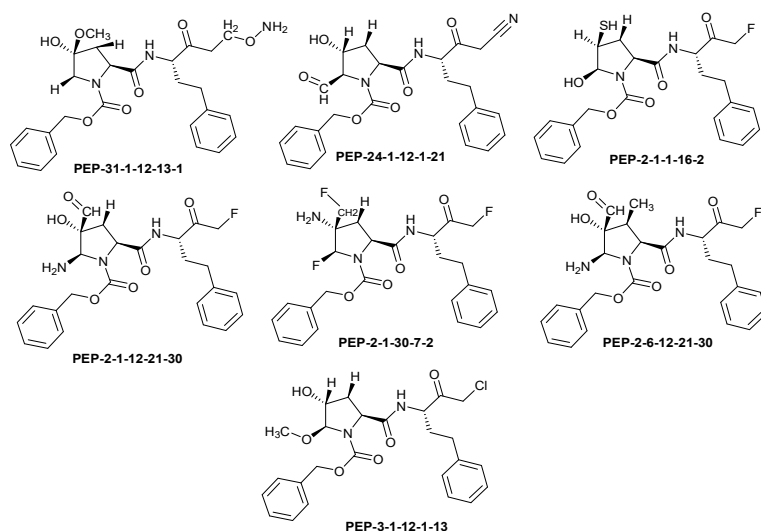


Fig. 9. Chemical structures towards *Pf* FP-3 of seven most potent PEP analogs

Table 7 Complexation Gibbs free energy and its components for the top 80 scoring virtually designed analogs. The analog numbering concatenates the index of each substituent R1 to R5 numbered in Table 6.

| Analog ^[a] | M _w ^[b] | $\Delta\Delta H_{MM}^{[c]}$ | $\Delta\Delta G_{sol}^{[d]}$ | $\Delta\Delta TS_{vib}^{[e]}$ | $\Delta\Delta G_{comp}^{[f]}$ | IC ₅₀ ^{the[g]} |
|-----------------------|----------------------------------|-----------------------------|------------------------------|-------------------------------|-------------------------------|------------------------------------|
| PEP1 | 482 | 0.00 | 0.00 | 0.00 | 0.00 | 36360 |
| 12-1-6-16-30 | 485 | -2.52 | 2.20 | 2.93 | -3.25 | 242 |
| 12-1-1-16-30 | 471 | -2.01 | 0.89 | 2.93 | -4.05 | 56 |
| 30-6-1-16-29 | 495 | 0.58 | 0.07 | 1.30 | -0.66 | 29516 |
| 27-1-1-30-6 | 498 | -4.90 | -0.92 | 0.90 | -6.71 | 0.5 |
| 30-11-1-16-30 | 498 | -0.83 | 0.74 | 0.26 | -0.35 | 52241 |
| 31-1-12-13-1 | 485 | -5.72 | -0.27 | 1.31 | -7.30 | 0.2 |
| 12-1-12-13-1 | 470 | -5.91 | 0.36 | 0.29 | -5.84 | 2 |
| 30-30-6-12-30 | 483 | -0.13 | -0.05 | 0.46 | 0.28 | 168136 |
| 16-21-30-30-1 | 498 | 3.33 | 2.72 | -4.98 | 1.07 | 725269 |
| 16-21-24-1-1 | 493 | 2.24 | 2.50 | 6.56 | -1.82 | 3429 |
| 16-1-21-12-30 | 499 | -5.11 | 1.90 | 1.62 | -4.83 | 13 |
| 16-1-12-1-13 | 486 | -2.48 | 2.20 | 1.90 | -2.19 | 1750 |
| 30-6-12-1-12 | 469 | -2.43 | 1.59 | 3.34 | -4.18 | 44 |
| 16-1-12-1-1 | 456 | -0.29 | -0.31 | 3.82 | -4.42 | 28 |
| 30-1-12-6-12 | 469 | -5.79 | 1.52 | -0.36 | -3.91 | 72 |
| 30-1-12-24-13 | 494 | -3.05 | 1.79 | 1.20 | -2.46 | 1045 |
| 30-1-12-7-29 | 497 | -5.14 | 1.48 | 1.02 | -4.68 | 17 |
| 30-1-3-12-29 | 500 | -4.53 | 1.04 | 1.87 | -5.36 | 5 |
| 30-1-1-16-3 | 490 | -4.14 | 0.14 | 0.85 | -4.85 | 13 |
| 30-1-1-16-29 | 481 | -3.67 | 1.14 | 1.63 | -4.16 | 45 |
| 30-1-1-30-7 | 452 | 0.38 | 0.22 | 1.20 | -0.61 | 32506 |
| 30-1-2-24-29 | 492 | -5.05 | 0.53 | -0.19 | -4.32 | 34 |
| 16-29-30-1-30 | 496 | 2.53 | 0.73 | 3.39 | -0.13 | 77834 |
| 17-21-12-1-1 | 498 | 0.87 | 2.62 | 4.30 | -0.81 | 22384 |
| 16-31-1-6-1 | 499 | 1.52 | 2.23 | 3.36 | 0.39 | 204328 |
| 16-29-12-1-12 | 498 | 1.03 | 1.46 | 5.53 | -3.03 | 364 |
| 16-30-30-1-30 | 485 | 1.30 | 1.18 | 3.63 | -1.15 | 11877 |

| | | | | | | |
|----------------|-----|-------|-------|-------|-------|---------|
| 16-30-30-6-30 | 499 | 0.17 | 0.81 | 3.10 | -2.12 | 1980 |
| 30-1-30-30-35 | 495 | -5.08 | 2.34 | -0.33 | -2.41 | 1149 |
| 30-1-30-6-16 | 484 | -2.11 | 0.25 | -0.49 | -1.38 | 7806 |
| 30-1-30-7-29 | 478 | -3.77 | 1.59 | -0.66 | -1.53 | 5854 |
| 30-1-35-12-30 | 496 | -3.32 | 4.25 | -2.38 | -1.45 | 6854 |
| 30-1-34-1-29 | 475 | -3.14 | -1.31 | -0.77 | -3.69 | 108 |
| 30-2-12-6-29 | 514 | -2.99 | 1.40 | -1.66 | 0.07 | 113594 |
| 30-2-28-1-12 | 497 | 0.44 | 0.85 | 1.63 | -0.34 | 53133 |
| 30-6-12-24-30 | 493 | -1.52 | 0.48 | 2.85 | -3.90 | 73 |
| 24-1-12-1-21 | 477 | -4.04 | 0.55 | 3.42 | -6.91 | 0.3 |
| 24-1-6-6-.21 | 489 | -3.08 | -0.45 | 0.36 | -3.89 | 75 |
| 24-29-30-1-12 | 490 | 0.49 | -0.08 | 1.02 | -0.62 | 31859 |
| 24-21-12-1-12 | 493 | -1.57 | 3.38 | 6.88 | -5.07 | 8 |
| 24-21-34-1-1 | 487 | 2.98 | -0.05 | 4.76 | -1.84 | 3338 |
| 28-21-12-1-2 | 494 | -3.78 | 5.35 | 6.37 | -4.80 | 14 |
| 29-21-6-12-1 | 492 | -6.27 | 4.79 | 3.19 | -4.68 | 17 |
| 28-1-31-30-1 | 492 | 5.00 | 1.19 | 5.05 | 1.15 | 833593 |
| 28-30-30-30-30 | 492 | -0.37 | 1.97 | 4.39 | -2.79 | 568 |
| 28-12-31-1-1 | 493 | 2.62 | 2.75 | 6.84 | -1.47 | 6608 |
| 16-1-1-35-12 | 498 | -2.58 | 3.49 | 5.05 | -4.14 | 47 |
| 16-1-21-12-30 | 499 | -5.23 | 2.54 | 1.01 | -3.69 | 107 |
| 12-30-6-30-29 | 494 | -0.18 | -0.51 | -2.91 | 2.22 | 6070564 |

