Article

QSAR-Based Virtual Screening of Natural Products Database for Identification of Potent Antimalarial Hits

Letícia Tiburcio Ferreira [1] , Joyce V. B. Borba [1,2] , José Teófilo Moreira-Filho [2], Aline Rimoldi [1], Carolina Horta Andrade [2] and Fabio Trindade Maranhão Costa [1,*]

[1] Laboratory of Tropical Diseases Prof. Dr. Luiz Jacintho da Silva, Department of Genetics, Evolution, Microbiology and Immunology, University of Campinas-UNICAMP, Campinas, SP 13083-864, Brazil; ltiburciof@gmail.com (L.T.F.); joycevillaverde@gmail.com (J.V.B.B.); rimoldialine@gmail.com (A.R.)
[2] Laboratory of Molecular Modeling and Drug Design, LabMol, Faculty of Pharmacy, Federal University of Goiás, Goiânia, GO 74605-170, Brazil; teofarma1@gmail.com (J.T.M.-F.); andradech@yahoo.com (C.H.A.)
[*] Correspondence: fabiotmc72@gmail.com; Tel.: +55-19-3521-6288

Abstract: With about 400,000 annual deaths worldwide, malaria remains a public health burden in tropical and subtropical areas, especially in low-income countries. Selection of drug-resistant *Plasmodium* strains has driven the need to explore novel antimalarial compounds with diverse modes of action. In this context, biodiversity has been widely exploited as a resourceful channel of biologically active compounds, as exemplified by antimalarial drugs such as quinine and artemisinin, derived from natural products. Thus, combining a natural product library and quantitative structure–activity relationship (QSAR)-based virtual screening, we have prioritized genuine and derivative natural compounds with potential antimalarial activity prior to in vitro testing. Experimental validation against cultured chloroquine-sensitive and multi-drug-resistant *P. falciparum* strains confirmed the potent and selective activity of two sesquiterpene lactones (LDT-597 and LDT-598) identified *in silico*. Quantitative structure–property relationship (QSPR) models predicted absorption, distribution, metabolism, and excretion (ADME) and physiologically based pharmacokinetic (PBPK) parameters for the most promising compound, showing that it presents good physiologically based pharmacokinetic properties both in rats and humans. Altogether, the in vitro parasite growth inhibition results obtained from in silico screened compounds encourage the use of virtual screening campaigns for identification of promising natural compound-based antimalarial molecules.

Keywords: *Plasmodium falciparum*; natural products; virtual screening; experimental validation; QSAR; ADME

1. Introduction

Malaria is a mosquito-borne disease transmitted by the bite of Anopheles mosquitoes infected with *Plasmodium* parasites. Mainly caused by *Plasmodium falciparum* and *Plasmodium vivax*, malaria still imposes a heavy burden upon developing countries as it is responsible for high rates of morbidity and mortality. In 2018, the World Health Organization (WHO) recorded 228 million cases of the disease worldwide with an estimate of approximately 400,000 malaria-related deaths [1].

The decreasing rate of cases tracked by the WHO in the last decade has stalled in the last 5 years [1], portraying how fragile the gains achieved can be. The COVID-19 pandemic, which arose in 2019, menaces these gains even further, as the spread of SARS-CoV-2 leads to extra pressure on health systems, especially in low-resourced settings [2]. It is estimated that disruption of malaria control endeavors in Africa due to the COVID-19 pandemic could lead to a malaria burden in 2020 that is the double of that registered in 2019 [3]. Moreover, selection of parasite strains resistant to currently available antimalarial drugs [4] has raised serious concerns about maintaining global achievements in the battle against malaria. The scenario of parasite resistance has threatened the front-line treatment with

Artemisinin Combined Therapies (ACTs), as the clinical efficacy of artemisinin (ART) and its derivatives has been compromised by delayed parasite clearance in Southeast Asia [5]. Even more alarming is the recent evidence of *de novo* emergence of *Pf* kelch13-mediated artemisinin resistance in Africa [6]. Together, these reports highlight the compelling call for new therapies based on alternative molecules with potent activity against the malaria parasite, especially drug-resistant *Plasmodium* strains.

The storyline of antimalarial drug development cannot be distinguished from the relevance of biodiversity-derived molecules. One of the first antimalarial indications in history, the use of *Chinchona* decoctions for recurring fevers even before the characterization of malaria parasites [7], motivated the isolation of quinine—an alkaloid with antiplasmodial properties and one of the most representative antimalarial drugs to date. Likewise, the use of *Artemisia annua* decoctions for fever treatment inspired the description of its active principle, artemisinin, which now constitutes the first line of malaria treatment indicated by the WHO [8].

For decades, the discovery of therapeutic molecules derived from natural compounds relied solely on laborious rounds of phenotypic screenings and subsequent cycles of compound purification [9], which is an onerous and uncertain process. However, progress in computational software and the availability of massive online commercial datasets of chemical compounds with their biological properties annotated [10,11] have boosted the development of computer-aided drug design (CADD). In this new setting of drug discovery and development, CADD opens avenues for reducing time and costs associated with identification of new drug candidates and consequent more effective pathways until the workbench [12,13]. Among CADD techniques, quantitative structure–activity relationships (QSARs) have been widely applied to hit identification by virtual screening [14–16] in order to predict physicochemical properties and biological activity of molecules. Different QSAR-based virtual screenings have been applied by our group for the effective description of new candidate hits for neglected tropical diseases, including malaria [17–20]. In light of this, commercial natural products libraries link virtual screening campaigns to the benefits of natural product-based drug discovery. Even though such libraries tend to deliver molecules not synthetically tractable due to their complex scaffolds, their advantages rely on the higher structural and physicochemical diversity, and wider coverage of the chemical space of natural products compared to synthetic drugs [21].

The goal of the present study was to develop a virtual screening pipeline for identifying potent natural products and natural product-derived molecules with antimalarial activity available in a commercial database by integrating QSAR-based virtual screening and further experimental evaluation of chloroquine-sensitive and multi-drug-resistance *P. falciparum* strains, in addition to testing their in vitro therapeutic indexes in human hepatoma cells. Finally, we analyzed multiple absorption, distribution, metabolism, and excretion (ADME) and physiologically based pharmacokinetic (PBPK) parameters to ensure drug efficacy and tolerability.

