Scientific Research Publishing

Computational Chemistry, 2023, 11, 1-23
https://www.scirp.org/journal/cc
ISSN Online: 2332-5984
ISSN Print: 2332-5968

Designing Artemisinins with Antimalarial Potential, Combining Molecular Electrostatic Potential, Ligand-Heme Interaction and Multivariate Models

Josué de Jesus Oliveira Araújo[1], Ricardo Morais de Miranda[2], Jeferson Stiver Oliveira de Castro[2], Antonio Florêncio de Figueiredo[2], Ana Cecília Barbosa Pinheiro[3], Sílvia Simone dos Santos Morais[4], Marcos Antonio Barros dos Santos[5], Andréia de Lourdes Ribeiro Pinheiro[6], Fábio dos Santos Gil[1], Heriberto Rodrigues Bitencourt[7], Gustavo Nery Ramos Alves[1], José Ciríaco Pinheiro[1*]

[1]Laboratório de Química Teórica e Computacional, Universidade Federal do Pará, Belém, PA, Brasil
[2]Instituto Federal de Educação, Ciência e Tecnologia do Pará, Belém, PA, Brasil
[3]Universidade do Estado do Amapá, Macapá, AP, Brasil
[4]Fundação Santa Casa de Misericórdia do Pará, Belém, PA, Brasil
[5]Universidade do Estado do Pará, Barcarena, PA, Brasil
[6]Departamento de Biologia, Universidade do Estado do Maranhão, São Luís, MA, Brasil
[7]Laboratório de Síntese Orgânica, Universidade Federal do Pará, Belém, PA, Brasil
Email: *ciríaco@ufpa.br

How to cite this paper: de Jesus Oliveira Araújo, J., de Miranda, R.M., de Castro, J.S.O., de Figueiredo, A.F., Pinheiro, A.C.B., dos Santos Morais, S.S., dos Santos, M.A.B., de Lourdes Ribeiro Pinheiro, A., dos Santos Gil, F., Bitencourt, H.R., Alves, G.N.R. and Pinheiro, J.C. (2023) Designing Artemisinins with Antimalarial Potential, Combining Molecular Electrostatic Potential, Ligand-Heme Interaction and Multivariate Models. *Computational Chemistry*, **11**, 1-23.
https://doi.org/10.4236/cc.2023.111001

Received: September 7, 2022
Accepted: November 28, 2022
Published: December 1, 2022

Abstract

Artemisinins tested against W-2 strains of malaria falciparum are investigated with molecular electrostatic potential (MEP), in an attempt to identify key features of the compounds that are necessary for their activities, as well as to investigate likely interactions with the receptor in a biological process and to use that information to propose new molecules. In order to discover the best geometry involving the ligand-receptor complexes (heme) studied and help in the proposition of the new derivatives, molecular simulations of interactions between the most negative charged region around the peroxide and heme locates (the ones around the Fe^{2+} ion) were carried out. In addition, PCA (principal components analysis), HCA (hierarchical cluster analysis), SDA (stepwise discriminant analysis), and KNN (K-nearest neighbor) multivariate models were employed to investigate which descriptors are responsible for the classification between the higher and lower antimalarial activity of the compounds, and also this information was used to propose new potentially active molecules. The information accumulated in studies of MEP, molecular docking, and multivariate analysis supported the proposal of new structures with potential antimalarial activities. The multivariate models constructed

were applied to the new structures and indicated numbers **19** and **20** as the most prominent for syntheses and biological assays.

Keywords

Artemisinins, Antimalarial Potential, Molecular Electrostatic Potential, Ligand-Heme Interaction, Multivariate Models

1. Introduction

Malaria is still one of the main public health problems. In 2020, the World Health Organization (WHO) estimated the occurrence of 241 million cases of malaria around the world [1], resulting in 627,000 deaths, with the most of them caused by infection with *P. falciparum* [2].

Currently, WHO recommends 14 drugs for the curative treatment of malaria and 4 drugs for the prophylactic treatment. Among these treatments, the most effective are artemisinin-based combinations that use an artemisinin derivative (short-acting) in combination with one or more complementary compounds (long-acting and having different mechanisms of action) [2].

Artemisinin was first isolated in 1971 from the plant Artemisia *annua* L., a herb that has been commonly used in traditional Chinese medicine [2] [3] [4] [5]. Artemisinin has shown to be effective against all multi-resistant forms of *P. falciparum*. The most common artemisinin derivatives are artemether, artesunate and arteter. These semi-synthetic pro-drug derivatives are transformed into the active metabolite, dihydroartemisinin. The use of artemisinins has been an integral part of the fight against malaria in most modern treatments [2] [5].

The mechanism of action by which artemisinin acts has been widely debated [2] [6] [7] [8]. The most accepted theory is that the molecule is activated by heme to generate free radicals which, in turn, damage the proteins necessary for the parasite's survival [2] [8] [9] [10] [11] [12]. Still, evidence for a range of other possible mechanisms has been found: in 2013, a computational approach was used to determine the mechanism of action based on previous studies that identified heme and PfATP6 (Ca^{2+} transporter) as potential mechanisms of action [2] [8] [13]. More recently, in 2015, it was shown that artemisinin is associated with regulation of the ascending pathway of protein response, which may be linked to decreased parasite development [2] [8] [14]. Another study showed that artemisinin is a potent inhibitor of phosphatidylinositol-3-kinase (PfPI3K) [2] [8] [15].

Bernardinelli *et al.* stated that active artemisinin and related molecules have similar MEP around the essential trioxane-ring, and this property is due to the peroxide linkage [16]. Here, the HF/6-31G** theory [17] [18] was used to build MEP maps to sixteen (16) artemisinins with different activities degree against W-2 strains of *P. falciparum* from Indochina [19] [20].

The MEP maps were then evaluated and used in an attempt to identify key features of the molecules that are necessary for their activities, to investigate

likely interactions with the receptor in a biological process, and to use this information to propose new promising molecules. In addition, in order to discover the best geometry involving the ligand-heme complex studied, and help in the proposition of the new derivatives, molecular simulations of interactions between the most negative MEP region around the peroxide of the ligand with heme locates around the Fe^{2+} ion were carried out. Mutivariate analysis (MA) was used and PCA, HCA, SDA, and KNN models were built with the artemisisnins being classified into greater antimalarial potential and lower antimalarial potential in terms of activity.

The information accumulated in studies of MEP, interaction with biological target, multivariate analysis and chemical intuition supported the proposal of new artemisinins with antimalarial potential. The multivariate models constructed were applied to the new artemisinins and indicated 19 and 20 as those with the most promising antimalarial potential.

2. Computational Methodology

2.1. Structures Studied, Calculations and Molecular Properties

Figure 1 shows the molecular structures of the artemisinins used in the present study. In structure 1 (artemisinin) of the figure, it is shown the atomic numbering adopted to study the compounds. The artemisinins were tested *in vitro* against W-2 strains of *P. falciparum* from Indochina that are resistant to the antimalarial drug mefloquine [19] [20]. The compounds were defined as having greater antimalarial potential (GAP) against the *P. falciparum* if logarithms of the relative activity (log AR = IC_{50} artemisinin/IC_{50} artemisinins X molecular mass of artemisinins/molecular mass of artemisinin) were equal or above 0 (zero) and as having lower antimalarial potential (LAP) if the logs AR were lower than this value.

The molecular calculations have, as starting point, the construction of the structure of artemisinin **1** using Gauss View software [21], and complete geometry optimization with the HF/6-31G** method/basis set for the artemisinin 1,2,4-trioxane-ring [17] [18] was done. The molecular parameters of the 1,2,4-trioxane ring were computed and compared to calculated ones in different basis sets [22] [23] and with experimental values from literature [24] [25], so as to check the reliability of the geometry obtained. All calculations reproduced most the geometrical parameters of the artemisinin 1,2,4-trioxane-ring (-$O_1O_2C_3O_{13}C_{12}C_{12a}$-) seen in the X-ray data in **Table 1**. This applies especially to the bond length of the endoperoxide bridge (-O_1O_2-) which, as mentioned before, seems to be responsible for the antimalarial activity [16]. All the other structures (**Figure 1**) were built with the optimized geometry of the artemisinin, by also using the GaussView program. The calculations were carried out by using the Gaussian 98 program and the DIRECT-SCF approach [26].

To give us valuable information about the influence of electronic, steric, hydrophilic and hydrophobic features on the interactions between artemisinins