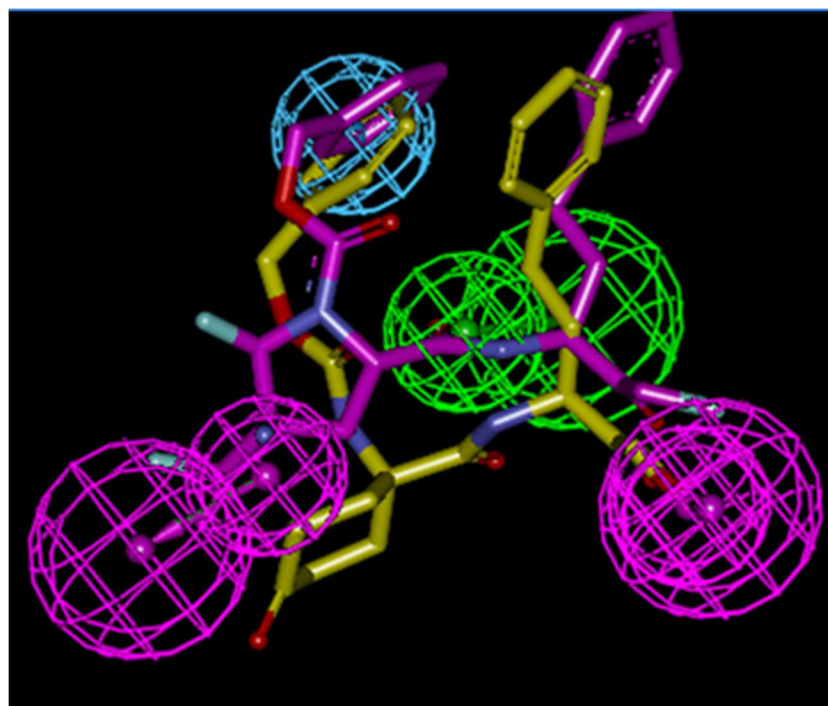


Figure 10. The best virtual hit, analog PEP-2-1-30-7-2 (with purple carbons), the inhibitor PEP7 (Figure 1) shown in yellow carbons, mapped a PH4.

(Table 7 continued)

| Analog ^[a] | M _w _[b] | $\Delta\Delta H_{MM}^{[c]}$ | $\Delta\Delta G_{sol}^{[d]}$ | $\Delta\Delta TS_{vib}^{[e]}$ | $\Delta\Delta G_{comp}^{[f]}$ | IC ₅₀ ^{pred} _[g] |
|-----------------------|----------------------------------|-----------------------------|------------------------------|-------------------------------|-------------------------------|---|
| PEP1 | 482 | 0.00 | 0.00 | 0.00 | 0.00 | 36360 |
| 12-32-1-12-12 | 496 | 0.78 | 1.18 | -0.23 | 2.19 | 5721022 |
| 12-30-12-6-29 | 495 | -1.72 | 3.76 | 0.25 | 1.79 | 2729506 |
| 12-30-6-30-13 | 498 | -2.11 | 0.57 | -2.93 | 1.39 | 1295250 |
| 12-21-1-30-13 | 497 | -0.41 | 1.05 | 1.13 | -0.50 | 39472 |
| 12-21-12-1-13 | 498 | -4.94 | 3.24 | 1.87 | -3.56 | 137 |
| 12-21-12-6-12 | 498 | -1.21 | 1.34 | 1.17 | -1.05 | 14,416 |
| 12-21-24-6-1 | 491 | 2.64 | -0.85 | 3.51 | -1.72 | 4,114 |
| 12-1-1-12-13 | 470 | -7.81 | 2.59 | 0.77 | -5.98 | 2 |
| 12-1-1-35-2 | 484 | -1.26 | 1.65 | 1.02 | -0.63 | 31,045 |
| 12-11-1-16-30 | 499 | 0.89 | 1.65 | 1.94 | 0.60 | 304,897 |
| 12-11-12-1-29 | 494 | -2.66 | 2.76 | -1.03 | 1.13 | 804,379 |
| 12-11-12-1-13 | 498 | -4.80 | 2.71 | -0.78 | -1.31 | 8,907 |
| 12-2-1-12-13 | 488 | -8.52 | 1.68 | -0.65 | -6.19 | 1 |
| 12-6-2-30-29 | 497 | -1.18 | -0.83 | 1.92 | -3.93 | 69 |
| 12-6-1-12-13 | 480 | -3.48 | 1.75 | 2.40 | -4.13 | 48 |
| 16-21-30-1-30 | 498 | 2.03 | 1.40 | 4.78 | -1.35 | 8,209 |
| 16-6-12-1-23 | 499 | -2.96 | 1.02 | 3.71 | -5.65 | 3 |
| 6-21-12-1-30 | 481 | -3.85 | 2.14 | 5.50 | -7.21 | 0.2 |
| 7-21-12-1-30 | 499 | -5.90 | 2.76 | 2.79 | -5.93 | 2 |
| 8-21-30-1-1 | 500 | -0.44 | 1.24 | 3.63 | -2.84 | 523 |
| 11-1-30-30-21 | 494 | -3.30 | 0.08 | 2.44 | -5.66 | 3 |
| 11-21-12-1-2 | 498 | -5.74 | 2.44 | 1.36 | -4.66 | 18 |
| 12-1-16-30-28 | 495 | -3.01 | 2.67 | 0.32 | -0.65 | 29,704 |
| 2-1-1-16-2 | 476 | -3.46 | -0.58 | 5.19 | -9.22 | 0.01 |
| 2-1-12-21-30 | 485 | -6.91 | 1.70 | 2.75 | -7.96 | 0.1 |
| 2-1-30-7-2 | 491 | -6.13 | 0.50 | 3.76 | -9.39 | 0.01 |
| 2-2-12-1-30 | 475 | -4.78 | 1.51 | 4.03 | -7.31 | 0.2 |
| 2-6-12-21-30 | 499 | -3.41 | 0.66 | 3.72 | -6.47 | 1 |
| 3-1-12-1-13 | 488 | -5.63 | 1.72 | 2.33 | -6.24 | 1 |
| 3-24-12-1-30 | 498 | 1.85 | 2.75 | 5.19 | -0.59 | 33,461 |
| 6-1-30-7-2 | 487 | -6.91 | 0.44 | -1.12 | -5.35 | 5 |

^[a] Designed analogues

^[b] M_w represents molecular mass of the inhibitor;

^[c] $\Delta\Delta H_{MM}$ represents the relative enthalpic contribution to the Gibbs free energy change related to the FP3-PEP complex formation $\Delta\Delta G_{comp}$;

^[d] $\Delta\Delta G_{sol}$ represents the relative solvation Gibbs free energy contribution to $\Delta\Delta G_{comp}$;

^[e] $\Delta\Delta TS_{vib}$ represents the relative entropic (vibrational) contribution to $\Delta\Delta G_{comp}$;

^[f] $\Delta\Delta G_{comp}$ represents the relative Gibbs free energy change related to the enzyme-inhibitor FP3-PEP complex formation $\Delta\Delta G_{comp} = \Delta\Delta H_{MM} - \Delta\Delta TS_{vib} + \Delta\Delta G_{sol}$

^[g] IC₅₀^{pred} represents the predicted inhibition constant towards PjFP-3 calculated from $\Delta\Delta G_{comp}$ using correlation equation (1), Table 3.

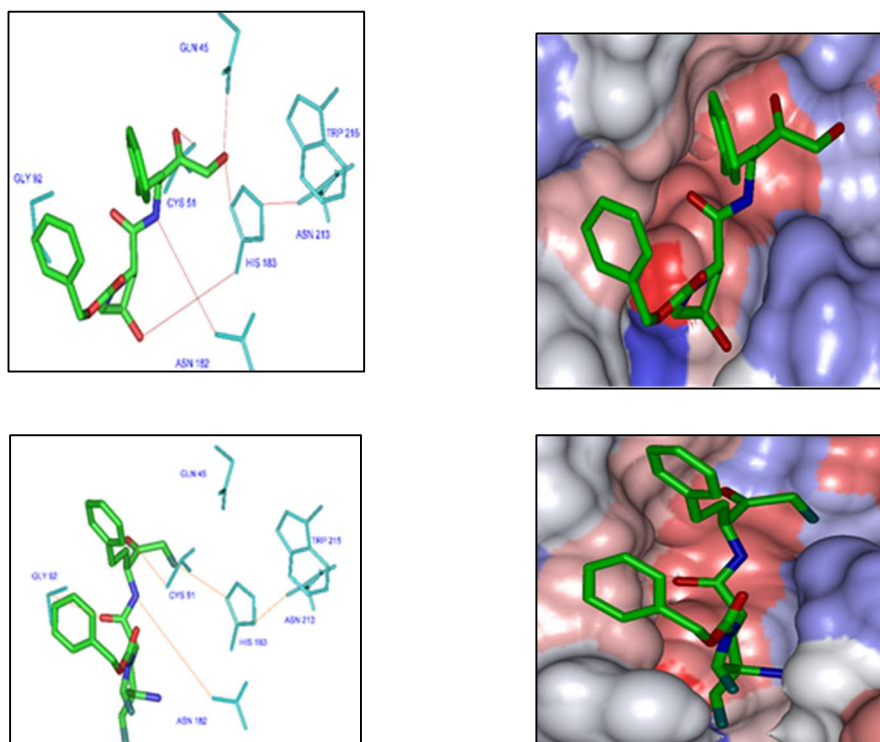


Figure 11. (Top) Left: 2D schematic interaction diagram of the most potent inhibitor PEP2 at the active site of *PjFP-3*. Right. Connolly surface of the active site of *PjFP-3* with the most potent inhibitor PEP2. The binding site surface is colored according to residue hydrophobicity: red - hydrophobic, blue - hydrophilic and white - intermediate. (Bottom) Left: 2D schematic interaction diagrams of the most potent analog design 2-1-30-7-2 at the active site of *PjFP-3*. Right. Connolly surfaces of the active site of *PjFP-3* with the predicted most active PEP inhibitor 2-1-30-7-2. The binding site surface is colored according to residue hydrophobicity: red - hydrophobic, blue - hydrophilic and white - intermediate.