2. Materials and Methods

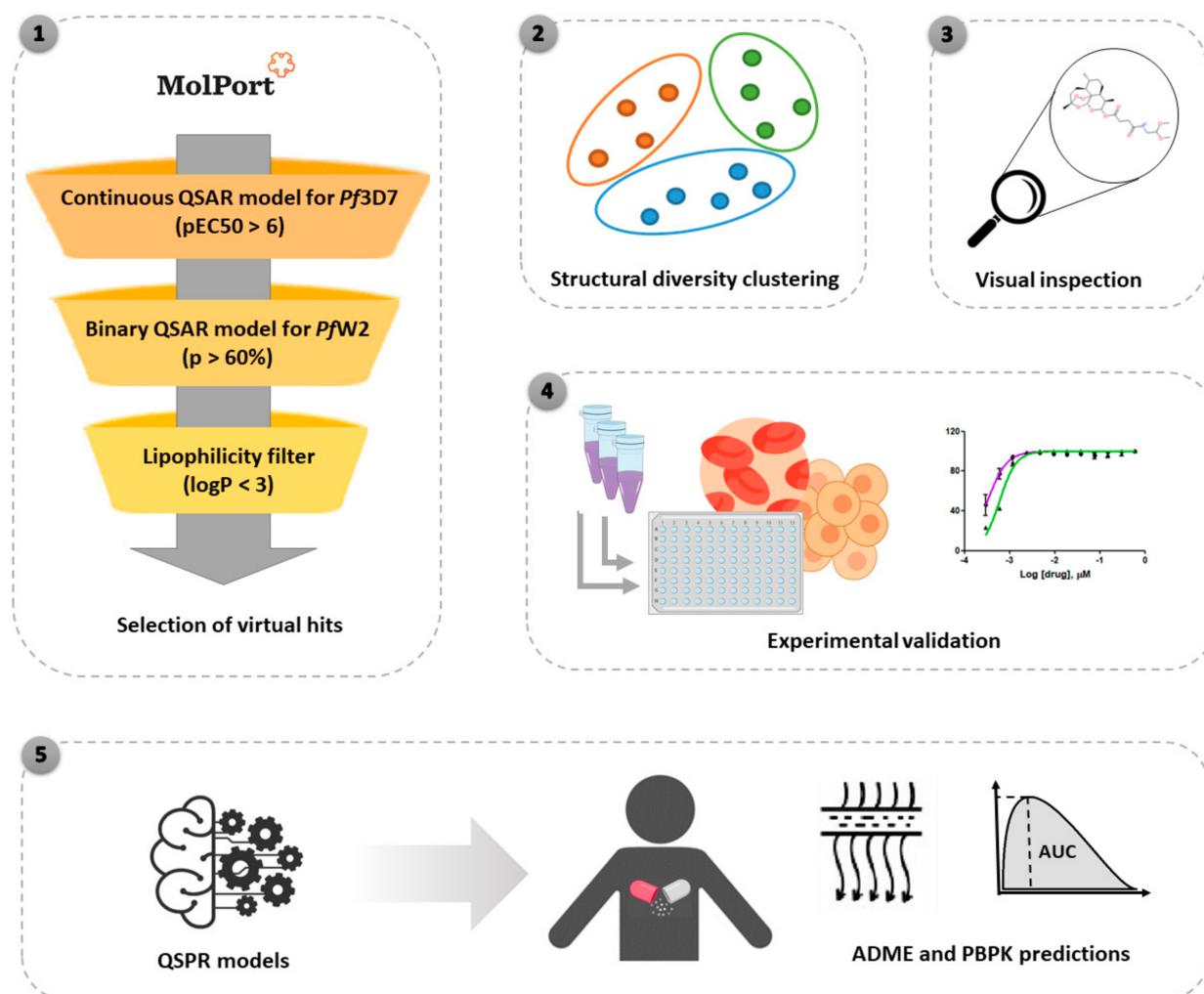The workflow for the study design is summarized in Figure 1.

Figure 1. General workflow of the virtual screening of natural compound database followed by experimental validation. The following steps were conducted: (**1**) quantitative structure–activity relationship (QSAR)-based virtual screening of MolPort natural products and derivatives database and selection of compounds with best predicted antimalarial activity; (**2**) analysis of structural diversity by clustering; (**3**) visual inspection; (**4**) experimental validation against *P. falciparum* blood stages and mammalian HepG2 cells and (**5**) quantitative structure–property relationship (QSPR) models for prediction of absorption, distribution, metabolism, and excretion (ADME) and physiologically based pharmacokinetic (PBPK) properties.

2.1. Virtual Screening and Structural Diversity Clustering

Steps were performed in our in-house QSAR workflow implemented on KNIME 3.2.2 [22]. In-house QSAR models [23] were applied for virtual screening of the natural products branch of MolPort commercial database (www.molport.com, accessed on 3 May 2019). The virtual screening aims to identify, within this database, a subset of compounds with potential biological activity against *P. falciparum* chloroquine-sensitive (3D7) and multi-drug-resistant (W2) strains. Prior to QSAR-based virtual screening predictions, the compounds were curated according to data curation protocols established by Fourches and colleagues [24–26]. Then, QSAR models were used to predict the *P. falciparum* 3D7 and W2 inhibition activities of compounds. The in-house QSAR models used for virtual screening were previously developed for 3D7 and W2 strains using learning sets compiled from PubChem database (PubChem IDs for 3D7 dataset: AID_1828, AID_449703, AID_524790, AID_660866 and for W2 dataset: AID_1883, AID_449704, AID_524796, AID_606570). These datasets were compiled and curated according to best practices of QSAR modeling and after curation, the 3D7 learning set contained 7873 compounds (3497 actives and 4376 inactives) and the W2 learning set contained 7403 compounds (3637 actives and 3766

inactives). Both datasets were then used to build consensus models with a Random Forest algorithm and five different molecular descriptors (Avalon, MACCS, Morgan, FeatMorgan and AtomPair) [23]. Three criteria were used for selection of virtual hits: (i) predicted pEC_{50} against *P. falciparum* 3D7 should be ≥ 6; (ii) probability of activity (p) against *P. falciparum* W2 should be ≥ 0.6 (p > 60%) and (iii) logP filters were also added to predict good lipophilicity with logP < 3 using XLogP [27]. For prioritization of structurally diverse compounds, molecules predicted to be active by the virtual screening were clustered through the Butina method [28] implemented in Python 3.6 and using the workflow proposed by Sydow and colleagues [29], which groups compounds based on Tanimoto similarity and picks a set of diverse compounds from these groups. Finally, the selected virtual hits were purchased and submitted to in vitro experimental evaluation. The similarity map and was generated using OSIRIS DataWarrior software v.05.02.01 [30].

2.2. Compound Preparation

All natural compounds and derivatives selected were purchased from MolPort and dissolved in DMSO at 10 mM. Chloroquine and artesunate (standard antimalarial drugs) were purchased from Sigma-Aldrich (St. Louis, MO, USA) and prepared at 10 mM as well.