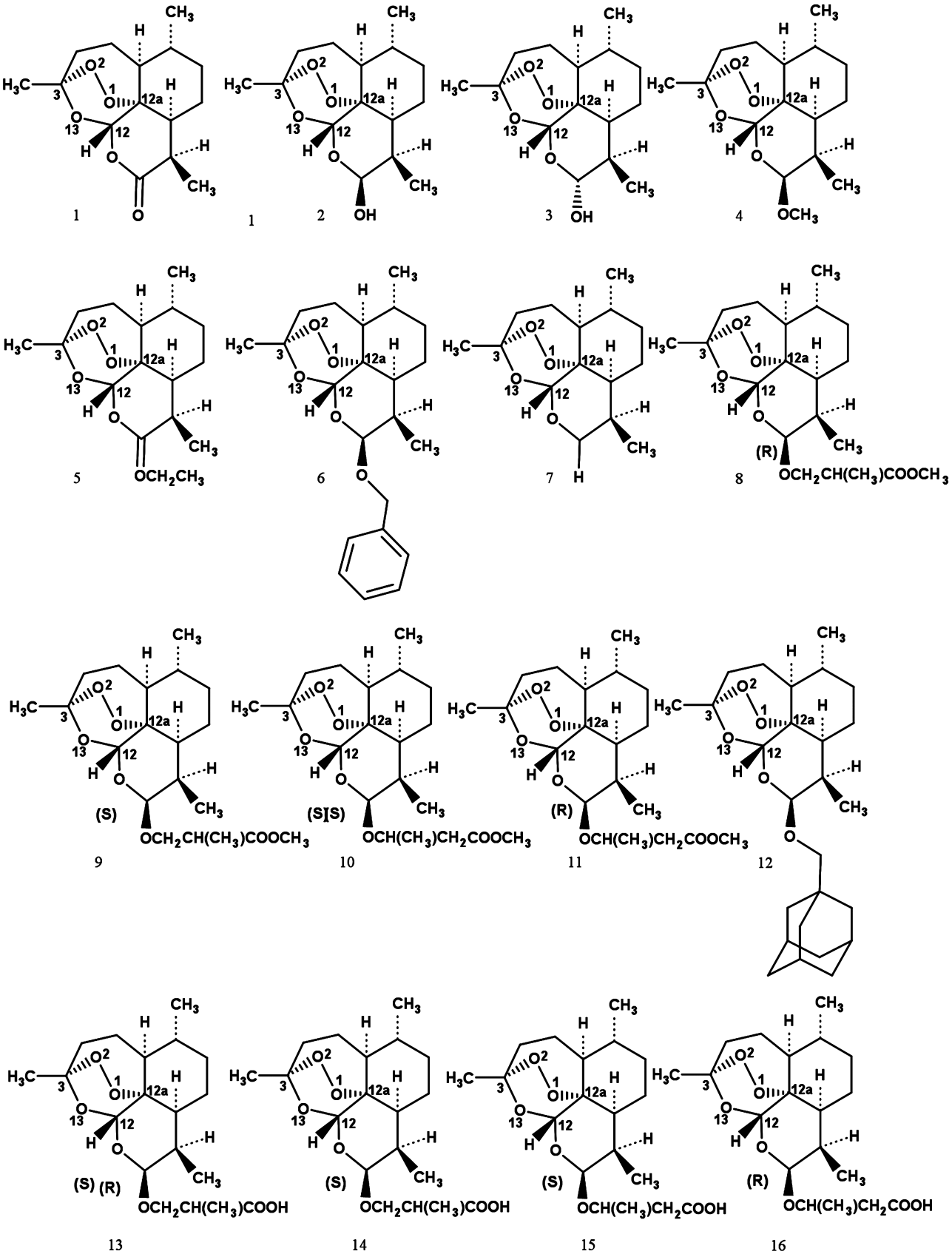

Figure 1. 2D structures of artemisinins with antimalarial activity against W-2 strains of *P. falciparum.*

Table 1. Experimental and calculated geometrical parameters of the 1,2,4-trionane ring for artemisinin (compound 1).

Parameters	Theory level				Experimental[c]	Experimental[d]
	HF/3-21G[a]	HF/6-21G[a]	HF/6-31G*[b]	HF/6-31G** (this work)		
Bond lengths (Å)						
O_1O_2	1.462	1.447	1.390	1.390	1.474 (4)	1.4.69 (2)
O_2C_3	1.441	1.435	1.396	1.396	1.418 (4)	1.416 (3)
C_3O_{13}	1.436	1.435	1.408	1.409	1.451 (4)	1.445 (3)
$O_{13}C_{12}$	1.407	1.403	1.376	1.376	1.388 (4)	1.380 (3)
$C_{12}C_{12a}$	1.529	1.533	1.532	1.532	1.528 (5)	1.523 (2)
$C_{12a}O_1$	1.477	1.469	1.429	1.429	1.450 (4)	1.462 (3)
Bond angles (degree)						
$O_1O_2C_3$	107.1	108.8	109.5	109.5	107.7 (2)	108.1 (2)
$O_2C_3O_{13}$	107.3	106.8	107.8	107.8	107.0 (2)	106.6 (2)
$C_3O_{13}C_{12}$	115.7	117.3	115.3	115.3	113.6 (3)	114.2 (3)
$O_{13}C_{12}C_{12a}$	112.1	112.3	112.3	112.3	114.7 (2)	114.5 (2)
$C_{12}C_{12a}O_1$	111.6	110.9	110.5	110.5	111.1 (2)	110.7 (2)
$C_{12a}O_1O_2$	111.3	113.2	112.7	112.7	111.5 (2)	111.1 (2)
Torsion angles (degree)						
$O_1O_2C_3O_{13}$	−74.7	−71.8	−73.4	−73.4	−75.5 (3)	−75.5 (2)
$O_2C_3O_{13}C_{12}$	32.4	33.4	31.1	31.1	36.3 (4)	36.0 (2)
$C_3O_{13}C_{12}C_{12}$a	28.3	25.3	27.4	27.4	24.7 (4)	25.3 (2)
$O_{13}C_{12}C_{12}$a O_1	−50.7	−49.4	−50.1	−50.2	−50.8 (4)	−51.3 (2)
$C_{12}C_{12}aO_1O_2$	9.9	12.5	10.9	10.9	12.2 (3)	12.6 (2)
$C_{12}aO_1O_2C_3$	50.4	46.7	48.7	48.7	47.7 (3)	47.8 (2)

[a]Values from Ref [22]. [b]Values from Ref [23]. [c,d]Values from Ref. [24] and [25], respectively.

and heme, quantum, holistic and physicochemical molecular descriptors were obtained with the aid of the Gaussian 98 [26], Dragon [27] and HyperChem [28] programs, respectively.

The quantum descriptors were the highest occupied molecular orbital (HOMO) energy, one level below HOMO orbital (HOMO-1) energy, the lowest unoccupied molecular orbital (LUMO) energy, one level above LUMO orbital (LUMO+1) energy, the HOMO-LUMO gap, Mülliken's electronegativity (χ), molecular hardness, molecular softness, dipole moment, atomic charges on the Nth atom (Q_N), bond lengths, bond angles and torsion angles of the artemisinins 1,2,4-trioxane-ring. The atomic charges were generated by the CHELPEG keyword through electrostatic potential [29] [30]. The charges derived from elec-

trostatic potential are, in general, physically more satisfactory than the Mülliken charges [31], especially regarding biological activity. The holistic descriptors were selected with of the purpose of representing different sources of chemical information in terms of molecular size and shape, symmetry and distribution in the molecule. The physicochemical descriptors were total surface area, molecular volume, molecular refractivity, and hydration energy.

2.2. Biological Recognition Process Ligand/Receptor through the Molecular Electrostatic Potential

The electrostatic potential, $V(\vec{r})$, created by the nuclei and electrons of a molecule at each point $\vec{r}$ in the surrounding space, is given by the Equation (1),

$$V(\vec{r}) = \sum_{i}^{N} \frac{Z_i}{\left|\overrightarrow{R_i} - \vec{r}\right|} - \int \frac{\rho(\vec{r}')\mathrm{d}\vec{r}'}{\left|\vec{r}' - \vec{r}\right|} \quad (1)$$

In Equation (1), Z_i is the charge on nucleus i, located at R_i, and $\rho(\vec{r}')$ is the electronic density of the molecule with the sign of $V(\vec{r})$ in a given region, depending on the positive contributions of the nuclei and negativity of the electrons [32] [33].

The electrostatic potential chemical-quantum approach has been an important tool to analyze biological recognition processes of one molecule by another, such as drug-receptor and enzyme-substrate interactions, because it is through their potentials that one species sees another, in a process of biological recognition [32].

The MEP was computed from the electronic density and the MEP maps were displayed by using the MOLEKEL software [34].

2.3. Interaction between Artemisinins and Heme

Another approach that has played an important role in the design of molecules of pharmacological interest, dealing with the interaction between ligand-receptor, is known as "molecular docking" [35] [36] [37] [38] [39]. Independently or in combination with other theoretical approaches, this methodology has been used efficiently to rationally design molecules of interest in the pharmaceutical industry [35] [36] [37] [38] [39].

The interaction between ligands (artemisinins) and heme was studied with aid of the molecular docking in order to discover the best geometry involving the complexes formed by these molecules. Artemisinins had their geometry build with HF/6-31G** theory, whereas for heme the geometry was obtained from the Protein Data Bank (PDB) RCSB, identified by the code 1A6M, according to the manuscript of Vojtechovsky *et al.* [40]. Besides that, for a more real description of the biological environment where this interaction takes place, a histidine residue was added to the heme unity. This histidine unity is, as usual, coordinated perpendicularly to the Fe^{2+} ion through its sp2 nitrogen atom of its imidazole ring. Such an arrangement will allow the Fe^{2+} ion to attain a nearly octahedral hexacoordinated arrangement after binding to the artemisinin molecule [41].

Then, for each artemisinins/heme interaction, 100 conformations were calculated, and the most probable, the one based on the lowest energies of interaction, was chosen to evaluate the distance between the O_1 atom from artemisinins and the iron atom from heme and the distance between the O_2 atom from artemisinins and the iron atom from heme. Automated dockings calculations were carried out to develop possible conformations for the complex by Lamarckian Genetic Algorithm as implemented in the Autodock 4.2 software [42] [43].

2.4. Multivariate Analysis

MA is especially useful for classifying objects into discrete classes based on measured characteristics. A set of characteristics of an object is considered an abstract pattern that contains information about a non-measured property of the object [44] [45].

Several researches carried out in this laboratory were conducted using the methods of MA mentioned in this section. Brief descriptions of these methods can be found in the articles [45] [46] [47] [48]. In this work, all MA methods were performed with the Pirouette program [49].

3. Results and Discussion

3.1. MEP Maps for Studied Artemisinis

It is very well-documented [16] that when artemisinin is compared with its derivatives, the MEP maps of the active compounds are similar in form to the artemisinin in the round of the 1,2,4-trioxane-ring of **Figure 1**. All the molecules in **Figure 1** are active and this evidence is supported by the experimentally determined data (IC_{50}) shown in **Table 2**.