2.8. Predicted ADMET profiles of designed PEPs

The ADMET values for the best active designed PEPs were compared with those computed for drugs used for the treatment of malaria individually or in Artemisinin combined therapy (ACT) or are currently undergoing clinical trials, Table 8. Among the 11 analogs compared, 9 display #stars descriptor equal to zero, meaning that the optimal value range of none of the drug-likeness descriptors has been violated. Additionally, oral bioavailability is among the main requirements of the WHO with regard to new drugs in general also anti-malarial drugs. In the last column of table 8, a high level drug-likeness descriptor is displayed. It is the percentage of human oral absorption in gastrointestinal tract (HOA). All the best provisionally active compounds used for comparison have this descriptor comprise between 58.3% and 100% (< 25% - poor, > 80% high) is the range of 95% of drugs. It would be interesting to note that the #stars = 0 for the designed PEP analogues, which is not the case with most of the ACT antimalarials.

Table 8. Predicted ADME-related properties of the best-designed PEP analogs.

| Analog ^a | # Stars ^b | M _w ^c | Smol ^d | Smol _{hf} ^e | Vmol ^f | RotB ^g | HBdon ^h | HBacc ⁱ | LogP _{o/w} ^j | LogS _{wat} ^k | LogK _{SHA} ^l | LogB/B ^m | BIP _{caco} ⁿ | #meta ^o | IC ₅₀ ^{pred} | HOA ^p | %HOA ^q |
|---------------------|----------------------|-----------------------------|-------------------|---------------------------------|-------------------|-------------------|--------------------|--------------------|----------------------------------|----------------------------------|----------------------------------|---------------------|----------------------------------|--------------------|----------------------------------|------------------|-------------------|
| 27-1-1-30-6 | 0 | 511 | 818 | 632.6 | 1622.7 | 13 | 4 | 9 | 2.71 | -2.8 | 0.22 | -0.7 | 23.4 | 6 | 0.5 | 2 | 59 |
| 31-1-12-13-1 | 0 | 498 | 774.4 | 594.2 | 1543.5 | 13 | 2 | 9 | 3.23 | -4.7 | 0.17 | -1.9 | 160.6 | 3 | 0.2 | 3 | 85 |
| 12-1-12-13-1 | 0 | 483 | 795.5 | 642 | 1538.5 | 12 | 2 | 9 | 3.16 | -4.7 | 0.08 | -1.7 | 245.9 | 3 | 2 | 3 | 88 |
| 24-1-12-1-21 | 1 | 490 | 806.9 | 523.7 | 1561.6 | 12 | 0 | 10 | 2.14 | -5.1 | -0.24 | -2.9 | 21.4 | 5 | 0.3 | 2 | 63 |
| 6-21-12-1-30 | 0 | 494 | 808 | 591.6 | 1571.8 | 12 | 1 | 9 | 2.72 | -3.9 | 0.26 | -1.9 | 17.1 | 6 | 0.2 | 2 | 65 |
| 2-1-1-16-2 | 1 | 489 | 847.4 | 622.5 | 1562.2 | 9 | 1 | 6 | 5.5 | -7.8 | 0.80 | -0.73 | 904 | 4 | 0.01 | 1 | 100 |
| 2-1-12-21-30 | 0 | 498 | 866.1 | 607.9 | 1606 | 11 | 2 | 9 | 2.9 | -5.2 | 0.34 | -2.1 | 15.2 | 5 | 0.1 | 2 | 65 |
| 2-1-30-7-2 | 0 | 504 | 813.6 | 593.7 | 1546.8 | 9 | 2 | 7 | 4.2 | -5.7 | 0.63 | -0.6 | 150.3 | 3 | 0.01 | 3 | 77 |
| 2-2-12-1-30 | 0 | 488 | 762.4 | 525.9 | 1490.5 | 10 | 2 | 8 | 3.05 | -3.9 | 0.34 | -1.17 | 39.07 | 5 | 0.2 | 3 | 73 |
| 2-6-12-21-30 | 0 | 512 | 881.1 | 649.9 | 1649.1 | 11 | 2 | 9 | 3.35 | -5.18 | 0.46 | -1.77 | 24.02 | 5 | 1 | 2 | 58 |
| 3-1-12-1-13 | 0 | 501 | 793.1 | 600.6 | 1537.9 | 10 | 1 | 9.7 | 3.67 | -5.47 | 0.11 | -1.08 | 614.1 | 4 | 1 | 3 | 85 |
| Chloroquine | 1 | 294 | 594.1 | 188.9 | 982.9 | 6 | 0 | 3 | 4.56 | -5.3 | 0.41 | -0.1 | 3718.1 | 0 | - | 3 | 100 |
| Amodiaquine | 1 | 334 | 603.2 | 131.7 | 1018.7 | 6 | 0 | 5 | 3.61 | -4.4 | -0.02 | -0.4 | 1689.1 | 0 | - | 3 | 100 |
| Dapsone | 1 | 236 | 431.6 | 0 | 687.9 | 2 | 0 | 7 | -0.37 | -0.5 | -1.34 | -0.9 | 289.1 | 0 | - | 3 | 69 |
| Trimethoprim | 0 | 272 | 500.2 | 223.9 | 835.9 | 5 | 0 | 6.5 | 0.59 | -1.5 | -0.91 | -1.2 | 282.8 | 3 | - | 3 | 74 |
| Mefloquine | 2 | 362 | 533.1 | 0 | 925.1 | 2 | 0 | 4 | 4.14 | -4.9 | 0.15 | 0.5 | 2903.1 | 0 | - | 3 | 100 |
| pamaquine | 0 | 315 | 654.8 | 443.4 | 1148.1 | 9 | 1 | 4.75 | 4.02 | -3.8 | 0.43 | 0.2 | 1475.2 | 5 | - | 3 | 100 |
| Sulfametopyrazine | 1 | 268 | 473.4 | 77.8 | 773.3 | 4 | 0 | 9 | -1.03 | 0.2 | -1.7 | -1.3 | 195.8 | 1 | - | 2 | 62 |
| Quinacrine | 0 | 370 | 680.5 | 268.8 | 1163.6 | 7 | 0 | 3.5 | 5.57 | -6.5 | 0.8 | -0.1 | 4435.7 | 1 | - | 1 | 100 |
| Tetracycline | 5 | 422 | 604.5 | 173.1 | 1111.8 | 2 | 0 | 7 | -3.43 | 1.1 | -2.5 | -2.6 | 6.8 | 5 | - | 1 | 22 |
| Lumefantrine | 5 | 497 | 819.1 | 160.7 | 1437.5 | 7 | 0 | 3 | 8.27 | - | 1.7 | 0.2 | 4337.2 | 0 | - | 1 | 100 |
| Bulaquine | 0 | 369 | 560.2 | 360.2 | 1097.8 | 9 | 1 | 5.8 | 3.62 | -2.98 | 0.1 | -0.4 | 3099.7 | 7 | - | 3 | 100 |
| Hydroxychloroquine | 1 | 310 | 609.5 | 119.5 | 1006.5 | 6 | 0 | 5 | 3.36 | -4.51 | -0.1 | -0.7 | 1023.7 | 0 | - | 3 | 100 |
| Sulfadoxine | 1 | 296 | 510.6 | 152.3 | 849.5 | 5 | 0 | 9.5 | -0.79 | -0.11 | -1.7 | -1.4 | 213.4 | 2 | - | 2 | 64 |
| Halofantrine | 5 | 470 | 785.4 | 160.2 | 1351.8 | 5 | 0 | 3 | 7.63 | -9.9 | 1.5 | 0.2 | 2844.1 | 0 | - | 1 | 100 |
| Proguanil | 1 | 238 | 478.2 | 125.3 | 768.6 | 6 | 0 | 6 | 1.09 | -1.5 | -1.1 | -0.7 | 834.6 | 0 | - | 3 | 86 |
| Doxycycline | 4 | 422 | 602.2 | 174.1 | 1104.2 | 2 | 0 | 17.2 | -3.99 | 1.73 | -2.88 | -2.45 | 9.17 | 4 | - | 1 | 21 |
| Arteether | 1 | 312 | 531.1 | 506.1 | 970.2 | 2 | 0 | 5.7 | 2.7 | -2.99 | -0.2 | 0.2 | 5731.8 | 0 | - | 3 | 100 |
| dihydroartemisinin | 1 | 284 | 477.4 | 395.7 | 864.6 | 1 | 1 | 5.7 | 1.84 | -2.92 | -0.1 | -0.1 | 1664.9 | 0 | - | 3 | 95 |
| Artemisinin | 0 | 282 | 456.6 | 380.6 | 848.4 | 0 | 0 | 5.3 | 1.7 | -2.1 | -0.30 | 0.001 | 1886 | 1 | - | 3 | 96 |