2.3. Plasmodium falciparum In Vitro Culture

Chloroquine-sensitive (3D7) and multi-drug-resistant (W2) *P. falciparum* strains were cultured following the candle jar method as described by Trager and Jensen [31]. Briefly, parasites were maintained in RPMI 1640 medium supplemented with 10% A+ human plasma at 5% CO_2 atmosphere. Ring stage synchronized cultures were obtained by two consecutive treatments with a 5% D-sorbitol solution [32] in 48 h intervals.

2.4. In Vitro Assays for P. falciparum Growth Inhibition

Parasite growth inhibition assays were conducted by distributing ring stage synchronized parasites at 2% hematocrit and 0.5% parasitemia in 96-well plates (NEST Biotechnology Co., Ltd., Wuxi, Jiangsu, China). Parasites were incubated in the presence of drugs in a two-fold 12-point serial dilution starting at 10 μM in duplicate. Chloroquine was used as an antimalarial standard. At the end of a 72 h incubation, parasitemia was determined by fluorescence reading at 490 nm excitation and 540 nm emission (BMG CLARIOstar, BMG Labtech Inc., Durham, NC, USA) using SybrGreen fluorescent dye as described by Hartwig et al. [33]. Parasite growth inhibition was determined as a percentage relative to drug-free control. EC_{50} values were interpolated from log doses vs. inhibition curves in GraphPad Prism 6 (GraphPad, La Jolla, CA, USA).

2.5. Citotoxicity Assays

Cytotoxicity of selected compounds was evaluated using the MTT reduction assay (3-[4,5-dimethyl-thiazol-2-yl]-2,5-diphenyltetrazolium chloride) [34]. Human hepatoma (HepG2) cells were cultured in Dulbecco's Modified Eagle Medium supplemented with gentamycin (40 mg/L) and 10% heat-inactivated fetal bovine serum at 37 °C and 5% CO_2. Cells were seeded in 96-well plates (NEST Biotechnology Co., Ltd., Wuxi, Jiangsu, China) at 10^5 cells per well and incubated in the presence of a serial dilution of the drugs starting at 100 μM. After 72 h of incubation, MTT was added to the wells. The optical density was measured at 570 nm (CLARIOStar, Labtech BMG Inc, Durham, NC, USA) and cell viability was expressed as a percentage relative to the untreated control. CC_{50} values were calculated by plotting a log dose vs. viability curve in GraphPad Prism 6 (GraphPad, La Jolla, CA, USA). The selectivity index (SI) of the compounds was determined by the expression:

$$SI = (HepG2\ CC_{50})/(Pf\ EC_{50}),$$

where HepG2 CC_{50} corresponds to the cytotoxic concentration of compounds in HepG2 cells and *Pf* EC_{50} relates to the 50% inhibitory concentration on *P. falciparum* 3D7 strain.

2.6. In Silico Predictions of Metabolism, ADME and PBPK

For the in silico predictions of the compound's metabolism, we used the web tool BioTransformer [35], which is a software tool that predicts small molecule metabolism in mammals, their gut microbiota, as well as the soil/aquatic microbiota. The program Detoxie® (http://insilicall.com/, accessed on 17 December 2020) was used to predict absorption, distribution, metabolism, and excretion (ADME), and physiologically based pharmacokinetic (PBPK) properties. This software is an artificial intelligence de-risking application based on quantitative structure–property relationship (QSPR) models for each endpoint and human PBPK model.

3. Results

3.1. QSAR-Based Virtual Screening

In-house built and validated QSAR models were used for virtual screening, whose results are summarized in Figure 2. The entire MolPort database of commercially available natural products and derivatives containing approximately 120,000 chemical compounds was screened for identification of potentially active compounds against both *P. falciparum* chloroquine-sensitive (3D7) and multi-drug-resistant (W2) strains. Compounds were downloaded and prepared for screening by analyzing their chemical structures according to data curation protocols proposed by Fourches et al. [24–26]. Briefly, explicit hydrogens were added, salts were removed and specific chemotypes were normalized. Moreover, polymers, inorganic salts, organometallic compounds and mixtures were also removed. The first filter applied, which predicted activity against *P. falciparum* 3D7 on a continuous QSAR model (pEC$_{50}$ $\geq$ 6) selected 41,207 compounds. Next, the binary model for activity against multi-drug-resistant *P. falciparum* W2 (p $\geq$ 0.6) narrowed the list down to 1257 compounds predicted to be active. Once poor physicochemical properties could be a relevant bottleneck in late-stage drug development, we added a filter to select compounds with good lipophilicity (logP < 3), which resulted in 265 compounds being selected after the virtual screening pipeline (Supplemental File S1).
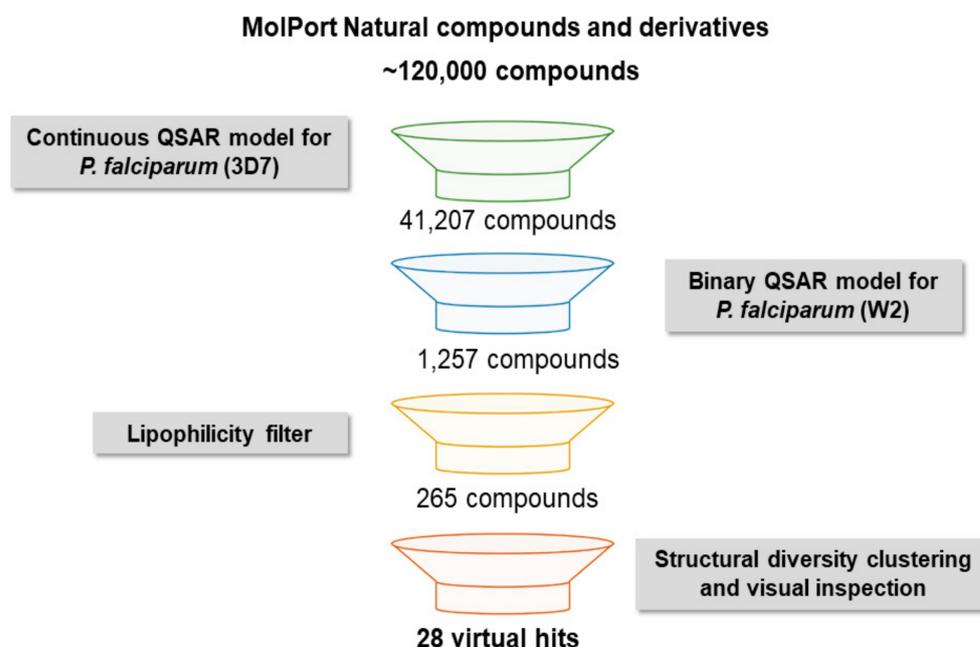

Figure 2. Virtual screening workflow for the identification of natural compounds and derivatives active against *P. falciparum*.