Table 2. Molecular descriptors of the training set artemisinins that classify the 16 compounds studied, experimental IC_{50} and relative activity (RA), and correlation matrix, LUMO+1 energy, HE (hidratation energy), MAXDN (electrotopological negative variation) and ALOGPS_logs (Squared Ghose-Crippen octanol-water partition coeff.(log P^2).

Compound	LUMO + 1 energy[a]	HE[a]	MAXDN	ALOGPS_logs	IC_{50} [b,c,d]	RA[e]	Activity[f]
1+ (artemisinin)	131.8	−2.82	2.043	−2.34	2.60[c]	1	GAP
2+	139.3	−3.69	2.115	−1.93	0.69[c]	3.77	GAP
3+	140.6	−5.67	2.115	−1.93	0.69[c]	3.77	GAP
4+ (arthemeter)	160.0	−0.89	1.948	−2.80	1.3542[d]	1.9199	GAP
5+ (arteether)	141.2	−0.89	1.950	−3.04	1.2264[d]	2.1200	GAP
6+	140.6	−5.25	2.115	−4.54	1.38[c]	1.88	GAP
7+	136.8	−2.10	1.906	−2.99	0.43[c]	6.05	GAP
8+	137.4	−1.70	2.040	−3.90	0.057[c]	45.6	GAP
9+	134.9	−2.20	2.040	−3.90	0.015[c]	173.3	GAP

Continued

10+	131.1	−2.11	2.041	−3.90	0.48[c]	5.42	GAP
11+	134.9	−3.00	2.041	−3.9	0.071[c]	36.6	GAP
12+	138.7	−0.14	1.953	−4.86	2.11[c]	2.23	GAP
13-	131.1	−5.59	2.537	−3.09	5.32[c]	0.49	LAP
14-	131.1	−5.84	2.537	−3.09	8.33[c]	0.31	LAP
15-	128.0	−3.67	2.544	−3.14	7.43[c]	0.35	LAP
16-	134.3	−8.29	2.544	−3.14	4.42[c]	0.59	LAP
LUMO + 1 energy		0.218	−0.087	0.213			
HE			−0.427	−0.278			
MAXDN				0.263			

[a]kcal.mol^{-1}. [b]ng/mL. [c]From Ref [19]. [d]From Ref [20]. [e]RA = IC_{50} of artemisisinin (1)/IC_{50} of artemisinins. [f]Activity: GAP (greater antimalarial potential) and LAP (lower antimalarial potential).

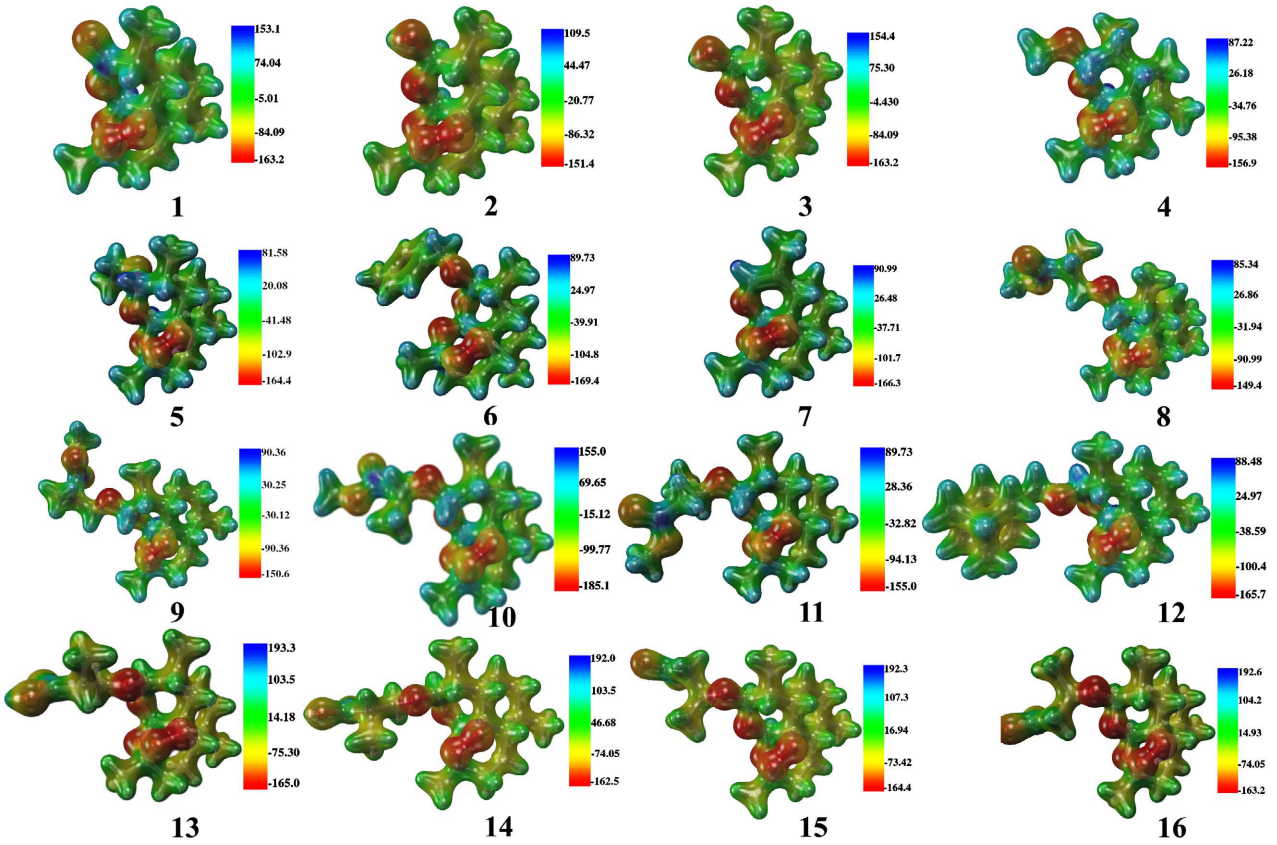

Figure 2. MEP (kcal.mol^{-1}) maps of artemisinins with antimalarial activity against W-2 strains of *P. falciparum.*

In Figure 2, the analysis of the MEP maps for the studied artemisinins shows similar contour surfaces close to the 1,2,4-trioxane-ring, characterized by negative electrostatic potential (red color), on which the lowest value for charge is localized around −166.3 kcal·mol^{-1}, indicating a concentration in electron density due to lone electron pairs on oxygen atoms (O_1, O_2, and O_{13}) susceptible to

electrophilic attack by a biological target.

These molecules also have contour surfaces characterized by positive electrostatic potential (blue color), whose the highest value is about +193.3 kcal·mol^{-1}. The distribution of electron density on the artemisinins around the trioxane-ring induces them to display activity against malaria, a belief supported by the fact that the complexation of artemisinin with heme involves particularly the interaction between peroxide bond, the most negative charged zone on the ligand, and Fe^{2+} ion, the most positive charged zone on heme (the receptor molecule) [16] [50] [51].

3.2. Interaction between Studied Artemisinins and Heme

The interaction between the studied artemisinins and heme (molecular docking) shows the ligand molecules placed in parallel to the plane of the heme porphyrin ring with the polar region of artemisinins containing the peroxide bond directed to the polar region of the heme system containing Fe^{2+} ion. **Figure 3(a)** shows the structure of the heme and **Figure 3(b)** shows the molecular docking for complexation between artemisinins and heme, where one can visualize the interactions for the molecules. Those orientations were assumed as the most favorable and then represent the real system under investigation, considering they were chosen based on the lowest binding free-energy (interaction energy). For the studied artemisinins, the FeO_1 distances are situated between 2.279 and 2.718 Å. On the other hand, the interval for the FeO_2 distances is from 3.078 to 3.744 Å. For artemisinin (**1**), the calculated value for the FeO_1 distance is 2.701 Å, reproducing the value (2.7 Å) by other authors reported in the literature [51] [52]. This result reinforces the perspective that the O_1 atom from artemisinins binds to the Fe^{2+} ion from heme more preferably than the O_2 atom. Such a preference for attack may be due to a greater steric hindrance at O_2 than at O_1 and a more negative charge on the latter atom; both factors are essential for intermolecular approach. A similar conclusion is supported by other studies [53] [54] that investigated the complex artemisinin/heme.

3.3. Multivariate Analysis

To performing the exploratory data analysis, all variables were auto-scaled as preprocessing so that they could be standardized and, this way, they have the same importance regarding the scale. Furthermore, given a large quality of multivariate data available, it was necessary to reduce the number of variables. Thus, if two any descriptors had a high Pearson Correlation Coefficient ($r > 0.8$), one of the two was randomly excluded from the matrix, since theoretically they describe the same property [55], that is, they also have a high correlation with antimalarial activity and only one of them is enough to be used as independent variable in a predictive model. Moreover, those properties that showed either the same values for most of the samples or a small correlation with activity ($r < 0.2$) were eliminated too.