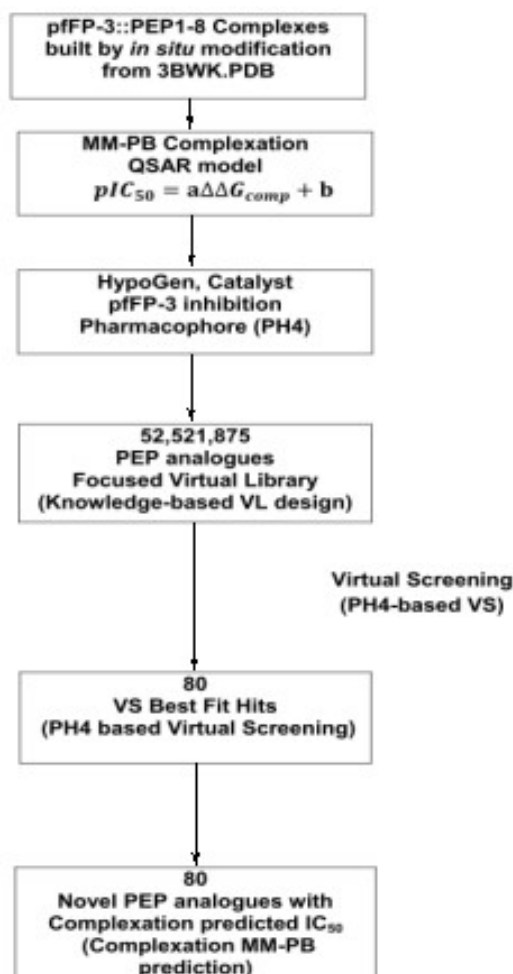
a) best designed APPs analogues, Table 7;

- b) drug likeness, number of property descriptors (from 24 out of the full list of 49 from of QikProp, ver. 3.7 (Schrödinger 2014)) that fall outside of the range of values for 95% of known drugs;
- c) molecular weight in g.mol^{-1} (range for 95% of drugs: 130-725 g.mol^{-1});
- 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 non-trivial (not CX3), non-hindered (not alkene, amide, small ring) rotatable bonds (range for 95% of drugs: 0-15);
- h) estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer (range for 95% of drugs: 0.0-6.0);
- i) estimated number of hydrogen bonds that would be accepted by the solute from water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be non-integer (range for 95% of drugs: 2.0-20.0);
- j) logarithm of partitioning coefficient between *n*-octanol and water phases (range for 95% of drugs: -2 - 6.5);
- k) logarithm of predicted aqueous solubility, $\log S$. S in mol dm^{-3} is the concentration of the solute in a saturated solution that is in equilibrium with the crystalline solid (range for 95% of drugs: -6.0 - 0.5);
- l) logarithm of predicted binding constant to human serum albumin (range for 95% of drugs: -1.5 - 1.5);
- m) logarithm of predicted brain/blood partition coefficient. Note: QikProp predictions are for orally delivered drugs so, for example, dopamine and serotonin are CNS negative because they are too polar to cross the blood-brain barrier (range for 95% of drugs: -3.0 - 1.2);
- n) predicted apparent Caco-2 cell membrane permeability in Boehringer-Ingelheim scale, in $[\text{nm/s}]$ (range for 95% of drugs: < 25 poor, > 500 great);
- o) number of likely metabolic reactions (range for 95% of drugs: 1 - 8);
- p) predicted inhibition constants $\text{IC}_{50}^{\text{pre}}$ (nM). $\text{IC}_{50}^{\text{pre}}$ was predicted from computed $\Delta\Delta G_{\text{comp}}$, using the regression equation shown in Table 5;
- q) human oral absorption (1 - low, 2 - medium, 3 - high);
- r) percentage of human oral absorption in gastrointestinal tract (<25% - poor, >80% high);

(*) Star Indicating that the property descriptor value falls outside the range of values for 95% of known drugs and numbered consecutively with Arabic numerals in parentheses on the right hand side of the page (if referred to explicitly in the text). They should also be separated from the surrounding text by one space.

3. Materials and Methods

Scheme 1 displays the workflow of different steps involved for the computer-aided drug design of the new PEP analogues.



Scheme 1: Novel PEP analogues design methodology workflow.

3.1. Biological activities of compounds included in the training and test sets

The biological activities (IC_{50}^{exp}) of the compounds included in training and validation sets of PEP PfFP-3 inhibitors were found in the literature, covering a range of activities from 60 nM to about 50,000 nM (Weldon et al. 2014). Weldon and coworkers have synthesized 22 molecules, but not all showed detected biological activities to be included in a QSAR study (e.g. activities recorded as >50 μ M would not be included in our study). Two compounds are isomers (one of them was included in the study). The absolute configuration of one compound was not specified. Finally, 10 compounds (almost the threshold for an acceptable QSAR study) have been used in our study.

3.2. Molecular modeling

3D models of the enzyme-inhibitor (E:I) complexes were built starting from the free enzyme (E) and the free inhibitors (I), both derived from a well-refined X-ray crystal structure (PDB ID: 3BWK) of the co-crystallised potent inhibitor K11017 (or Mu-Leu-Hph-VSPH where VSPH: phenyl vinyl sulfone; Hph: homophenylalanyl; Mu: morpholino urea) retrieved from the PDB (Berman et al. 2000). Chain A was employed in all computations and modeling done on the graphical user interface of Insight II (Accelrys 2005), using a previously well described protocol (Dali et al. 2012;

Esmel et al. 2017; Frecer et al. 2008; Frecer et al. 2005; Frecer et al. 2010; Frecer et al. 2004; Keita et al. 2014; Kily Herve Fagnidi et al. 2018; Kouassi et al. 2015; Megnassan et al. 2012; N'Guessan et al. 2017; N'Guessan and Megnassan 2017; Owono Owono et al. 2013; Owono Owono et al. 2015). This implies that the pH values were kept at 7, while all N- and C-terminal residues were kept neutral, while all water molecules originally in the crystal structure were deleted and protonated and ionised amino acid residues were charged. Each inhibitor was built into the crystal reference structure by modifying the original K11017 inhibitor *in situ*. During this process, all rotatable bonds of the replacing residues underwent an exhaustive conformational search by a careful and gradual energy minimisation of each modified inhibitor within the active site residues of FP-3 within 5 \AA of the inhibitor, leading to the identification of low-energy bound conformations of each modified inhibitor. The various low-energy structures of the E:I complexes were then carefully refined by energy minimisation the whole complex.

3.3. Molecular mechanics

The simulation of each inhibitor, FP-3 and E:I complex were carried out by molecular mechanics (MM) as implemented in the consistent force field CFF91 (Accelrys 2002). All MM calculations used a dielectric constant of 4 for representing dielectric shielding effects in the proteins. The optimisation (energy minimisation process) of the free E, free I and E:I complexes were carried out by a gradual relaxation of the structures, beginning by adding H-atoms, then the residue side chain heavy atoms, and ending up with the relaxation of the protein backbone. A large number of steepest descent, followed by conjugate gradient iterative cycles were employed. A convergence criterion for the average gradient was set to of $0.01\text{ kcal mol}^{-1}\text{ \AA}^{-1}$ in each geometry optimisation procedure.