3.2. Structural Diversity Clustering

Starting with the 265 virtual hits obtained from screening the MolPort natural products and derivatives database, we performed an analysis of structural diversity based on

clustering of structurally similar compounds. The grouping search for a centroid compound which consists of compounds containing bigger amounts of neighbor compounds (or compounds that are structurally similar) and selects 10 of the nearest neighbor compounds to form a cluster. We observed that among more than 120 clusters obtained, more than half contained only one molecule—i.e., the chemical space itself is already diverse (Supplementary Figure S1). Inside each cluster, a structure–activity relationship (SAR) analysis was performed in order to identify the compounds predicted to be the most active within each structural group. Afterwards, just a few compounds were selected from each cluster in order to obtain the most structural diversity possible. Finally, we obtained 188 compounds with significative structural diversity. Among those, 28 were selected by visual inspection.

3.3. In Vitro Screening against P. falciparum

The 28 virtual hits selected were purchased and experimentally evaluated. In order to optimize the screening process, we carried out a preliminary in vitro screen to determine each compound's ability to inhibit the growth of the chloroquine-sensitive *P. falciparum* 3D7 strain at a concentration of 5 μM (Figure 3). Among the 28 compounds tested, 8 (LDT-598, LDT-588, LDT-599, LDT-597, LDT-592, LDT-614, LDT-586 and LDT-601) inhibited more than 70% of parasite growth at 5 μM. Therefore, we aimed to investigate the 50% inhibitory concentration (EC_{50}) of the eight most promising compounds against both chloroquine-sensitive and multi-drug-resistant parasite strains (Table 1).
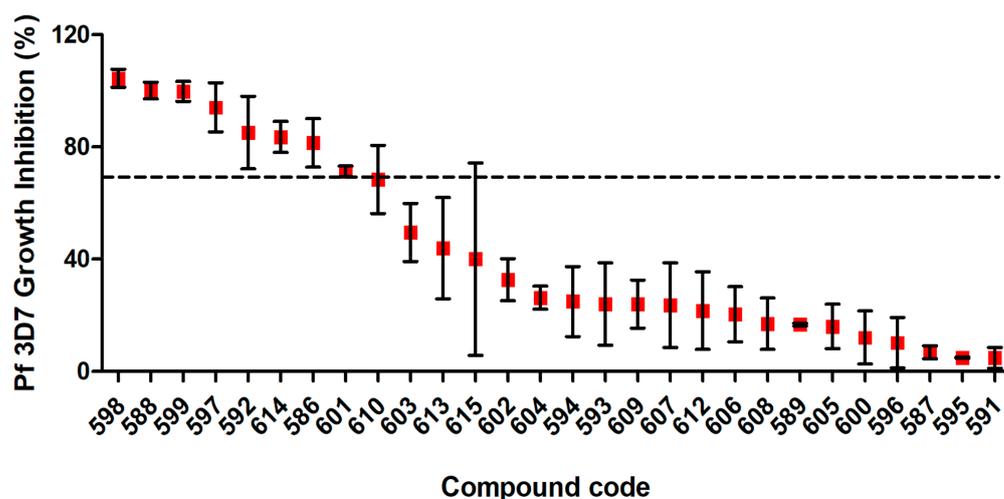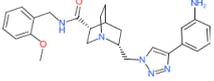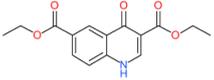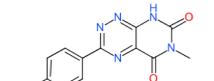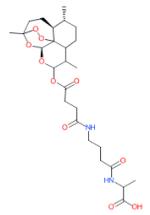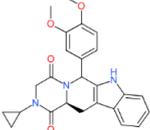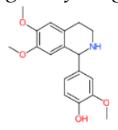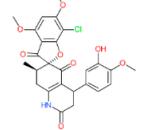

Figure 3. In vitro growth inhibition of asexual blood stage *P. falciparum* (3D7) for prioritized natural compounds and derivatives from virtual screening. The inhibitory potential of different compounds was tested at a concentration of 5 μM and the inhibition of parasitemia was measured after 72 hours of incubation. The dashed line represents the cutoff used to highlight the most promising compounds above 70% of inhibition.

Table 1 shows the five compounds (LDT-586, LDT-588, LDT-597, LDT-598 and LDT-599) that demonstrated good antimalarial activity with $EC_{50} < 5$ μM for both *P. falciparum* strains tested, including a multi-drug-resistant one (W2). It is worth noting that compounds LDT-597 and LDT-598 showed remarkable activity with EC_{50} of 0.54 and 0.78 nM, respectively, against *P. falciparum* 3D7 and EC_{50} of 0.56 and 0.59 nM against *P. falciparum* W2.

Table 1. The most promising compounds predicted to be active against *P. falciparum* asexual stages selected by virtual screening.

Compound Code (MolPort ID)	2D Structure (Biological Source of NP Precursor)	EC$_{50}$ [a] (μM)		CC$_{50}$ [b] (μM)	In vitro Therapeutic Index [c]
		*Pf*3D7	*Pf*W2	HepG2	
LDT-586 (MolPort-001-745-423)	(No data available)	4.52 ± 0.91	3.44 ± 1.30	98.59 ± 0	21.81
LDT-588 (MolPort-000-651-065)	(No data available)	3.84 ± 1.45	2.09 ± 1.17	67.04 ± 2.28	17.46
LDT-592 (MolPort-002-323-504)	(No data available)	9.09 ± 3.74	3.11 ± 0.54	17 ± 0	1.87
LDT-597 (MolPort-001-732-360)	(*Artemesia annua*)	0.0005 ± 0.00	0.0005 ± 0.00	18.29 ± 3.51	33,870.37
LDT-598 (MolPort-001-732-370)	(*Artemesia annua*)	0.0007 ± 0.00	0.0006 ± 0.00	25.94 ± 1.13	33,299.10
LDT-599 (MolPort-001-737-485)	(*Aspergillus fumigatus*)	3.68 ± 1.92	2.74 ± 0.78	20.96 ± 2.51	5.70
LDT-601 (MolPort-002-506-405)	(*Pachycereus weberi*, *Pachycereus pringlei*, *Pachycereus pecten-aboriginum*, *Backebergia militaris* and *Carnegiea gigantea*)	6.61 ± 3.20	0.65 ± 0.47	21.79 ± 4.16	3.30
LDT-614 (MolPort-044-180-513)	(*Penicillium patulum*)	5.26 ± 0.52	5.65 ± 3.11	23.55 ± 1.82	4.48
Chloroquine		0.0079 ± 0.00	0.147 ± 0.04	ND	-
Artesunate		0.0016 ± 0.00	ND	ND	-
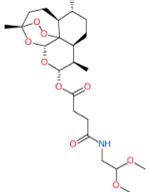
[a] EC$_{50}$: half of the maximum inhibitory concentration in 3D7 and W2 strains and their respective standard deviations; [b] CC$_{50}$: half the maximum cytotoxic concentration in HepG2 cells; [c] SI: Selectivity index calculated from CC$_{50}$/EC$_{50}$ (3D7). NP: natural products. ND: not determined. The data derive from at least two independent experiments.