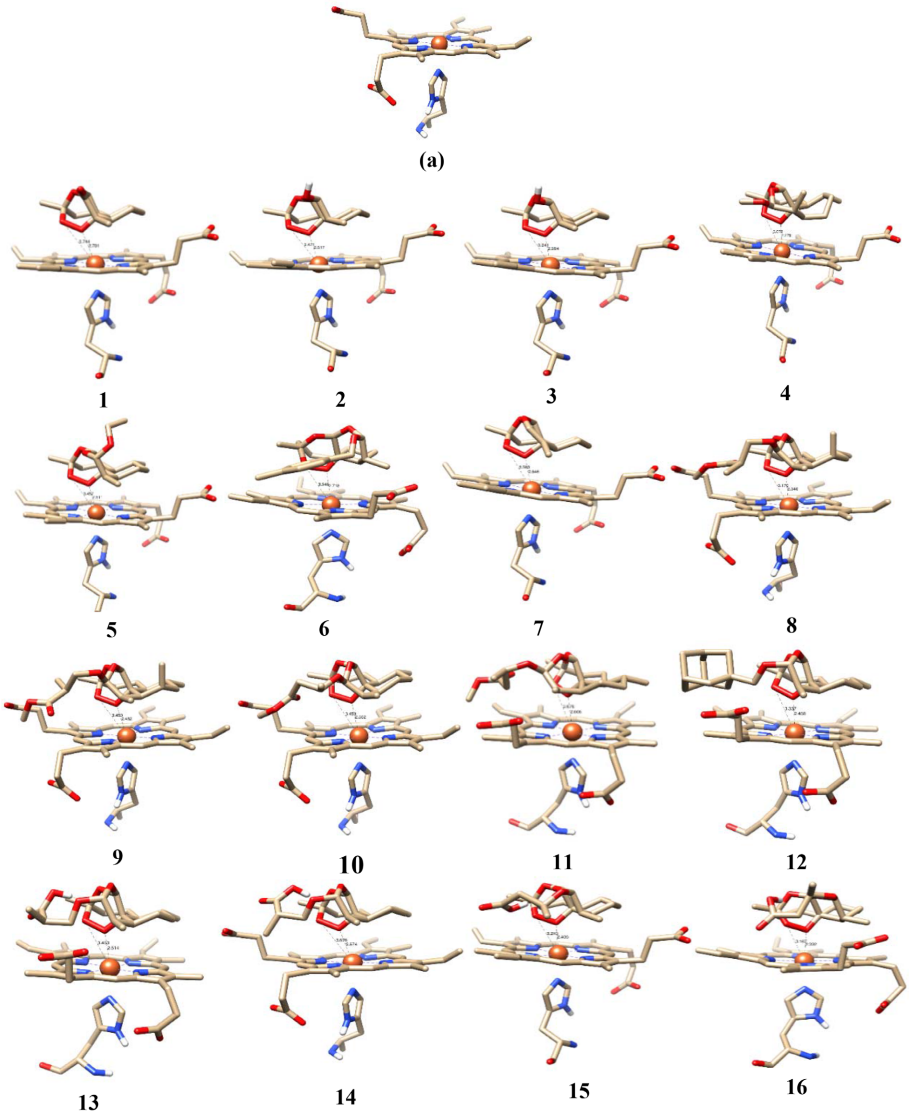

Figure 3. The 2D structure of heme (a) and the final docking configurations of artemisisnins-heme complex (b).

3.3.1. PCA and HCA Multivariate Models

After the matrix data compression, PC analysis was employed so as to reduce the dimensionality of the datum, and find descriptors that could be useful in characterizing the outlier samples. While processing PC analysis, several attempts to obtain a good classification of the compounds were done. At each attempt, the score and loading plots were analyzed based on the variables employed in the analysis. The score plot gives information about the compounds (similarity and differences). The loading plot gives information about the variables (how they are connected to each other and which are the best to describe the variance in original data).

The first three principal components (PC1, PC2, and PC3) explained 96.85% of the total variance in the data distributed as: PC1 = 55.18, PC2 = 26.71%, and PC3 = 14.96%. **Figure 4(a)** shows PC1-PC2 scores plot for the samples (**1-16**),

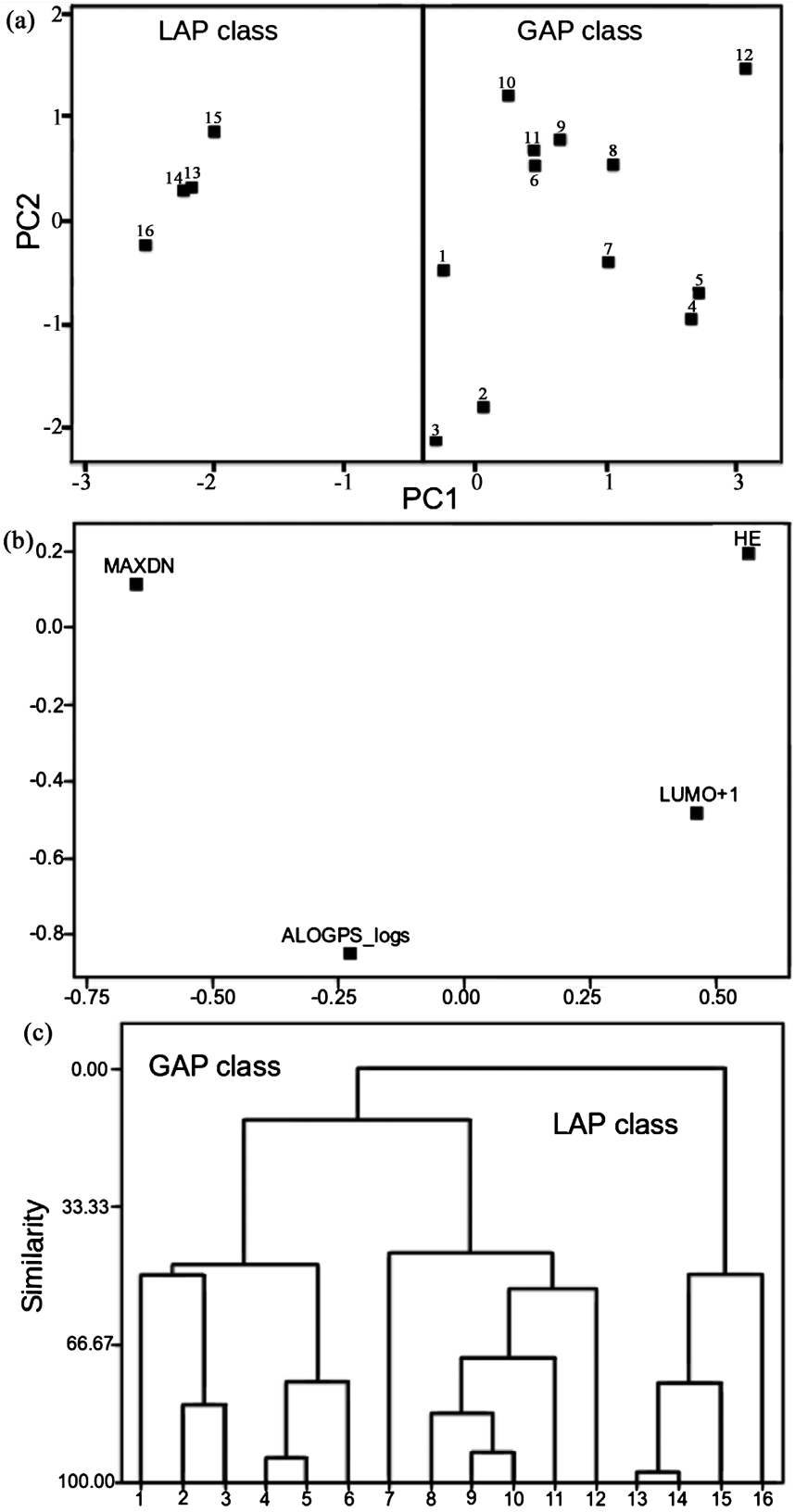

Figure 4. Scores (a), loadings (b) vector plots of the first two PCs (PC1 and PC2) for the four properties responsible for the separation and (c) dendrogram obtained with HCA for the separation of artemisinins into two classes: GAP class and LAP class against *P. falciparum.*

which are discriminated into two classes separated in PC1. The right side has samples corresponding to GAP class artemisinins (associated to plus sign in Table 2), whereas the left side has samples corresponding to LAP class artemisinins (associated to minus sign in Table 2). The molecular descriptors responsible for such a distinct classification were the LUMO + 1 energy, hydration energy (HE), maximal electrotopological negative variation (MAXDN), and Squared Ghose-Crippen octanol-water partition coeff. ($logP^2$) (ALOPS_logs), shown in Table 2. The correlation between them is less than 0.428. They were selected among the descriptors initially computed and are believed to be closely related to the biological response investigated. Figure 4(b) shows the plot of the loadings of these first two PC1-PC2. According to the figure, the LUMO + 1 energy HE, and ALONGPS_logs descriptors are responsible by the display of GAP class artemisinins of the right side, while the MAXDN descriptor displayed LAP class artemisinins on the left side of this figure.

Table 3 shows the loadings for the first principal components (PC1, PC2, and PC3). According to this table, PC1 can be expressed through the following equation:

$$\begin{aligned} \mathrm{PC1} = {} & 0.467(\mathrm{LUMO}+1\ \mathrm{energy}) + 0.573\ (\mathrm{HE}) \\ & - 0.641(\mathrm{MAXDN}) - 0.206(\mathrm{ALOGPS\_logs}) \end{aligned} \tag{2}$$

From this equation, it can be seen that more active artemisinins can be obtained when we have higher values for the LUMO + 1 energy and HE combined with lower values for the MAXDN and more negative values of ALOGPS_logs.

The objective of HC analysis was also to present artemisinins distributed in natural groupings and confirm the PC analysis. Thus, several approaches to establish links between samples/cluster were tried. All of them were of an agglomerative type, since each sample was firstly defined as its own cluster, others were then grouped together to form new clusters until all samples were part of a single cluster. The representation of the clustering results is shown in a dendrogram which depicts the similarity, with a value of 100 assigned to identical samples and a value of 0.00 to the most dissimilar samples.

The best approach chosen in HC analysis to group samples into two main classes (one for more active artemisinin and other for less active artemisinins) was based on Euclidean distances and the incremental method. The descriptors employed were the same in PC analysis, that is, LUMO + 1 energy, HE, MAXDN, and ALOGPS_logs.

Table 3. Loadings for the first principal components (PC1, PC2, and PC3).