3.4. Conformational search

The conformation of each free inhibitor was obtained from its bound conformation in the E:I complex, which had been previously obtained by the gradual relaxation to the nearest local energy minimum (see section 2.3). The next step was a Monte Carlo search of about 50000 iterations conducted on low-energy conformations of all rotatable bonds, not including those in the rings. Discovery Studio 2.5 (DS 2.5) molecular modeling program (Accelrys 2009) was used to implement this step, which consisted in the generation of 200 unique conformations per inhibitor after randomly varying all torsional angles of the last accepted conformer by $\pm 15^\circ$ at 5000 K, followed by energy minimisation. During this minimisation, the dielectric constant was kept at $\epsilon = 80$. This was in order to approximately take the dielectric screening effect of hydration upon the generated conformers into account. The lowest total energy conformer was then selected and re-minimised at $\epsilon = 4$.

3.5. Solvation Gibbs free energy

The electrostatic component of solvation Gibbs free energy was computed using the DelPhi package in Discovery Studio (Accelrys 2009). This incorporates the effects of ionic strength by solving the nonlinear Poisson–Boltzmann equation (Gilson and Honig 1991; Rocchia et al. 2001) This DelPhi program treats the solvent as a continuous medium of high dielectric constant ($\epsilon_0 = 80$) while the solute is treated as a cavity with low dielectric ($\epsilon_i = 4$). Boundaries are linked to the solute's molecular surface, which enclose the solute's atomic charges. The molecular electrostatic potential and reaction

field around the solute are solved by a finite difference method on a (235 x 235 x 235) cubic lattice grid for the complexes and free enzyme and (65 x 65 x 65) grid for the free inhibitors, implementing the full Coulombic boundary conditions. In both cases, two (02) subsequent focusing steps led to a similar final resolution of about 0.3 Å per grid unit at 70% filling of the grid by the solute. Physiological ionic strength of 0.145 mol dm⁻³, atomic partial charges and radii defined in the CFF91 parameter set (Maple et al. 1994) and a probe sphere radius of 1.4 Å were used. In this implementation, the electrostatic component of the solvation Gibbs free energy was calculated as the reaction field energy (Bötcher 1973; Frecer and Miertuš 1992; Miertuš et al. 1981).

3.6. Calculation of the entropic term

During the simulation, the vibrational entropy change which occurs as the inhibitor binds to the enzyme was computed by normal mode analysis of the inhibitor vibrations, by using a simplified method previously described by Fischer and co-workers (Fischer et al. 2001; Schwarzl et al. 2002). This approach involves the vibrational analyses of the inhibitor bound at the active site of a 'frozen' enzyme, while the low-energy conformers of the free inhibitor were computed for fully minimised structures. This was done using the Discover package (Accelrys 2005) and the formula $T\Delta S_{\text{vib}} = T\Delta S_{\text{vib}}\{\text{I}\}_{\text{PR}} - T\Delta S_{\text{vib}}\{\text{I}\}$. The method has previously provided a good approximation of the vibrational entropy change of the fully flexible system for small and relatively stiff ligands, i.e. including the degrees of freedom of the protein receptor (Fischer et al. 2001; Schwarzl et al. 2002). In this calculation, the $T\Delta S_{\text{vib}}\{\text{I}\}$ term can explain vibrational motions of the free inhibitor and the conformational flexibility of the molecule, i.e. low frequency vibrations, which correspond to collective motions of atoms with larger amplitudes (conformational changes contribute mostly to this term). The relative values of $\Delta\Delta T S_{\text{vib}}$, with respect to the reference inhibitor, were used to compensate partially for the restricted flexibility of the E. In this respect, the entropy term is also recognised as an important factor for drug optimisation, even though an enthalpy contribution to binding affinity is known to be more essential (Freire 2008).

3.7. Binding affinity calculations

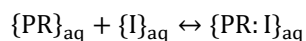
It has been previously proven that the concentration of a competitive tight binding inhibitor that causes 50% reduction of the rate of catalytic substrate conversion (IC_{50}) of a reversible inhibitor depends on the enzyme-inhibition constant K_i as follows:

$$IC_{50}^{\text{exp}} = K_i + [S] \cdot (K_i/K_M) + [E]/2$$

where [S] and [E] are the substrate and enzyme concentrations, respectively, while K_M represents the Michaelis constant (Copeland et al. 1995). The IC_{50} value can thus be predicted from the standard Gibbs free energy change during the enzyme:inhibitor complex formation:

$$\Delta G_{\text{comp}} = -RT \ln K_i,$$

By assuming that there is equilibrium in solution between the solvated protein (or enzyme), $\{\text{PR}\}_{\text{aq}}$, the inhibitor, $\{\text{I}\}_{\text{aq}}$ and the protein-ligand complex, $\{\text{PR}:\text{I}\}_{\text{aq}}$:



Freundlich and co-workers (Freundlich et al. 2009) showed that for molecular simulations of the complex and the free reactants PR and I, the standard Gibbs free energy change of the above equilibrium (1) can be written as:

$$\Delta G_{\text{comp}} = G\{\text{PR: I}\} - G\{\text{PR}\} - G\{\text{I}\}$$

In our calculations, the exact values of standard Gibbs free energies for larger systems such as enzyme: inhibitor complexes were approximated by the derived expressions from the works of Frecer and Miertus (Frecer et al. 2008; Frecer and Miertus 2002; Frecer et al. 1998):

$$G\{\text{E: I}\} \cong [E_{\text{MM}}\{\text{E: I}\} + RT - TS_{\text{trv}}\{\text{E: I}\} + G_{\text{sol}}\{\text{E: I}\}]$$

with E_{MM} representing the molecular mechanics total energy of the complex (including bonding and non-bonding contributions), $G_{\text{sol}}\{\text{E: I}\}$ representing the solvation Gibbs free energy and $TS_{\text{trv}}\{\text{E: I}\}$ representing the entropic term:

$$TS_{\text{trv}}\{\text{E: I}\} = TS_{\text{trans}}\{\text{E: I}\} + TS_{\text{rot}}\{\text{E: I}\} + TS_{\text{vib}}\{\text{E: I}\},$$

The entropic term $TS_{\text{trv}}\{\text{E: I}\}$ here represents a sum of contributions arising from translational, rotational and vibrational motions. When assuming that the translational and rotational terms for the free enzyme and for the enzyme: inhibitor complex were approximately taken to be equal ($TS_{\text{trans}}\{\text{E: I}\} \approx TS_{\text{trans}}\{\text{E}\}$) and ($TS_{\text{trv}}\{\text{E: I}\} \approx TS_{\text{trv}}\{\text{E}\}$), thus leading to:

$$\begin{aligned} \Delta G_{\text{comp}} &\cong [E_{\text{MM}}\{\text{E: I}\} - E_{\text{MM}}\{\text{E}\} - E_{\text{MM}}\{\text{I}\}] + [G_{\text{sol}}\{\text{E: I}\} - G_{\text{sol}}\{\text{E}\} - G_{\text{sol}}\{\text{I}\}] + TS_{\text{trans}}\{\text{I}\} + TS_{\text{rot}}\{\text{I}\} \\ &\quad - [TS_{\text{vib}}\{\text{E: I}\} - TS_{\text{vib}}\{\text{E}\} - TS_{\text{vib}}\{\text{I}\}] \\ &= \Delta H_{\text{MM}} + TS_{\text{trans}}\{\text{I}\} + TS_{\text{rot}}\{\text{I}\} - \Delta TS_{\text{vib}} + \Delta G_{\text{sol}} \end{aligned}$$

with $TS_{\text{trans}}\{\text{I}\}$ and $TS_{\text{rot}}\{\text{I}\}$ describing the translational and rotational entropy terms of the free inhibitor, respectively, and ΔTS_{vib} representing the vibrational entropy change during the formation of the enzyme-inhibitor complex. By comparing between different inhibitors *via* relative changes in their respective complexation Gibbs free energies, with respect to a reference inhibitor, I_{ref} (in this case PEP1), and by assuming the ideal gas behaviour for the rotational and translational motions of the inhibitors, it can be shown that:

$$\Delta \Delta G_{\text{comp}} = \Delta G_{\text{comp}}(\text{I}) - \Delta G_{\text{comp}}(I_{\text{ref}}) = \Delta \Delta H_{\text{MM}} - \Delta \Delta TS_{\text{vib}} + \Delta \Delta G_{\text{sol}}$$

The advantage of such an approach is that the evaluation of relative changes is preferable, since it results in the partial cancellation of errors caused by the approximate nature of the MM method. Additionally, addition solvent and entropic effects are included in the description.