3.4. Cytotoxicity against Human Cells

Cytotoxicity was assessed in human hepatoma cells (HepG2). Among the eight compounds tested, four of them showed favorable therapeutic indexes for hit compounds with an SI > 10 (LDT-586, LDT-588, LDT-597 and LDT-598). Two compounds showed interesting results as they were highly selective: LDT-597 and LDT-598 showed in vitro therapeutic indexes of 33,870.37 and 33,299.1, respectively, likely due to their low nanomolar half maximal inhibitory concentrations (EC_{50}) in *P. falciparum* 3D7.

3.5. In Silico Predictions of Metabolism, ADME and PBPK

We have also analyzed the metabolism of LDT-597 and LDT-598 using the software BioTransformer. The program showed that the carboxylic esters of these compounds might be hydrolyzed into dihydroartemisinin (DHA) (Figure 4) by plasma carboxylesterases, showing the same type of metabolization of artesunate [36].
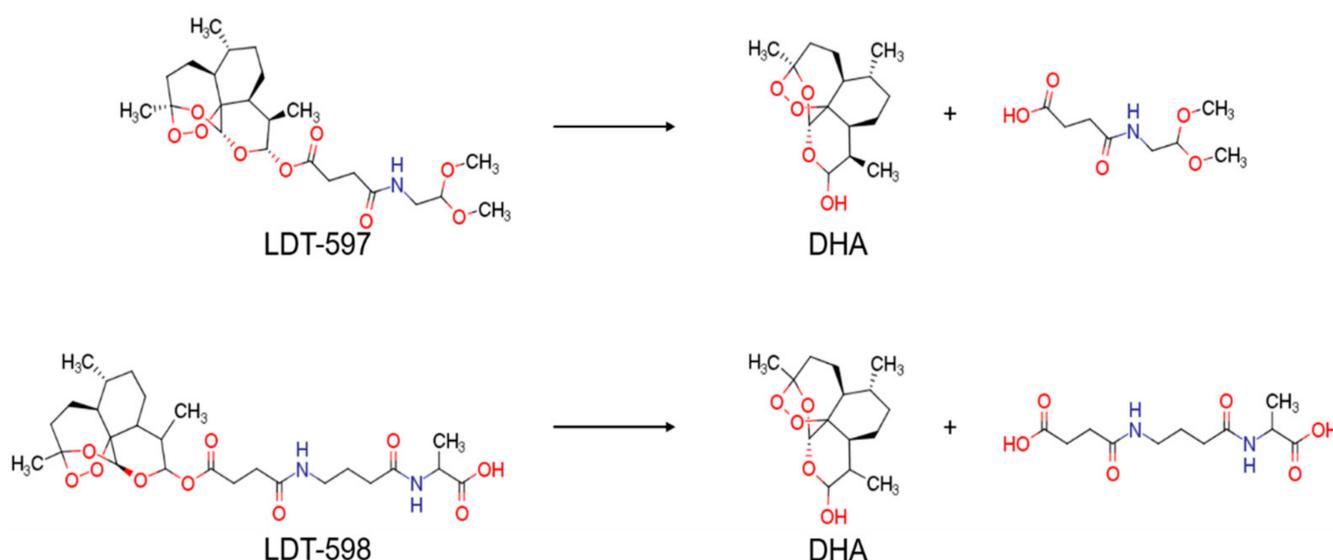

Figure 4. Predicted metabolism of compounds LDT-597 and -598 via plasma carboxylesterases predicted using the software BioTransformer.

Artesunate has significantly greater solubility in water than either artemisinin, dihydroartemisinin or artemether, which influences its diffusion across mucosal membranes as well as other pharmacokinetic properties [37]. For this reason, we decided to compare the in silico ADME and PBPK profiles of LDT-597 (Figure 5), Artesunate and DHA (Supplemental Figures S2 and S3, respectively) using the program Detoxie®which is an artificial intelligence de-risking application based on QSPR models for each endpoint and human PBPK model. While compound LDT-597 was predicted to be slightly more soluble than artesunate and a lot more soluble than DHA, its predicted intestinal permeability is $1.29 \text{ cm/s} \times 10^4$, while artesunate's predicted intestinal permeability is $2.29 \text{ cm/s} \times 10^4$. Moreover, the volume of distribution (VD) of compound LDT-597 was considerably higher than that of artesunate. As compound LDT-597 showed to have good physiologically based pharmacokinetic properties both in rats and humans, it was predicted to have a higher fraction unbound to protein plasma (UF) when compared to artesunate and DHA, which results in better compound bioavailability.

PERMEABILITY

Human Intestinal Absorption (HIA)	1.29 cm/s $\times 10^4$
MDCK Permeability	230.42 cm/s $\times 10^6$
Molecular Diffusion Coefficient	0.59 cm²/s $\times 10^5$
Human Skin	0.71 cm/s $\times 10^7$
Rabbit cornea	44.65 cm/s $\times 10^7$
CNS (Blood-brain barrier)	BBB(+)
Molar Volume (normal BP)	483 cm³/mol

ABSORPTION

Absorption risk	Size, Too many HBA, Too flexible (RotB), Excessive charge
logP (octanol-water, also known as Kow)	1.88
logD (octanol-water - Lipophilicity ionisable cpds)	1.88

SOLUBILITY

Aqueous	0.160 mg/mL
Aqueous solubility at pH 7.4	0.160 mg/mL
Aqueous intrinsic solubility	0.160 mg/mL
Aqueous solubility (pH at saturation)	7.29
Aqueous salts solubility factor	251.23
Supersaturated solution	Potential supersaturate (+)

SIMULATED GASTROINTESTINAL FLUIDS

Fasted state gastric fluid	0.14 mg/mL
Fasted state intestinal fluid	0.05 mg/mL
Fed state intestinal fluid	0.16 mg/mL

METABOLISM

Metabolism Risk	-
Clearance Classification System	Metabolism clearance

HUMAN LIVER MICROSOMES (HML)