Descriptor	PC1	PC2	PC3
LUMO + 1 energy	0.467	0.483	−0.650
HE	0.573	−0.213	0.560
MAXDN	−0.641	−0.109	−0.133
ALOGPS_logs	−0.206	0.842	0.496

In Figure 4(c), the dendrogram also shows the PC analysis plot for the artemisinins into two different classes, according to their activities. GAP-class artemisinins are on the left side, while LAP-class artemisinins are on the right side.

3.2.2. SDA Multivariate Model

The multivariate SDA method separates objects from distinct populations and allocates new objects to previously defined populations. The stepwise procedure is used and at each step the most powerful variable is inserted into the discriminant function. The selection criteria for the next variable depend on the number of groups specified.

The method is based on the F-test to establish the significance of the variables. Thus, at each step a variable will be selected on the basis of its significance and after several steps, the most significant variables are extracted from the whole set in question.

Equation 3(a) and Equation 3(b) show the discriminant functions for the more active and less active groups, respectively:

$$\text{Group more active}(\text{GAP class}) = 0.260\text{LUMO} + 1\ \text{energy} - 2.66\text{HE} - 13.1\text{MAXDN} - 0.007\text{ALOGPS\_logs} - 3.01 \quad (3a)$$

$$\text{Group less active}(\text{LAP class}) = 0.780\ \text{LUMO} + 1\ \text{energy} + 7.99\ \text{HE} + 39.4\ \text{MAXDN} + 0.021\text{ALOGPS\_logs} - 27.1 \quad (3b)$$

Table 4 shows the classification matrix obtained with Equation 3(a) and Equation 3(b) using the value of each variable for the compounds of the training set compounds. According to this table, one can note that, in the classification, the error rate was 0% and the separation into artemisinins with GAP and artemisinins with LAP is satisfactory. Now, to allocate new artemisinins with antimalarial potential in one of the Groups (Classes), according to the SDA multivariate model, it is needed just to follow the rule: 1) calculate, for the new artemisinin, the value of the most important auto scaled variables obtained in the construction of the SDA multivariate model; 2) substitute these auto scaled values in the two discrimination functions performed in this work; 3) check which discrimination function (more active compounds group or less active compounds group) presents the highest value.

The new artemisinin is more active if it is related to discrimination function of

Table 4. Classification matrix obtained by SDA.

Class[a]	Class[a]	
	GAP	LAP
GAP	12	0
LAP	0	4
Total	12	4
Percentage of correct information	100	100

[a]Class: GAP (greater antimalarial potential) and LAP (lower antimalarial potential).

the group more active and vice versa.

To determine if the model obtained is reliable, a cross-validation test was performed by using the leave-one-out method. That is, one compound is omitted from the data set and classification functions are built based on the remaining compounds. Then, the omitted compound is classified according to the classification of the generated functions. Subsequently, the omitted compound is included and a new compound is removed, and the procedure continues until the last compound is removed. Table 5 summarizes the results obtained with the cross-validation approach.

3.2.3. KNN Multivariate Model

The KNN analysis classifies the objects (artemisinins) based on the comparison of distances among them. The multivariate Euclidean distances between every pair of samples (artemisinins) with known class membership is calculated. The closest K samples (artemisinins) are used to build the model. The optimal K is determined by cross-validation applied to the studied artemisinins. The classification of a test sample is determined based on the multivariate distance of this sample regarding the K samples (studied artemisinins).

The KNN analysis was used for the validation of the initial data set. Table 6 presents the results obtained with one (1NN), two (2NN), three (3NN), four (4NN), and five (5NN) nearest neighbors. In the five cases (1NN, 2NN, 3NN, 4NN, and 5NN), the percentage of information was 100%. The 5NN model

Table 5. Classification matrix obtained by SDA with cross-validation.

Class[a]	Class[a]	
	GAP	LAP
GAP	12	0
LAP	0	4
Total	12	4
Percentage of correct information	100	100

[a]Class: GAP (grater antimalarial potential) and LAP (lower antimalarial potential).

Table 6. Classification obtained with the KNN method for or the training set (the 16 compounds).

Class[a]	Number of compounds	Compounds incorrectly				
		1NN	2NN	3NN	4NN	5NN
GAP	12	0	0	0	0	0
LAP	4	0	0	0	0	0
Total	16	0	0	0	0	0
Percentage of correct information		100	100	100	100	100

[a]GAP (greater antimalarial potential) and LAP (lower antimalarial potential).

was used [56].

As reported above, the properties: HOMO + 1 energy, HE, MAXDN, and ALOGPS_logs are important in the construction of multivariate models and, therefore, in the description of the antimalarial activity of artemisinins. For this reason, the following considerations may be relevant to understand the behavior of artemisinins with GAP.

According to the literature, the HOMO does not always provide a reliable indication for the location of the most energetic electrons in a molecule and the most reactive location for electrophiles [57]. Furthermore, it is possible that an orbital lower than the HOMO contains the most energetic electrons in the molecule [57]. In **Table 2**, one can see that, in general, the HOMO + 1 energy for artemisinins with GAP should present higher values than those corresponding to artemisinins with LAP. This may be an indication of the occurrence of biological processes resulting from the interaction between the most energetic electrons of the heme and the LUMO + 1 of artemisinins with GAP.

HE is the energy change that accompanies the hydration of one mole of ions in a solution process and involves steps associated with a change in enthalpy. In a biological environment, the apolar molecules (artemisinins) are surrounded by polar molecules (water molecules), by verifying the formation of hydrogen bonds between them [58]. In **Table 2**, it can be seen, generally, that artemisinins with GAP have higher values for HE compared to artemisinins with LAP. This may be an indicative of the importance of the mediation of water molecules in the biological process between artemisinins with LAP and the molecular receptor.

MAXDN, an electrotopological state index of atoms in molecules, represents the maximum intrinsic state difference in the molecule, and can be related to its nucleophilicity [59] [60]. In **Table 2**, in general, artemisinins with GAP have lower values for MAXDN, when compared to artemisinins with LAP. This may indicate a lower nucleophilic character of artemisins with GAP in the interaction with heme in the biological process.

ALOGPS_logs are a hydrophobic index and measures the lipophilic contributions of atoms in the molecule [61] [62]. As it can be seen in **Table 2**, in general, artemisinins with lower ALOGPS_logs values correspond to those with GAP. This may indicate the importance of hydrophobic interactions between artemisinins with GAP and heme in the biological process.

3.2.4. Designed Artemisinins with Insights from MEP, Ligand-Heme Interaction, and Multivariate Models (Prediction Set)

With the information accumulated in previous studies (MEP, interaction between artemisinins and heme, and multivariate analysis) and chemical intuition, eight new artemisinins with antimalarial potential were designed. **Figure 5** shows the 2D structures of these arteminins (prediction set).

The quality demonstrated by the previously obtained multivariate models (PCA, HCA, SDA, and KNN) led us to apply them to artemisinins from prediction set, and the results of this application are shown in **Table 7**.

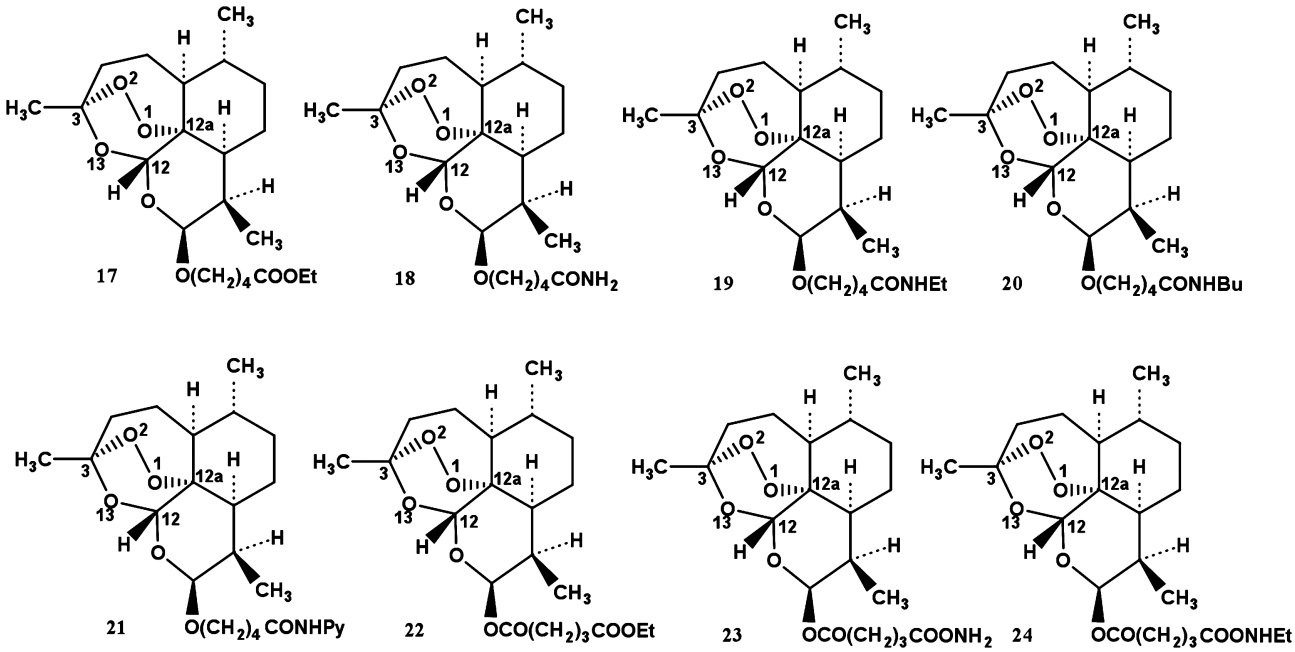

Figure 5. 2D structures of new artemisinins designed from the study of MEP, molecular docking and multivariate analysis with antimalarial potential against W-2 strains of *P. falciparum.*

Table 7. Prediction results obtained with four multivariate models for artemisinins designed from the study of MEP, molecular docking, and multivariate analysis[a].