3.8. Interaction energy calculations

Interaction energy values were computed using Discovery Studio 2.5 (Accelrys 2009). The MM interaction energy (E_{int}) protocol available in this program computes the (non-bonded) van der Waals and electrostatic interactions between enzyme residues and each inhibitor. The CFF force field (Accelrys 2009) was used during the calculations, with a dielectric constant set at 4. The breakdown of E_{int} into the contributions by active site residues reveals the significance of individual interactions and permits us to carry out a comparative analysis. The approach leads to the identification of affinity values which would enhance the prediction of favourable and unfavourable PEP substitutions.

3.9. Pharmacophore (PH4) modeling

By definition, a pharmacophore is often regarded as a set of features arranged in 3D space which are essential for a molecule to exert a certain biological activity. The perception of a pharmacophore is

essential for understanding the interaction between a ligand and its receptor. The PH4 concept is based on the assumption that a set of structural features in a molecule is recognised at the receptor site and is responsible for the molecule's biological activity. Bound conformations of inhibitors taken from E:I complexes were used to construct 3D-QSAR pharmacophore models using Catalyst HypoGen algorithm implemented in Discovery Studio 2.5 (Accelrys 2009). This consisted in building a top scoring pharmacophore hypothesis from the most active inhibitor. Three stages (construction, subtraction and optimisation) are involved, meanwhile the inactive ones were used to define the exclusion volume. A maximum number of five features were selected according to the PEP scaffold and substituents, i.e. hydrophobic aliphatic (HYd), hydrophobic aromatic (HYdAr), hydrogen-bond acceptor (HBA) hydrogen-bond donor (HBD) and ring aromatic (Ar). As per the adjustable parameters in the HypoGen protocol, all were kept by default excepting for the uncertainty on the activity and the minimum inter-feature distance, which were set to 1.6 and 1.25 Å, respectively. These parameters were carefully chosen in order to bring the uncertainty interval of experimental activity from a wide span [$IC_{50}/3$, $3 \times IC_{50}$] to a relatively narrow one [$4 \times IC_{50}/5$, $5 \times IC_{50}/4$]. This is important because the accuracy and homogeneity of the measured inhibitory activities based on the fact that the experimental biological activities were derived from the same laboratory must be taken into account (Weldon et al. 2014). During the generation of 10 pharmacophores, 0 was set as the number of missing features and the best pharmacophore models were selected.

3.10. Generation of the virtual library

Molecular models of new analogues were generated using the Molecular Operating Environment (MOE) program (Chemical Computing Group 2014). This was carried out by attaching the R-groups (fragments, building blocks) onto the PEP scaffolds using the Quasar CombiDesign module of MOE program. Chemical reagents considered in this study were taken from the directories of chemicals available from the commercial suppliers of chemicals (Available Chemicals Directory). Each analog was built as a neutral molecule in MOE program and its molecular geometry has been refined by the MM optimization, implemented in the Discovery Studio 2.5 smart minimizer. Convergence criteria (energy difference of 10^{-4} kcal mol⁻¹, root-mean-square displacement (RMSD) of 10^{-5} Å and a dielectric constant of 4 using the CFF force field) were set, as described in section 2.3.

3.12. In silico screening

The conformer with the best mapping on PH4 pharmacophore in each cluster was selected for virtual screening using the complexation QSAR model. For each E:I complex, the relative complexation Gibbs free energies, $\Delta\Delta G_{\text{comp}}$, was calculated. This was then used to compute the predicted activities (pIC_{50}^{pred}) of each of the newly designed analogs against FP-3. The IC_{50}^{pred} values were then calculated using the formular $IC_{50}^{\text{pred}} = 10^{-pIC_{50}^{\text{pred}}(\text{FP-3})}$.

3.13. Computing ADMET-related properties

The calculation of molecular descriptors related to adsorption, distribution, metabolism, excretion and toxicity of the designed inhibitors was carried out by the QikProp program (Schrödinger 2014). The models are based on the methods developed by Jorgensen and Duffy (Duffy and Jorgensen 2000;

Jorgensen and Duffy 2000; Jorgensen and Duffy 2002). The library focusing strategy was based on ADMET-related properties and the pharmacophore-based library focusing procedure, as previously described (Owono Owono et al. 2015; Esmel et al. 2017).

4. Conclusions

Structural information from the crystal structure of FP3-K11017 complex has been successfully used to establish of reliable QSAR model of non-covalent inhibition of the FP3 of *Pf* by peptidomimetics (PEP) inhibitors. This model correlates the unique descriptor, namely the computed Gibbs free energies (GFE) upon complex formation, with observed inhibitory potencies and is able to identify a few compounds which could be low nM range inhibitors of *P. falciparum*. As GFE is a combined descriptor involving the enthalpic gas phase, entropic contributions and solvation free energy, a precise insight into S1' and S2 pockets filling have been performed from the model by analysis of interactions between the enzyme active-site residues and the inhibitor. For this purpose, the breakdown of the interaction energy clearly indicated the residues involved in the affinity with the most active inhibitors. This information has helped to design an initial diversity virtual combinatorial library of new analogs to be screened by the pharmacophore models derived from the GFE QSAR. The focused library filtered by a set of ADME-related descriptors and screened by matching of the analogs to the PH4 pharmacophore permitted selection of a library subset of orally bioavailable PEP. This subset of 80 best virtual hits was submitted to the computation of predicted *Pf*FP-3 inhibitory potencies by the formerly established complexation QSAR model. The best analogues showed predicted activities in the low nanomolar concentration range, with the most promising hits being **31-1-12-13-1**, **24-1-12-1-21**, **2-1-1-16-2**, **2-1-12-21-30**, **2-1-30-7-2**, **2-6-12-21-30**, and **3-1-12-1-13** with predictive IC_{50}^{pred} of 0.2, 0.3, 0.01, 0.1, 0.01, 1, and 1 nM respectively against *Pf*FP-3. These are being proposed for synthesis and biological screening and may lead to a discovery of novel potent orally bioavailable anti-malarial.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: A-groups (fragments, building blocks, substituents) used in the design of the additional analogs. Dashed bonds indicate the attachment points of the fragments. Table S2: Molar masses and energy contributions towards $\Delta\Delta G_{comp}$ for most potent set of analogs inhibiting *Pf*FP-3 set as well as analogs with most frequently occurring fragments from the highly focused library of *Pf*FP-3 inhibitors.

Author Contributions: Conceptualization, L.C.O.O and E.M.; methodology, F.N.K., L.C.O.O and E.M.; software, F.N.K., L.C.O.O and E.M.; validation, E.E.A. B.D., M.K. and E.M.; formal analysis, E.E.A.; investigation, B.D.B., B.D. and E.E.A.; resources, F.N.K., L.C.O.O and E.M.; data curation, E.M.; writing—original draft preparation, B.D.B., B.D., F.N.K., E.E.A., M.K. and E.M.; writing—review and editing, B.D.B., B.D., F.N.K., E.E.A., M.K., L.C.O.O. and E.M.; supervision, L.C.O.O. and E.M.; project administration, E.M.; funding acquisition, B.D.B and F.N.K.