Overall unbound Intrinsic Clearance	15.93 µL/min/mg

RAT LIVER MICROSOMES (RML)

Overall unbound Intrinsic Clearance	31.23 µL/min/mg

CYP3A4 (human liver microsomes)

Inhibition/Substrate	Potential substrate (+)
Michaelis-Menten constant (Km)	195.32 µM
Maximum metabolic rate (Vmax)	1.82 nmol/min/nmol
Intrinsic clearance (CLint)	9.21 µL/min/mg

CYP3A4

Inhibition/Substrate	Potential substrate (+)
Michaelis-Menten constant (Km)	7.92 µM
Maximum metabolic rate (Vmax)	21.43 nmol/min/nmol
Intrinsic clearance (CLint)	169.23 µL/min/mg

CYP2C9

Inhibition/Substrate	Non inhibitor/substrate (-)
Michaelis-Menten constant (Km)	-
Maximum metabolic rate (Vmax)	-
Intrinsic clearance (CLint)	-

CYP2D6

Inhibition/Substrate	Non inhibitor/substrate (-)
Michaelis-Menten constant (Km)	-
Maximum metabolic rate (Vmax)	-
Intrinsic clearance (CLint)	-

ALDEHYDE OXIDASE (AOX)	Non-substrate (-)
UDP-GLYCOSYLTRANSFERASE	Substrate (+) (UGT2B7)

CYP1A2

Inhibition/Substrate	Potential inhibitor (+)
Michaelis-Menten constant (Km)	-
Maximum metabolic rate (Vmax)	-
Intrinsic clearance (CLint)	-

RAT PBPK EXPLORER

Volume of distribution (Vd)	0.90 L/Kg
Fraction unbound to protein plasma (fup)	48.99 %
Ratio of blood to plasma (RBP)	0.84
Fraction absorbed (Fa%) @ 1 mg	96.14 %
Oral bioavaliability (F%) @ 1 mg	68.43 %
Max. plasma conc. (Cmax) @ 1 mg	682.23 ng/mL
Time to max. plasma conc. (Tmax) @ 1 mg	1.66 hr
Area under the curve (AUC) @ 1 mg	2086.23 ng-hr/mL

HUMAN PBPK EXPLORER

Volume of distribution (Vd)	1.18 L/Kg
Fraction unbound to protein plasma (fup)	46.23 %
Fraction unbound to microsomes (fumic)	0.81
Ratio of blood to plasma (RBP)	0.71
Fraction absorbed (Fa%) @ 10 mg	99.16 %
Oral bioavaliability (F%) @ 10 mg	69.2 %
Max. plasma conc. (Cmax) @ 10 mg	56.23 ng/mL
Time to max. plasma conc. (Tmax) @ 10 mg	2.45 hr
Area under the curve (AUC) @ 10 mg	319.32 ng-hr/mL

Figure 5. ADME and PBPK multiparametric prediction of LDT-597 using the Detoxie® software (http://insilicall.com/, accessed on 17 December 2020).

4. Discussion

Considering the relevance of natural products and derivatives for antimalarial drug development, we have applied a QSAR-based virtual screening pipeline for identification of potent antimalarial candidates against chloroquine-sensitive (3D7) and multi-drug-resistant (W2) *P. falciparum* strains. Firstly, the MolPort database for natural compounds and derivatives was screened with the aim of identifying molecules with predicted EC_{50} against *P. falciparum* 3D7 ≤ 1 µM and predicted probability of being active against *P. falciparum* W2 $> 60\%$. These molecules were also passed by a lipophilicity prediction filter and were verified for agreement within the applicability domain of QSAR models used in this study. It is important to note that we did not use a molecular weight filter due to large carbonic chains in natural products and derivatives, which could be a huge restriction and lead to loss of potential candidate molecules. The selected virtual hits were further divided into structurally diverse clusters and visually inspected for compound prioritization. Experimental in vitro assays against intraerythrocytic *P. falciparum* 3D7 and W2 strains were conducted to pinpoint active compounds.

Among the compounds evaluated, some of them do not show impressive phenotypic activity (high EC_{50} values and consequent low in vitro therapeutic indexes), which might be due to the limited size of natural products-based commercial databases compared to those of synthetic drugs. Even though the universe of natural products covers a more diverse chemical space, it is limited by being heavily dependent on the identification, isolation, and characterization of bioactive natural products from the biodiversity. However, our QSAR models were able to identify compounds with relevant activity in vitro, as highlighted by the nanomolar levels of parasite growth inhibition for compounds LDT-597

and LDT-598 against multi-drug-resistant *P. falciparum* W2 strains. These two compounds, which are artesunate derivatives, show activity about 10-fold more potent than artesunate itself (EC_{50} = 5.7 nM) against this same multi-resistant strain [38]. The identification of sesquiterpene lactones containing an endoperoxide bridge as the most promising antimalarial candidates within a diverse set of natural products and derivatives illustrates the agreement in the results obtained with the virtual screening pipeline, as this class of compounds has been widely characterized as potent antimalarial molecules [39–41]. By describing structural moieties that support antimalarial activity, HQSAR analysis performed by Avery and colleagues [42] showed that the lactone ring has a strong positive contribution for the antimalarial activity of artemisinin. Once this lactone ring is the main common substructure among artemisinin, artesunate and compounds LDT-597 and LDT-598, the robust potency of the latter two compounds can be explained by this shared feature that is responsible for antimalarial activity even against drug-resistant *P. falciparum* parasites.

Since the pace of drug resistance selection in *P. falciparum* parasites has been as fast as ever, there is an urgent need to ensure efficacy of compounds that stand out within the pipeline of drug discovery. For this reason, in silico models for pharmacokinetic antimalarial drugs have been developed, including those that have analyzed properties of artemisinin derivatives [43–45]. Altogether, the results highlight a better pharmacokinetics profile of LDT-597 when compared with artesunate, which might enhance efficacy of treatments using this compound instead of the traditional artemisinin derivatives.

5. Conclusions

Herein, we report a QSAR-based virtual screening study for predicting the antimalarial potential of a library of natural compounds and derivatives and explore the potential molecular target for the most potent compounds. Aiming to validate our in silico approach, we identified two compounds with low nanomolar inhibition levels against multi-drug-resistant *P. falciparum* strain in vitro. Moreover, in silico ADME/PBPK analysis of one of these compounds showed that it presents favorable physicochemical characteristics. The work presented endorses the applicability of natural compound databases for virtual screening campaigns and supports QSAR as a means of identifying potent and selective compounds against the malaria parasite.