New artemisinins	PCA multivariate model	HCA multivariate model	SDA multivariate model	KNN multivariate model
17	GAP	LAP	GAP	LAP
18	GAP	LAP	GAP	LAP
19	GAP	GAP	GAP	GAP
20	GAP	GAP	GAP	GAP
21	LAP	LAP	GAP	LAP
22	GAP	GAP	LAP	GAP
23	LAP	LAP	LAP	LAP
24	GAP	GAP	LAP	GAP

GAP (greater antimalarial potential) and LAP (lower antimalarial potential).

According to Table 7, artemisinins 17 and 18 were predicted as GAP by the PCA and SDA models, and as LAP by the HAC and KNN models, respectively. Artemisinins **19** and **20** were predicted as GAP by all four models; and artemisinin **21** as LAP by the PCA, HCA and KNN models, presenting GAP prediction by the SDA model. Furthermore, artemisinin **22** was predicted as GAP by the PCA, HCA and KNN models, and LAP by the SDA model. On the other hand, artemisinin **23** was predicted LAP by all four models; with artemisinin **24** showing the same prediction as artemisinin **22**. According to the prediction reported

in Table 7, artemisinins **19** and **20** are the most promising for synthesis and biological tests against W-2 falciparum malaria.

Table 8 shows the properties of the new artemisinins used in the prediction reported in Table 7. Here it is worth noting that the properties of artemisinins **19** and **20**, in general, follow the same behavior shown by artemisinins with GAP in Table 2.

Table 8. Molecular properties for artemisinins designed from the study of MEP, molecular docking, and multivariate analysis.

Compound	LUMO + 1 energy[a]	HE[a]	MAXDN	ALOGPS_logs
17	134.3	−6.91	2.01	−4.09
18	133.9	−6.98	2.01	−4.09
19	133.3	−2.94	2.00	−4.42
20	133.8	−1.87	2.01	−5.05
21	89.50	−5.95	2.02	−4.08
22	123.7	−2.22	2.10	−4.18
23	123.2	−7.24	2.10	−3.41
24	123.4	−3.62	2.10	−4.03

[a]kcal·mol^{-1}.

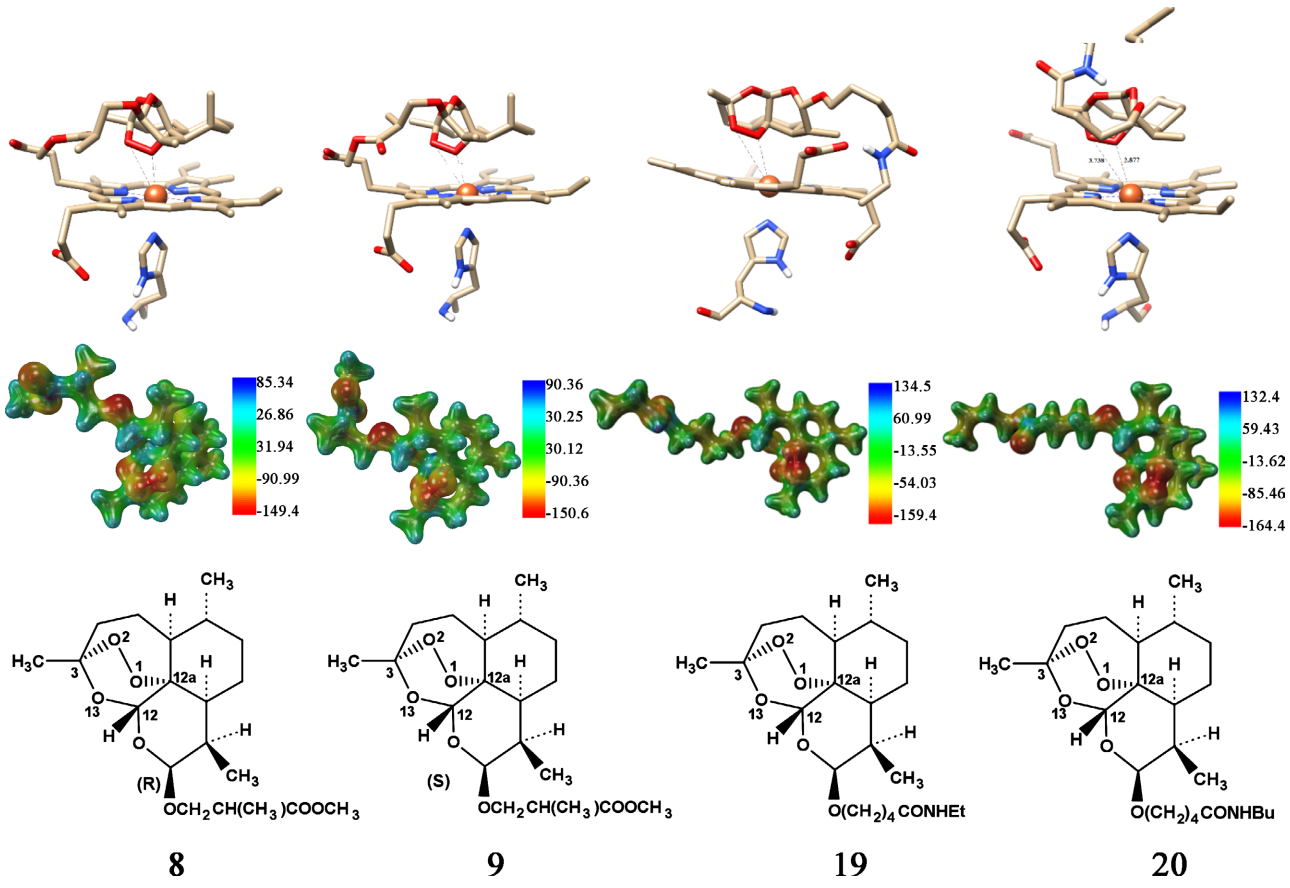

Figure 6. 2D structures (a), MEP maps (b), and molecular dockings (c) for 8, 9, 19, and 20 artemisinins from training and prediction sets, respectively.

3.2.5. MEP Maps, Ligand-Heme Interaction for Artemisinins from Training and Prediction Sets with GAP

In Figure 6, the 2D structures, the MEP maps, and the ligand-heme interactions for artemisinins from the training sets (**8** and **9**) of prediction (**19** and **20**) with GAP are shown. It can be noticed that the MEP maps of the four compounds in the trioxane ring region show the same behavior, *i.e.*: there is concentration of the electron density in the atoms O_1, O_2, and O_{13} of the 1,2,4-trioxane ring structure. It is also noted that artemisinins **19** and **20** present distances of interactions with the heme with values that are in the range evidenced in the study carried out with the training set, and when these values are compared to those artemisinins **8** and **9** with the highest GAP among the molecules of this set, it is verified that the values are very close to each other: FeO_1 and FeO_2 distances for artemisinins equal to 2.340 and 3.170 Å (**8**); 2.329 and 3.170 Å (**9**); 2.487 and 3.097 Å (**19**); 2.613 and 3.235 Å (**20**), respectively.

4. Conclusions

According to the MEP investigation, the key features necessary for the biological activities of artemisinins, Figure 1 (2-16), around the 1,2,4-trioxane ring are similar in form to those exhibited by artemisinin (1), also in this figure. This is supported by the following evidence: all artemisinins are active, according to Table 2. Furthermore, investigation has shown, as with artemisinin (1), that the MEP maps of artemisinins (2-16) exhibit contour surfaces close to the 1,2,4-trioxane ring, characterized by negative MEP, indicating an electron density concentration due to lone pairs of electrons on atoms O_1, O_2, and O_{13}, susceptible to electrophilic attack by the biological target. In addition, the electron density distribution in artemisinins around the trioxane ring induces their antimalarial behavior, this is supported by the evidence that the complexation of artemisinins with the heme involves the interaction between the peroxide region, the most negatively charged zone, with Fe^{2+} ion corresponding to the most positively charged zone in the heme.

The study of the interaction between artemisinins and the heme molecular receptor showed the polar region of the molecules containing the peroxide bond directed to the Fe^{2+} of the heme ion, with this ion preferentially approaching the O_1 atom of the ligands (2-16). The FeO_1 distances are between 2.279 and 2.718 Å, while for FeO_2 they are between 3.078 and 3.744 Å. For artemisinin (1), the calculated value of 2.701 Å practically reproduces the results of the literature.

The application of exploratory methods (PCA and HCA) and classification (SDA and KNN) to the data matrix built with molecular characteristics obtained from the calculated properties allowed the separation of artemisinins into GAP and LAP classes, with the variables LUMO + 1 energy, HE, MAXDN, and ALOPS_logs being responsible for this separation. Additionally, these variables may indicate the occurrence of biological processes involving electronic, hydration, nucleophilic, and hydrophobic interactions between artemisinins with GAP and the heme.

The accumulation of information in studies with MEP, the interaction of artemisinins and heme, and with multivariate analysis, as well as chemical intuition, allowed the design of eight new artemisinins (17-24), whose antimalarial potential was scrutinized through the PCA, HCA models, SDA, and KNN, resulting in the identification of two new artemisinins (19 and 20) with GAP. The synthesis studies, in progress at LABSínt-UFPA, and biological assays of these two new artemisinins will be able to validate the quality of the methodology used in the research. Furthermore, the reported studies can provide valuable insight for experimentalists on the processes of synthesis and biological evaluation that lead to the development of new artemisinins with antimalarial potential against *P. falciparum*.

Acknowledgements

We thank the Swiss Center for Scientific Computing for allowing the use of the MOLEKEL program. We also employed computing facilities at the Laboratório de Química Teórica e Computacional (LQTC) at the Universidade Federal do Pará (UFPA).