Funding: B.D.B. thanks the African-German Network of Excellence in Science (AGNES) for granting a Mobility Grant in 2017, generously sponsored by German Federal Ministry of Education and Research and the Alexander von Humboldt Foundation, Germany. F.N.K. acknowledges a Georg Forster Return Fellowship from the Alexander von Humboldt Foundation, Germany. Financial support for this work is acknowledged from a ChemJets fellowship from the Ministry of Education, Youth and Sports of the Czech Republic awarded to F.N.K.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Accelrys (2002) Cerius 2 Life Sciences molecular simulation software, 4.6 edn.,
- Accelrys (2005) Insight-II and Discover molecular modeling and simulation package, 2005 edn.,
- Accelrys (2009) Discovery Studio molecular modeling and simulation program, 2.5 edn.,
- Ang KKH, Ratnam J, Gut J, Legac J, Hansell E, Mackey ZB, Skrzypczyńska KM, Debnath A, Engel JC, Rosenthal PJ, McKerrow JH, Arkin MR, Renslo AR (2011) Mining a cathepsin inhibitor library for new antiparasitic drug leads. *PLoS Neglected Tropical Diseases* 5:e1023 doi:10.1371/journal.pntd.0001023
- Available Chemicals Directory.
- Bekono BD, Ntie-Kang F, Owono Owono LC, Megnassan E (2018) Targeting cysteine proteases from *Plasmodium falciparum*: a general overview, rational drug design and computational approaches for drug discovery. *Current Drug Targets* 19:501-526 doi:10.2174/1389450117666161221122432
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Research*. 28(1):235-342.
- Blackman M (2000) Proteases involved in erythrocyte invasion by the malaria parasite function and potential as chemotherapeutic. *Targets Current Drug Targets* 1:59-83 doi:10.2174/1389450003349461
- Böttcher CJF (1973) Non-linear effects. Elsevier. doi:10.1016/b978-0-444-41019-1.50013-4
- Chibale K, Musonda C (2003) The synthesis of parasitic cysteine protease and trypanothione reductase inhibitors. *Current Medicinal Chemistry* 10:1863-1889 doi:10.2174/0929867033456963
- Copeland RA, Lombardo D, Giannaras J, Decicco CP (1995) Estimating KI values for tight binding inhibitors from dose-response plots. *Bioorganic & Medicinal Chemistry Letters* 5:1947-1952 doi:10.1016/0960-894x(95)00330-v
- Coterón JM, Catterick D, Castro J, Chaparro MJ, Díaz B, Fernández E, Ferrer S, Gamo FJ, Gordo M, Gut J, de las Heras L, Legac J, Marco M, Miguel J, Muñoz V, Porrás E, de la Rosa JC, Ruiz JR, Sandoval E, Ventosa P, Rosenthal PJ, Fiandor JM (2010) Falcipain inhibitors: optimization studies of the 2-pyrimidinecarbonitrile lead series. *Journal of Medicinal Chemistry* 53:6129-6152 doi:10.1021/jm100556b
- Dali B, Keita M, Megnassan E, Freceer V, Miertus S (2012) Insight into selectivity of peptidomimetic inhibitors with modified statine core for plasmepsin II of *Plasmodium falciparum* over human cathepsin D. *Chemical Biology & Drug Design* 79:411-430 doi:10.1111/j.1747-0285.2011.01276.x
- Desai PV, Patny A, Gut J, Rosenthal PJ, Tekwani B, Srivastava A, Avery M (2006) Identification of novel parasitic cysteine protease inhibitors by use of virtual screening. 2. The Available Chemical Directory. *Journal of Medicinal Chemistry* 49:1576-1584 doi:10.1021/jm0505765
- Desai PV et al. (2004) Identification of novel parasitic cysteine protease inhibitors using virtual screening. 1. The ChemBridge database. *Journal of Medicinal Chemistry* 47:6609-6615 doi:10.1021/jm0493717
- Duffy EM, Jorgensen WL (2000) Prediction of Properties from simulations: free energies of solvation in hexadecane, octanol, and water. *Journal of the American Chemical Society* 122:2878-2888 doi:10.1021/ja993663t
- Dye C (2006) Global epidemiology of tuberculosis. *The Lancet* 367:938-940 doi:10.1016/s0140-6736(06)68384-0
- Eggleston KK, Duffin KL, Goldberg DE (1999) Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*. *Journal of Biological Chemistry* 274:32411-32417 doi:10.1074/jbc.274.45.32411
- Elliott DA, McIntosh MT, Hosgood HD, Chen S, Zhang G, Baeovova P, Joiner KA (2008) Four distinct pathways of hemoglobin uptake in the malaria parasite *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences* 105:2463-2468 doi:10.1073/pnas.0711067105

- Esmel A, Keita M, Eugene M, Toi B, Frecer V, Miertus S (2017) Insight into binding mode of nitrile inhibitors of *Plasmodium falciparum* falcipain-3, QSAR and pharmacophore models, virtual design of new analogues with favorable pharmacokinetic profiles. SDRP Journal of Computational Chemistry & Molecular Modelling 2 doi:10.25177/jcmp.2.1.5
- Fischer S, Smith JC, Verma CS (2001) Dissecting the vibrational entropy change on protein/ligand binding: burial of a water molecule in bovine pancreatic trypsin inhibitor. The Journal of Physical Chemistry B 105:8050-8055 doi:10.1021/jp0120920
- Francis SE, Sullivan DJ, Goldberg, Daniel E (1997) hemoglobin metabolism in the malaria parasite *Plasmodium falciparum*. Annual Review of Microbiology 51:97-123 doi:10.1146/annurev.micro.51.1.97
- Frecher V, Berti F, Benedetti F, Miertus S (2008) Design of peptidomimetic inhibitors of aspartic protease of HIV-1 containing -PheΨPro- core and displaying favourable ADME-related properties. Journal of Molecular Graphics and Modelling 27:376-387 doi:10.1016/j.jmglm.2008.06.006
- Frecher V, Jedinak A, Tossi A, Berti F, Benedetti F, Romeo D, Miertus S (2005) Structure-based design of inhibitors of aspartic protease of HIV-1. Letters in Drug Design & Discovery 2:638-646 doi:10.2174/157018005774717307
- Frecher V, Miertus S (2002) Interactions of ligands with macromolecules: Rational design of specific inhibitors of aspartic protease of HIV-1. Macromolecular Chemistry and Physics 203:1650-1657 doi:10.1002/1521-3935(200207)203:10/11<1650::aid-macp1650>3.0.co;2-e
- Frecher V, Miertuš S (1992) Polarizable continuum model of solvation for biopolymers. International Journal of Quantum Chemistry 42:1449-1468 doi:10.1002/qua.560420520
- Frecher V, Miertus S, Tossi A, Romeo D (1998) Rational design of inhibitors for drug-resistant HIV-1 aspartic protease mutants. Drug Design and Discovery 15:211-231
- Frecher V, Seneci P, Miertus S (2010) Computer-assisted combinatorial design of bicyclic thymidine analogs as inhibitors of *Mycobacterium tuberculosis* thymidine monophosphate kinase. Journal of Computer-Aided Molecular Design 25:31-49 doi:10.1007/s10822-010-9399-4
- Frecher V, Kabeláč M, De Nardi P, Pricl S, Miertuš S (2004) Structure-based design of inhibitors of NS3 serine protease of hepatitis C virus. Journal of Molecular Graphics and Modelling 22:209-220 doi:10.1016/s1093-3263(03)00161-x
- Freire E (2008) Do enthalpy and entropy distinguish first in class from best in class? Drug Discovery Today 13:869-874 doi:<https://doi.org/10.1016/j.drudis.2008.07.005>
- Freundlich JS et al. (2009) Triclosan derivatives: towards potent inhibitors of drug-sensitive and drug-resistant *Mycobacterium tuberculosis*. ChemMedChem 4:241-248 doi:10.1002/cmdc.200800261
- Ghasemi J, Shiri F, Pirhadi S, Heidari Z (2015) Discovery of new potential antimalarial compounds using virtual screening of ZINC database. Combinatorial Chemistry & High Throughput Screening 18:227-234 doi:10.2174/1386207318666141229123705
- Gilson M, Honig B (1991) The inclusion of electrostatic hydration energies in molecular mechanics calculations. Journal of Computer-Aided Molecular Design 5 doi:10.1007/bf00173467
- Jorgensen WL, Duffy EM (2000) Prediction of drug solubility from Monte Carlo simulations. Bioorganic & Medicinal Chemistry Letters 10:1155-1158 doi:10.1016/s0960-894x(00)00172-4
- Jorgensen WL, Duffy EM (2002) Prediction of drug solubility from structure. Advanced Drug Delivery Reviews 54:355-366 doi:10.1016/s0169-409x(02)00008-x
- Keita M, Kumar A, Dali B, Megnassan E, Siddiqi MI, Frecer V, Miertus S (2014) Quantitative structure-activity relationships and design of thymine-like inhibitors of thymidine monophosphate kinase of