References

1. World Health Organization. *World Malaria Report 2019*; World Health Organization: Geneva, Switzerland, 2019.
2. Rogerson, S.J.; Beeson, J.G.; Laman, M.; Poespoprodjo, J.R.; William, T.; Simpson, J.A.; Price, R.N.; The ACREME Investigators. Identifying and combating the impacts of COVID-19 on malaria. *BMC Med.* **2020**, *18*, 1–7. [CrossRef] [PubMed]
3. Sherrard-Smith, E.; Hogan, A.B.; Hamlet, A.; Watson, O.J.; Whittaker, C.; Winskill, P.; Ali, F.; Mohammad, A.B.; Uhomoibhi, P.; Maikore, I.; et al. The potential public health consequences of COVID-19 on malaria in Africa. *Nat. Med.* **2020**, *26*, 1411–1416. [CrossRef] [PubMed]
4. Haldar, K.; Bhattacharjee, S.; Safeukui, I. Drug resistance in Plasmodium. *Nat. Rev. Genet.* **2018**, *16*, 156–170. [CrossRef] [PubMed]
5. Noedl, H.; Se, Y.; Schaecher, K.; Smith, B.L.; Socheat, D.; Fukuda, M.M. Evidence of Artemisinin-Resistant Malaria in Western Cambodia. *N. Engl. J. Med.* **2008**, *359*, 2619–2620. [CrossRef] [PubMed]
6. Uwimana, A.; Legrand, E.; Stokes, B.H.; Ndikumana, J.-L.M.; Warsame, M.; Umulisa, N.; Ngamije, D.; Munyaneza, T.; Mazarati, J.-B.; Munguti, K.; et al. Emergence and clonal expansion of in vitro artemisinin-resistant Plasmodium falciparum kelch13 R561H mutant parasites in Rwanda. *Nat. Med.* **2020**, *26*, 1602–1608. [CrossRef] [PubMed]
7. Achan, J.; O Talisuna, A.; Erhart, A.; Yeka, A.; Tibenderana, J.K.; Baliraine, F.N.; Rosenthal, P.J.; D'Alessandro, U. Quinine, an old anti-malarial drug in a modern world: Role in the treatment of malaria. *Malar. J.* **2011**, *10*, 144. [CrossRef] [PubMed]
8. Bosman, A.; Mendis, K.N. A Major Transition in Malaria Treatment: The Adoption and Deployment of Artemisinin-Based Combination Therapies. *Am. J. Trop. Med. Hyg.* **2007**, *77*, 193–197. [CrossRef] [PubMed]
9. Chatterjee, A.K. Back to the Future: Lessons Learned in Modern Target-based and Whole-Cell Lead Optimization of Antimalarials. *Curr. Top. Med. Chem.* **2012**, *12*, 473–483. [CrossRef] [PubMed]
10. Kim, S.; Thiessen, P.A.; Bolton, E.E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B.A.; et al. PubChem substance and compound databases. *Nucleic Acids Res.* **2016**, *44*, D1202–D1213. [CrossRef]
11. Bento, A.P.; Gaulton, A.; Hersey, A.; Bellis, L.J.; Chambers, J.; Davies, M.; Krüger, F.A.; Light, Y.; Mak, L.; McGlinchey, S.; et al. The ChEMBL bioactivity database: An update. *Nucleic Acids Res.* **2014**, *42*, D1083–D1090. [CrossRef]
12. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57*, 4977–5010. [CrossRef]
13. Jorgensen, W.L. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303*, 1813–1818. [CrossRef]
14. Nantasenamat, C.; Isarankura-Na-Ayudhya, C.; Prachayasittikul, V. Advances in computational methods to predict the biological activity of compounds. *Expert Opin. Drug Discov.* **2010**, *5*, 633–654. [CrossRef]
15. Verma, J.; Khedkar, V.M.; Coutinho, E.C. 3D-QSAR in Drug Design—A Review. *Curr. Top. Med. Chem.* **2010**, *10*, 95–115. [CrossRef]
16. Zhang, L.; Fourches, D.; Sedykh, A.; Zhu, H.; Golbraikh, A.; Ekins, S.; Clark, J.; Connelly, M.C.; Sigal, M.; Hodges, D.; et al. Discovery of Novel Antimalarial Compounds Enabled by QSAR-Based Virtual Screening. *J. Chem. Inf. Model.* **2013**, *53*, 475–492. [CrossRef]
17. Neves, B.J.; Dantas, R.F.; Senger, M.R.; Melo-Filho, C.C.; Valente, W.C.G.; De Almeida, A.C.M.; Rezende-Neto, J.M.; Lima, E.F.C.; Paveley, R.; Furnham, N.; et al. Discovery of New Anti-Schistosomal Hits by Integration of QSAR-Based Virtual Screening and High Content Screening. *J. Med. Chem.* **2016**, *59*, 7075–7088. [CrossRef]
18. Lima, M.N.N.; Melo-Filho, C.C.; Cassiano, G.C.; Neves, B.J.; Alves, V.M.; Braga, R.C.; Cravo, P.V.L.; Muratov, E.N.; Calit, J.; Bargieri, D.Y.; et al. QSAR-Driven Design and Discovery of Novel Compounds With Antiplasmodial and Transmission Blocking Activities. *Front. Pharmacol.* **2018**, *9*, 146. [CrossRef]
19. Gomes, M.N.; Alcântara, L.M.; Neves, B.J.; Melo-Filho, C.C.; Freitas-Junior, L.H.; Moraes, C.B.; Ma, R.; Franzblau, S.G.; Muratov, E.N.; Andrade, C.H. Computer-aided discovery of two novel chalcone-like compounds active and selective against Leishmania infantum. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 2459–2464. [CrossRef]
20. Melo-Filho, C.C.; Dantas, R.F.; Braga, R.C.; Neves, B.J.; Senger, M.R.; Valente, W.C.G.; Rezende-Neto, J.M.; Chaves, W.T.; Muratov, E.N.; Paveley, R.A.; et al. QSAR-Driven Discovery of Novel Chemical Scaffolds Active against Schistosoma mansoni. *J. Chem. Inf. Model.* **2016**, *56*, 1357–1372. [CrossRef]
21. Xie, N.; Wang, C.; Lian, Y.; Wu, C.; Zhang, H.; Zhang, Q. Inhibition of mitochondrial fission attenuates Aβ-induced microglia apoptosis. *Neuroscience* **2014**, *256*, 36–42. [CrossRef]
22. Neves, B.; Moreira-Filho, J.; Silva, A.; Borba, J.; Mottin, M.; Alves, V.; Braga, R.; Muratov, E.; Andrade, C. Automated Framework for Developing Predictive Machine Learning Models for Data-Driven Drug Discovery. *J. Braz. Chem. Soc.* **2021**, 1–13. [CrossRef]
23. Lima, M.N.N.; Cassiano, G.C.; Tomaz, K.C.P.; Silva, A.C.; Sousa, B.K.P.; Ferreira, L.T.; Tavella, T.A.; Calit, J.; Bargieri, D.Y.; Neves, B.J.; et al. Integrative Multi-Kinase Approach for the Identification of Potent Antiplasmodial Hits. *Front. Chem.* **2019**, *7*, 773. [CrossRef] [PubMed]
24. Fourches, D.; Muratov, E.; Tropsha, A. Trust, But Verify: On the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research. *J. Chem. Inf. Model.* **2010**, *50*, 1189–1204. [CrossRef] [PubMed]