This work is dedicated to Professor Raimundo Dirceu de Paula Ferreira (*in memory*). Professor Dirceu was a member of this research group and contributed to many researches while he was among us.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] World Malaria Report (2021) https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021

[2] Tse, E.G., Korsik, M. and Todd, M.H. (2019) The Past, Present and Future of Anti-Malarial Medicines. *Malaria Journal*, **18**, Article No. 93. https://malariajournal.biomedcentral.com/track/pdf/10.1186/s12936-019-2724-z.pdf https://doi.org/10.1186/s12936-019-2724-z

[3] Qinghaosu Antimalarial Coordinating Research Group (1979) Antimalarial Studies on Qinghaossu. *Chinese Medical Journal*, **92**, 811-816.

[4] Li, Y. (2012) Qinghaosu (Artemisinin): Chemistry and Pharmacology. *Acta Pharmacologica Sinica*, **33**, 1141-1146. https://www.nature.com/articles/aps2012104 https://doi.org/10.1038/aps.2012.104

[5] Chang, Z. (2016) The Discovery of Qinghaosu (Artemisinin) as an Effective Anti-Malaria Drug: A Unique China Story. *Science China Life Sciences*, **59**, 81-88. https://link.springer.com/article/10.1007/s11427-015-4988-z https://doi.org/10.1007/s11427-015-4988-z

[6] O'Neill, P.M., Barton, V.E. and Ward, S.A. (2010) The Molecular Mechanism of Action of Artemisinin—The Debate Continues. *Molecules*, **15**, 1705-1721.

https://doi.org/10.3390/molecules15031705

[7] Denisov, E.T. and Denisova, T.G. (2011) Hydroxyl Mechanism of the Antimalarial Effect of Artemisinin and Its Analogs. *Russian Chemical Bulletin, International Edition*, **60**, 1421-1435. https://doi.org/10.1007/s11172-011-0213-9

[8] Rawe, S.L. (2020) Artemisinin and Artemisinin-Related Agents, In: Patrick, G.L., Ed., *Antimalarial Agents Design and Mechanism of Action*, Elsevier, London, 99-132. https://doi.org/10.1016/B978-0-08-101210-9.00004-4

[9] Wang, J., Zhang, C.-J., Chia, W.N., Loh, C.C.Y., Li, Z., Lee, Y.M., He, Y.; Yuan, L.-X., Lim, T.K., Liu, M., Liew, C.X., Lee, Y.Q., Zhang, J.Z., Lu, N., Lim, C.T., Hua, Z.C., Liu, B., Shen, H.-M., Tan, K.S.W. and Lin, Q. (2015) Haem-Activated Promiscuous Targeting of Artemisinin in *Plasmodium falciparum*. *Nature Communications*, **6**, Article No. 10111. https://doi.org/10.1038/ncomms10111

[10] Tilley, L., Straimer, J., Gnädig, N.F., Ralph, S.A. and Fidock, D.A. (2016) Artemisinin Action and Resistance in *Plasmodium falciparum*. *Trends Parasitology*, **32**, 682-696. https://doi.org/10.1016/j.pt.2016.05.010

[11] Tam, D.N.H., Tawfk G.M., El-Qushayri, A.E., Mehyar, G.M., Istanbuly, S., Karimzadeh, S., Tu, V.L., Tiwari, R., Dat, T.V., Nguyen, P.T.V., Hirayama, K. and Tien Huy, N.T. (2020) Correlation between Anti-Malarial and Anti-Haemozoin Activities of Anti-Malarial Compounds. *Malaria Journal*, **19**, Article No. 298. https://doi.org/10.1186/s12936-020-03370-x

[12] Jahan, M., Leon, F., Fronczek, F.R., Elokely, K.M, Rimoldi, J., Khan, S.I. and Avery, M.A. (2021) Structure-Activity Relationships of the Antimalarial Agent Artemisinin 10. Synthesis and Antimalarial Activity of Enantiomers of *rac*-5β-Hydroxy-D-Secoartemisinin and Analogs: Implications Regarding the Mechanism of Action. *Molecules*, **26**, Article No. 4163. https://www.mdpi.com/1420-3049/26/14/4163 https://doi.org/10.3390/molecules26144163

[13] Shandilya, A., Chacko, S., Jayaram, B. and Ghosh, I. (2013) A Plausible Mechanism for the Antimalarial Activity of Artemisinin: A Computational Approach. *Scientific Reports*, **3**, Article No. 2513. https://www.nature.com/articles/srep02513 https://doi.org/10.1038/srep02513

[14] Mok, S., Ashley, E.A., Ferreira, P.E., Zhu, L., Lin, Z., Yeo, T., Chotivanich, K., Imwong, M., Pukrittayakamee, S., Dhorda, M., Nguon, C., Lim, P., Amaratunga, C., Suon, C., Hien, T.T., Htut, Y., Faiz, M.A., Marie A. Onyamboko, M.A., Mayxay, M., Newton, P.N., Tripura, R., Woodrow, C.J., Miotto, O., Kwiatkowski, D.P., Nosten, F., Day, N.P.J, Preiser, P.R., White, N. J., Dondorp A.M., Fairhurst, R.M. and Bozdech, Z. (2014) Population Transcriptomics of Human Malaria Parasites Reveals the Mechanism of Artemisinin Resistance. *Science*, **347**, 431-435. https://doi.org/10.1126/science.1260403

[15] Mbengue, A., Bhattacharjee, S., Pandharkar, T., Liu, H., Estiu, G., Stahelin, R.V., Rizk, S.S, Njimoh, D.L., Ryan, Y., Chotivanich, K., Nguon, C., Ghorbal, M., Lopez-Rubio, J.-J., Pfrender, M., Emrich, S., Mohandas, N., Dondorp, A.M., Wiest, O. and Haldar, K. (2015) A Molecular Mechanism of Artemisinin Resistance in *Plasmodium falciparum* Malária. *Nature*, **520**, 683-687. https://www.nature.com/articles/nature14412.pdf https://doi.org/10.1038/nature14412

[16] Bernardinelli, G., Jefford, C.W., Maric, D., Thoson, C. and Weber, J. (1994) Computational Studies of the Structures and Properties of Potential Antimalarial Compounds Based on the 1,2,4-Trioxane Ring Structure. I. Artemisinin-Like Molecules. *International Journal Quantum Chemistry Quantum Biology Symposium*, **21**, 117-131. https://doi.org/10.1002/qua.560520710

[17] Roothaan, C.C.J. (1951) New Developments in Molecular Orbital Theory. *Review of Modern Physics*, **23**, 69-89. https://doi.org/10.1103/RevModPhys.23.69

[18] Binkley, J.S., Pople, J.A. and Hehre, W.J. (1980) Self-Consistent Molecular Orbital Methods. 2.1. Small Split-Valence Basis Sets for First-Row Elements. *Journal of the American Chemical Society*, **102**, 939-947. https://doi.org/10.1021/ja00523a008

[19] Lin, A.J., Lee, M. and Klayman, D.L. (1989) Antimalarial Activity of New Water-Soluble Dihydroartemisinin Derivatives. 2. Stereospecificity of the Ether Side Chain. *Journal of Medicinal Chemistry*, **32**, 1249-1252. https://doi.org/10.1021/jm00126a017

[20] Bossi, A, Venugopalan, B., Gerpe LD, Yeh, H.J.C., Flippen-Anderson, J.L., Buchs, P., Luo, X.D., Milhous, W. and Peters, W., (1988) Arteether, a New Antimalarial Drug: Synthesis and Antimalarial Properties. *Journal of Medicinal Chemistry*, **31**, 645-650. https://doi.org/10.1021/jm00398a026

[21] (1997) GaussView 1.0. Gaussian Inc., Pittsburg.

[22] Pinheiro, J.C., Ferreira, M.M.C. and Romero, O.A.S. (2001) Antimalarial Activity Dihydroartemisinin Derivatives against *P. falciparum* Resistant to Mefloquine: A Quantum Chemical and Multivariate Study. *Journal of Molecular Structure*: *Theochem*, **572**, 35-44. https://doi.org/10.1016/S0166-1280(01)00522-X

[23] Pinheiro, J.C, Kiralj, R., Ferreira, M.M.C. and Romero, O.A.S. (2003) Artemisinin Derivatives with Antimalarial Activity against *Plasmodium falciparum* Designed with the Aid of Quantum Chemical and Partial Least Squares Methods. *QSAR Combinatorial Science*, **22**, 830-842. https://doi.org/10.1002/qsar.200330829

[24] Leban, I., Golič, L. and Japelj, M. (1988) Crystal and Molecular Structure of Qinghaosu: A Redetermination. *Acta Pharmaceutica Jugoslavica*, **38**, 71-77.

[25] Lisgarten, J.N., Potter, B.S., Bantuzeko, C. and Palmer, R.A. (1998) Structure, Absolute Configuration, and Conformation on the Antimalarial Compound: Artemisinin. *Journal of Chemical Crystallography*, **28**, 539-543. https://doi.org/10.1023/A:1023244122450

[26] Frisch, A. and Frisch, M.J. (1998) Gaussian 98 User's Reference, Revision A.7. Gaussian, Inc., Pittsburg.

[27] Todeschini, R. and Consonni V. (2009) Molecular Descriptors for Chemoinformatics. Wiley-VCH, Weinheim.

[28] ChemPlus: Modular Extensions to HyperChem Release 8.06 (2008) Molecular Modeling for Windows. Hyperchem, Inc., Gainesville.