- Mycobacterium tuberculosis* with favourable pharmacokinetic profiles. RSC Adv 4:55853-55866 doi:10.1039/c4ra06917j
- Kerr ID et al. (2009a) Vinyl sulfones as antiparasitic agents and a structural basis for drug design. Journal of Biological Chemistry 284:25697-25703 doi:10.1074/jbc.m109.014340
- Kerr ID, Lee JH, Pandey KC, Harrison A, Sajid M, Rosenthal PJ, Brinen LS (2009b) Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. Journal of Medicinal Chemistry 52:852-857 doi:10.1021/jm8013663
- Kily Herve Fagnidi Y, Toi B, Megnassan E, Frecer V, Miertus S (2018) *In silico* design of *Plasmodium falciparum* cysteine protease falcipain 2 inhibitors with favorable pharmacokinetic profile. Journal of Analytical & Pharmaceutical Research 7 doi:10.15406/japlr.2018.07.00244
- Klemba M, Gluzman I, Goldberg DE (2004) A *Plasmodium falciparum* dipeptidyl aminopeptidase i participates in vacuolar hemoglobin degradation. Journal of Biological Chemistry 279:43000-43007 doi:10.1074/jbc.m408123200
- Kouassi A et al. (2015) Computer-aided design of orally bioavailable pyrrolidine carboxamide inhibitors of enoyl-acyl carrier protein reductase of *Mycobacterium tuberculosis* with favorable pharmacokinetic profiles. International Journal of Molecular Sciences 16:29744-29771 doi:10.3390/ijms161226196
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings 1 PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3–25. 1 Advanced Drug Delivery Reviews 46:3-26 doi:10.1016/s0169-409x(00)00129-0
- Liu J, Gluzman IY, Drew ME, Goldberg DE (2004) The role of *Plasmodium falciparum* food vacuole plasmepsins. Journal of Biological Chemistry 280:1432-1437 doi:10.1074/jbc.m409740200
- Maple JR, Hwang MJ, Stockfisch TP, Dinur U, Waldman M, Ewig CS, Hagler AT (1994) Derivation of class II force fields. I. Methodology and quantum force field for the alkyl functional group and alkane molecules. Journal of Computational Chemistry 15:162-182 doi:10.1002/jcc.540150207
- Megnassan E, Keita M, Bieri C, Esmel A, Frecer V, Miertus S (2012) Design of novel dihydroxynaphthoic acid inhibitors of *Plasmodium falciparum* lactate dehydrogenase. Medicinal Chemistry 8:970-984 doi:10.2174/157340612802084324
- Miertuš S, Scrocco E, Tomasi J (1981) Electrostatic interaction of a solute with a continuum. A direct utilization of *ab initio* molecular potentials for the prevision of solvent effects. Chemical Physics 55:117-129 doi:10.1016/0301-0104(81)85090-2
- Miller Louis H, Su X (2011) Artemisinin: discovery from the chinese herbal garden. Cell 146:855-858 doi:10.1016/j.cell.2011.08.024
- Chemical Computing Group (2014) Molecular Operating Environment (MOE), 2014 edn.,
- N'Guessan A, Megnassan E, Ziao N, Frecer V, N'Guessan YT, Miertus S (2017) Artemisinin-dipeptidyl vinyl sulfone hybrid inhibitors of *Plasmodium falciparum* falcipain 2 with favorable pharmacokinetic profile. Journal of Drug Design and Development 1:11-28
- N'Guessan H, Megnassan E (2017) *In silico* design of phosphonic arginine and hydroxamic acid inhibitors of *Plasmodium falciparum* M17 leucyl aminopeptidase with favorable pharmacokinetic profiles. Journal of Drug Design and Medicinal Chemistry 3:98 doi:10.11648/j.jddmc.20170306.13
- Nethavhani SA, van Ree T (2014) Synthesis and antimalarial activities of some novel 2-pyridones. Arabian Journal for Science and Engineering 39:6595-6598 doi:10.1007/s13369-014-1193-5

- Olson J (1999) Antimalarial effects in mice of orally administered peptidyl cysteine protease inhibitors. *Bioorganic & Medicinal Chemistry* 7:633-638 doi:10.1016/s0968-0896(99)00004-8
- Owono Owono LC, Keita M, Megnassan E, Frecer V, Miertus S (2013) Design of Thymidine analogues targeting thymidilate kinase of *Mycobacterium tuberculosis*. *Tuberculosis Research and Treatment* 2013:1-13 doi:10.1155/2013/670836
- Owono Owono LC, Ntie-Kang F, Keita M, Megnassan E, Frecer V, Miertus S (2015) Virtually Designed triclosan-based inhibitors of enoyl-acyl carrier protein reductase of *Mycobacterium tuberculosis* and of *Plasmodium falciparum*. *Molecular Informatics* 34:292-307 doi:10.1002/minf.201400141
- Potshangbam AM, Tanneeru K, Reddy BM, Guruprasad L (2011) 3D-QSAR and molecular docking studies of 2-pyrimidinecarbonitrile derivatives as inhibitors against falcipain-3. *Bioorganic & Medicinal Chemistry Letters* 21:7219-7223 doi:10.1016/j.bmcl.2011.09.107
- Ramjee Manoj K, Flinn Nicholas S, Pemberton Tracy P, Quibell M, Wang Y, Watts John P (2006) Substrate mapping and inhibitor profiling of falcipain-2, falcipain-3 and berghepains-2: implications for peptidase anti-malarial drug discovery. *Biochemical Journal* 399:47-57 doi:10.1042/bj20060422
- Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, Honig B (2001) Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *Journal of Computational Chemistry* 23:128-137 doi:10.1002/jcc.1161
- Rosenthal P, Sijwali P, Singh A, Shenai B (2002) Cysteine proteases of malaria parasites: targets for chemotherapy. *Current Pharmaceutical Design* 8:1659-1672 doi:10.2174/1381612023394197
- Rosenthal PJ, Lee GK, Smith RE (1993) Inhibition of a *Plasmodium vinckei* cysteine proteinase cures murine malaria. *Journal of Clinical Investigation* 91:1052-1056 doi:10.1172/jci116262
- Rosenthal PJ, Olson JE, Lee GK, Palmer JT, Klaus JL, Rasnick D (1996) Antimalarial effects of vinyl sulfone cysteine proteinase inhibitors. *Antimicrobial Agents and Chemotherapy* 40:1600-1603 doi:10.1128/aac.40.7.1600
- Rosenthal PJ, Wollish WS, Palmer JT, Rasnick D (1991) Antimalarial effects of peptide inhibitors of a *Plasmodium falciparum* cysteine proteinase. *Journal of Clinical Investigation* 88:1467-1472 doi:10.1172/jci115456
- Schrödinger (2014) QikProp, 3.7, release 14 edn.,
- Schwarzl SM, Tschopp TB, Smith JC, Fischer S (2002) Can the calculation of ligand binding free energies be improved with continuum solvent electrostatics and an ideal-gas entropy correction? *Journal of Computational Chemistry* 23:1143-1149 doi:10.1002/jcc.10112
- Shah F, Mukherjee P, Desai P, Avery M (2010) Computational approaches for the discovery of cysteine protease inhibitors against malaria and SARS. *Current Computer Aided-Drug Design* 6:1-23 doi:10.2174/157340910790980142
- Shah F, Mukherjee P, Gut J, Legac J, Rosenthal PJ, Tekwani BL, Avery MA (2011) Identification of novel malarial cysteine protease inhibitors using structure-based virtual screening of a focused cysteine protease inhibitor library. *Journal of Chemical Information and Modeling* 51:852-864 doi:10.1021/ci200029y
- Sijwali PS, Koo J, Singh N, Rosenthal PJ (2006) Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* 150:96-106 doi:10.1016/j.molbiopara.2006.06.013
- Verissimo E et al. (2008) Design and synthesis of novel 2-pyridone peptidomimetic falcipain 2/3 inhibitors.

Bioorganic & Medicinal Chemistry Letters 18:4210-4214 doi:10.1016/j.bmcl.2008.05.068

Weldon DJ et al. (2014) Synthesis, biological evaluation, hydration site thermodynamics, and chemical reactivity analysis of α -keto substituted peptidomimetics for the inhibition of *Plasmodium falciparum*. Bioorganic & Medicinal Chemistry Letters 24:1274-1279 doi:10.1016/j.bmcl.2014.01.062

White SW, Zheng J, Zhang Y-M, Rock CO (2005) The structural biology of type ii fatty acid biosynthesis. Annual Review of Biochemistry 74:791-831 doi:10.1146/annurev.biochem.74.082803.133524

WHO (2015) World Malaria Report.

Yeh I, Altman R (2006) Drug targets for *Plasmodium falciparum*: a post-genomic review/survey. Mini-Reviews in Medicinal Chemistry 6:177-202 doi:10.2174/138955706775475957

Zhang Y-M, Lu Y-J, Rock CO (2004) The reductase steps of the type II fatty acid synthase as antimicrobial targets. Lipids 39:1055-1060 doi:10.1007/s11745-004-1330-3



© 2019 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).