25. Fourches, D.; Muratov, E.; Tropsha, A. Trust, but Verify II: A Practical Guide to Chemogenomics Data Curation. *J. Chem. Inf. Model.* **2016**, *56*, 1243–1252. [CrossRef] [PubMed]

26. Fourches, D.; Muratov, E.; Tropsha, A. Curation of chemogenomics data. *Nat. Chem. Biol.* **2015**, *11*, 535. [CrossRef] [PubMed]

27. Wang, R.; Fu, A.Y.; Lai, L. A New Atom-Additive Method for Calculating Partition Coefficients. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 615–621. [CrossRef]

28. Butina, D. Unsupervised Data Base Clustering Based on Daylight's Fingerprint and Tanimoto Similarity: A Fast and Automated Way to Cluster Small and Large Data Sets. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 747–750. [CrossRef]

29. Sydow, D.; Morger, A.; Driller, M.; Volkamer, A. TeachOpenCADD: A teaching platform for computer-aided drug design using open source packages and data. *J. Cheminform.* **2019**, *11*, 29. [CrossRef]

30. Sander, T.; Freyss, J.; Von Korff, M.; Rufener, C. DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis. *J. Chem. Inf. Model.* **2015**, *55*, 460–473. [CrossRef]

31. Trager, J.J.W. Human malaria parasites in continuous culture. *Science* **1976**, *193*, 673–675. [CrossRef]

32. Radfar, A.; Méndez, D.; Moneriz, C.; Linares, M.; Marín-García, P.; Puyet, A.; Diez, A.; Bautista, J.M. Synchronous culture of Plasmodium falciparum at high parasitemia levels. *Nat. Protoc.* **2009**, *4*, 1899–1915. [CrossRef]

33. Hartwig, C.L.; Ahmed, A.O.A.; Cooper, R.A.; Stedman, T.T. SYBR Green I®-Based Parasite Growth Inhibition Assay for Measurement of Antimalarial Drug Susceptibility in Plasmodium Falciparum. In *Methods in Malaria Research*; Moll, K., Kaneko, A., Scherf, A., Wahlgren, M., Eds.; EVIMalaR: Manassas, Virginia, 2013; pp. 122–129.

34. Mosmann, T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *J. Immunol. Methods* **1983**, *65*, 55–63. [CrossRef]

35. Djoumbou-Feunang, Y.; Fiamoncini, J.; Gil-De-La-Fuente, A.; Greiner, R.; Manach, C.; Wishart, D.S. BioTransformer: A comprehensive computational tool for small molecule metabolism prediction and metabolite identification. *J. Cheminform.* **2019**, *11*, 1–25. [CrossRef]

36. Cui, L.; Su, X.-Z. Discovery, mechanisms of action and combination therapy of artemisinin. *Expert Rev. Anti-Infect. Ther.* **2009**, *7*, 999–1013. [CrossRef]

37. Karunajeewa, H.A. Artemisinins: Artemisinin, Dihydroartemisinin, Artemether and Artesunate. In *Treatment and Prevention of Malaria*; Metzler, J.B., Ed.; Springer Basel: Basel, Switzerland, 2011; pp. 157–190.

38. Smit, F.J.; van Biljon, R.A.; Birkholtz, L.-M.; N'Da, D.D. Synthesis and in vitro biological evaluation of dihydroartemisinyl-chalcone esters. *Eur. J. Med. Chem.* **2015**, *90*, 33–44. [CrossRef]

39. Pedersen, M.M.; Chukwujekwu, J.C.; Lategan, C.A.; Van Staden, J.; Smith, P.J.; Staerk, D. Antimalarial sesquiterpene lactones from Distephanus angulifolius. *Phytochemistry* **2009**, *70*, 601–607. [CrossRef]

40. Francois, G.; Passreiter, C.M. Pseudoguaianolide sesquiterpene lactones with high activities against the human malaria parasitePlasmodium falciparum. *Phytother. Res.* **2004**, *18*, 184–186. [CrossRef]

41. Bordignon, A.; Frédérich, M.; LeDoux, A.; Campos, P.-E.; Clerc, P.; Hermann, T.; Quetin-Leclercq, J.; Cieckiewicz, E. In vitro antiplasmodial and cytotoxic activities of sesquiterpene lactones from Vernonia fimbrillifera Less. (Asteraceae). *Nat. Prod. Res.* **2017**, *32*, 1463–1466. [CrossRef]

42. Avery, M.A.; Alvim-Gaston, M.; Rodrigues, C.R.; Barreiro, E.J.; Cohen, F.E.; Sabnis, Y.A.; Woolfrey, J.R. Structure−Activity Relationships of the Antimalarial Agent Artemisinin. The Development of Predictive In Vitro Potency Models Using CoMFA and HQSAR Methodologies. *J. Med. Chem.* **2002**, *45*, 292–303. [CrossRef]

43. Arey, R.; Reisfeld, B. Predicting the Disposition of the Antimalarial Drug Artesunate and Its Active Metabolite Dihydroartemisinin Using Physiologically Based Pharmacokinetic Modeling. *Antimicrob. Agents Chemother.* **2020**, *65*. [CrossRef]

44. Gordi, T.; Xie, R.; Huong, N.V.; Huong, D.X.; Karlsson, M.O.; Ashton, M. A semiphysiological pharmacokinetic model for artemisinin in healthy subjects incorporating autoinduction of metabolism and saturable first-pass hepatic extraction. *Br. J. Clin. Pharmacol.* **2005**, *59*, 189–198. [CrossRef]

45. Olafuyi, O.; Coleman, M.; Badhan, R.K. Development of a paediatric physiologically based pharmacokinetic model to assess the impact of drug-drug interactions in tuberculosis co-infected malaria subjects: A case study with artemether-lumefantrine and the CYP3A4-inducer rifampicin. *Eur. J. Pharm. Sci.* **2017**, *106*, 20–33. [CrossRef]