[29] Chirlian, L.E. and Francl, M.M. (1987) Atomic Charges Derived from Electrostatic Potentials: A Detailed Study. *Journal of Computational Chemistry*, **8**, 894-905. https://doi.org/10.1002/jcc.540080616

[30] Williams, D.E. and Yan, J.M. (1998) Point-Charge Models for Molecules Derived from Least Squares Fitting of the Electric Potential. *Advances in Atomic and Molecular Physics*, **23**, 87-130. https://doi.org/10.1016/S0065-2199(08)60106-2

[31] Singh, H.C. and Kollman, P.A. (1984) An Approach to Computing Electrostatic Charges for Molecules. *Journal of Computational Chemistry*, **5**, 129-145. https://doi.org/10.1002/jcc.540050204

[32] Politzer, P. and Murray, J.S. (2021) Electrostatic Potentials at the Nuclei of Atoms and Molecules. *Theoretical Chemistry Accounts*, **140**, Article No. 7. https://doi.org/10.1007/s00214-020-02701-0

[33] Politzer, J., Laurence, P.R. and Jayasuriya, K. (1985) Molecular Electrostatic Potentials: An Effective Tool for the Elucidation of Biochemical Phenomena. *Environ-*

mental Health Perspectives, **61**, 191-202. https://doi.org/10.1289/ehp.8561191

[34] Flukiger, P., Luth, H.P., Portmann, S. and Weber, J. (2000-2001) MOLEKEL. Swiss Center for Scientific Computing, Mano.

[35] Gohlke, H. (2012) Protein-Ligand Interactions. In: Mannhold, R., Kubinyi, H. and Folkers, G., Eds., *Methods and Principles in Medicinal Chemistry*, Wiley-VCH, Winheim, 331 p.

[36] Náray-Szabó, G. (2014) Protein Modelling. Springer, Berlin. https://doi.org/10.1007/978-3-319-09976-7

[37] Sethi, A., Joshi, K., Sasikala, K. and Alvala, M. (2019) Molecular Docking in Modern Drug Discovery: Principles and Recent Applications. In: Gaitonde, V., Karmakar, P. and Trivedi, A., Eds., *Drug Discovery and Development—New Advances*, IntechOpen, London, 27-48. https://doi.org/10.5772/intechopen.85991

[38] Pinzi, L. and Rastelli, G. (2019) Molecular Docking: Shifting Paradigms in Drug Discovery. *International Journal of Molecular Sciences*, **20**, 4331. https://doi.org/10.3390/ijms20184331

[39] Stanzione, F., Giangreco, I. and Cole, J.C. (2021) Use of Molecular Docking Computational Tools in Drug Discovery. In: Witty, D.R. and Cox, B., Eds., *Progress in Medicinal Chemistry*, Elsevier, New York, 273-343.

[40] Vojtechovsky, J., Chu, K., Berendzen, J., Sweet, R.M. and Schlichting, I. (1999) Crystal Structures of Myoglobin-Ligand Complexes at Near-Atomic Resolution. *Biophysical Journal*, **77**, 2153-2174. https://doi.org/10.1016/S0006-3495(99)77056-6

[41] Araújo, J.Q., Carneiro, J.M., Araújo, M.T., Laite, G.H. and Taranto, A.G. (2008) Interaction between Artemisinin and Heme. A Density Functional Theory Study of Structures and Interaction Energies. *Bioorganic & Medicinal Chemistry*, **16**, 5021-5029. https://doi.org/10.1016/j.bmc.2008.03.033

[42] Morris, G.M., Huey, R., Hart, W.E., Lindstron, W., Gillet, A., Goodsell, D. and Olson. A.J. (2009) Autodck 4.2 Program. The Scripps Research Institute, Department of Molecular Biology, MB-5, La Jolla.

[43] Trott, O. and Olson, A.J. (2009) Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *Journal of Computational Chemistry*, **31**, 455-461. https://doi.org/10.1002/jcc.21334

[44] Varmuza, K. (2018) Methods for Multivariate Data Analysis. In: Engel, T. and Gasleiger, J., Eds. *Chemoinformatics—Basic Concepts and Methods*, Wiley-VCH, Weinheim, 399-437. https://doi.org/10.1002/9783527816880.ch11_01

[45] Santos, M.A., Oliveira, L.F., Figueiredo, A.F., Gil, F.S., Farias, M.S., Bitencourt, H.R., Lobato, J.R.B., Ferreira, R. D.P., Alves, S.S.S., Aquino, E.L.C. and Ciríaco-Pinheiro, J. (2020) Molecular Electrostatic Potential and Chemometric Techniques as Tools to Design Bioactive Compounds. In: Stefaniu, A., Rasul, A. and Hussain, G., Eds., *Cheminformatics and Its Applications*, Intechopen, London, 1-28.

[46] Barbosa, J.P., Ferreira, J.E.V., Figueiredo, A.F., Almeida, R.C.O., Silva, O.P.P., Carvalho, J.R.C., Cristino, M.G.G., Ciríaco-Pinheiro, J., Vieira, J.L.F. and Serra, R.T.A. (2011) Molecular Modeling and Chemometric Study of Anticancer Derivatives of Artemisinin. *Journal of the Serbian Chemical Society*, **76**, 1263-1282. https://doi.org/10.2298/JSC111227111B

[47] Cristino, M.G.G., Meneses, C.C.F., Soeiro, M.M., Ferreira, J.E.V., Figueiredo, A.F., Barbosa, J.P., Almeida, R.C.O., Pinheiro, J.C. and Pinheiro, A.L.R. (2012) Computational Modeling of Antimalarial 10-Substituted Deoxoartemisinins. *Journal*

of Theoretical and Computational Chemistry, **11**, 241-263. https://doi.org/10.1142/S0219633612500162

[48] Oliveira, L.F.S., Cordeiro, H.C., Brito, H.G., Pinheiro, A.C.B., Santos, M.A.B., Bitencourt, H.R., Figueiredo, A.F., Araújo, J.J.O., Gil, F.S., Farias, M.S., Barbosa, J.P. and Pinheiro, J.C. (2021) Molecular Electrostatic Potential and Pattern Recognition Models to Design Potentially Active Pentamidine Derivatives against *Trypanosoma brucei* Rhodesiense. *Research, Society and Development*, **10**, e261101220207. https://doi.org/10.33448/rsd-v10i12.20207

[49] (2001) Pirouette 3.01. Infometrix, Inc., Woodinville.

[50] Jefford, C.W. (2001) Why Artemisinin and Certain Synthetic Peroxides Are Potent Antimalarials. Implications for the Mode of Action. *Current Medicinal Chemistry*, **8**, 1803-1826. https://doi.org/10.2174/0929867013371608

[51] Cheng, F., Shen, J., Luo, X., Zhu, W., Gu, J., Ji, R., Jiang, H. and Chen, K. (2002) Molecular Docking and 3-D-QSAR Studies on the Possible Antimalarial Mechanism of Artemisinin Analogues. *Bioorganic & Medicinal Chememistry*, **10**, 2883-2891. https://doi.org/10.1016/S0968-0896(02)00161-X

[52] Tonmunphean, S., Parasuk, V. and Kokpol, S. (2001) Automated Calculation of Docking of Artemisinin to Heme. *Journal of Molecular Modeling*, **7**, 26-33. https://doi.org/10.1007/s008940100013

[53] Polman, S., Kokpol, S., Hannongbua, S. and Rode, B.M. (1989) Quantum Pharmacological Analysis of Structure-Activity Relationships for Mefloquine Antimalarial Drugs. *Analytical Sciences*, **5**, 641-646. https://doi.org/10.2116/analsci.5.641

[54] White, N.J. (2004) Antimalarial Drug Resistance. *The Journal of Clinical Investigation*, **113**, 1084-1092. https://doi.org/10.1172/JCI21682

[55] Ferreira, M.M.C., Montanari, C.A. and Gaudio, A.C. (2002) Seleção de variáveis em QSAR. *Quim Nova*, **25**, 439-448. https://doi.org/10.1590/S0100-40422002000300017

[56] Ferreira, M.M.C. (2015) Químiometria: Conceitos, Métodos e Aplicações. Editora UNICAMP, Campinas. https://doi.org/10.7476/9788526814714

[57] Bulat, F.A., Murray, J.S. and Politzer, P. (2021) Identifying the Most Energetic Electrons in a Molecule: The Highest Occupied Molecular Orbital and the Average Local Ionization Energy. *Computational and Theoretical Chemistry*, **1199**, Article ID: 113192. https://doi.org/10.1016/j.comptc.2021.113192

[58] Soualmia, F., Belaidi, S., Tchouar, N., Lanez, T. and Boudergua, S. (2021) QSAR Studies and Structure Property/Activity Relationships Applied in Pyrazine Derivatives as Antiproliferative Agents against the BGC823. *Acta Chimica Slovenica*, **68**, 882-895. https://doi.org/10.17344/acsi.2021.6875

[59] Gramatica, P., Corradi, M. and Consonni, V. (2000) Modelling and Prediction of Soil Sorption Coefficients of Non-Ionic Organic Pesticides by Molecular Descriptors. *Chemosphere*, **41**, 763-777. https://doi.org/10.1016/S0045-6535(99)00463-4

[60] Kier, L.B., Hall, L.H. and Frazer, J.W. (1991) An Index of Electrotopological State for Atoms in Molecules. *Journal of Mathematical Chemistry*, **7**, 229-241. https://doi.org/10.1007/BF01200825

[61] Ghose, A.K., Viswanadhan, V.N. and Wendoloski, J.J. (1998) Prediction of Hydrophilic (Lipophilic) Properties of Small Organic Molecules Using Fragmental Methods: An Analysis of ALOGP and CLOGP Methods. *The Journal of Physical Chemistry*, **102**, 3762-3772. https://doi.org/10.1021/jp980230o

[62] Todeschini, R. and Consonni, V. (2000) Handbook of Molecular Descriptors, Wiley-VCH, Weinheiem. https://doi.org/10.1002/9783527613106