Bioorganic & Medicinal Chemistry 17 (2009) 5325–5331

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

In vitro erythrocytic uptake studies of artemisinin and selected derivatives using LC–MS and 2D-QSAR analysis of uptake in parasitized erythrocytes

Falgun Shah [a,†], Shuang-Qing Zhang [b,†], Shilpa Prakash Kandhari [a], Prasenjit Mukherjee [b], Amar Chittiboyina [b], Mitchell A. Avery [b], Bonnie A. Avery [a,*]

[a] Department of Pharmaceutics, School of Pharmacy, The University of Mississippi, University, MS 38677, USA
[b] Department of Medicinal Chemistry, School of Pharmacy, University of Mississippi, University, MS 38677, USA

ARTICLE INFO

Article history:
Received 4 February 2009
Revised 30 April 2009
Accepted 1 May 2009
Available online 18 May 2009

Keywords:
Malaria
Artemisinin
Erythrocytic uptake
Liquid chromatography–mass spectrometry
Multiple Linear Regression

ABSTRACT

The purpose of the present investigation was to characterize the partitioning of artemisinin and its derivatives into both non-parasitized as well as *Plasmodium falciparum* parasitized red blood cells (RBCs). Artemisinin and selected derivatives at concentrations of 3.55 μM were incubated in RBCs with a hematocrit of 33% for 2 h at 37 °C, extracted from RBCs by solid phase extraction, and analyzed using liquid chromatography–mass spectrometry in positive electro-spray ionization mode with methanol as mobile phase. The uptake percent of artemisinin and selected derivatives into the non-parasitized RBCs ranged between 35% and 45%, while that into parasitized RBCs was between 51% and 72%. The results suggested that artemisinin and selected derivatives were preferentially distributed in parasitized RBCs. A Multiple Linear Regression model was built to gain insight about the essential structural properties required for the uptake of this class of compounds in parasitized RBCs and will provide instruction for designing of new derivatives of this class of compounds with improved uptake.

Published by Elsevier Ltd.

1. Introduction

Malaria affects an estimated 270 million people each year, especially in Africa and Southern Asia.[1] Resistance to the existing antimalarial treatment has become widespread, and there is an urgent need for the development of new drugs.[1] One of the most promising new classes of antimalarial drugs is the derivatives of artemisinin, a drug that was originally isolated from an ancient Chinese herbal remedy.[2] Artemisinin derivatives are the most rapidly acting of all antimalarial drugs, effective by all routes of administration. They already have an established role in the treatment of multidrug-resistant *Plasmodium falciparum* malaria, and no serious adverse effects have been reported in clinical trials.[3,4]

Although artemisinin derivatives have been extensively studied in an attempt to uncover a mode of action, little information concerning red blood cell (RBCs) uptake with regard to structure is available despite the observation that the artemisinin class of antimalarials is preferentially taken up into malaria parasitized erythrocytes as compared to non-parasitized RBCs.[5–8] Our first attempts to understand this phenomenon revealed that artemisinin and derivatives are selectively taken into parasitized RBCs by a passive but facilitated uptake. Those results suggested that this specificity could be sensitive to the molecular properties of the peroxide moiety. Hence, the uptake of artemisinin and select derivatives into parasitized RBCs may be an important parameter in drug potency and is therefore worthy of investigation.

Although numerous methods for the analysis of specific artemisinin derivatives have been developed, most are not widely adaptable to the extensive range of analogs that need to be tested. The development of a selective and sensitive analytical method for the determination of artemisinin derivatives poses challenging problems because they lack ultraviolet absorption or fluorescent chromophores and do not possess functional groups with potential for derivatization. Liquid chromatography–mass spectroscopy (LC–MS) gives definitive identification and quantitative determination of compounds, and provides high sensitivity and selectivity for all analytes of interest.

The purpose of this investigation was to study the uptake of artemisinin and molecularly diverse derivatives by parasitized and non-parasitized RBCs analytically using LC–MS. The resulting quantitative data should provide an important parameter in the overall pharmacological equation defining drug potency (e.g., K_i, ED_{50}, and IC_{50}) in this class of schizonticidal-staged antimalarial agents. Thus, we expected to conduct quantitative structure–activity relationship studies (QSAR) that would, if statistically valid, establish a mathematical relationship between uptake and molecular properties of the peroxidic antimalarials to be studied.

* Corresponding author. Tel.: +1 662 915 5163; fax: +1 662 915 1177.
E-mail addresses: mavery@olemiss.edu (M.A. Avery), bavery@olemiss.edu (B.A. Avery).
† Both authors contributed equally to the work.

0968-0896/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.bmc.2009.05.026

While we and others have studied the QSAR of artemisinin and its derivatives extensively, all such efforts were conducted with biological potency as the dependent variable. Now, we hope to correlate uptake with the structural properties of artemisinin and its select derivatives which could lead to an expanded ability to conduct drug design and development in this class of antimalarial agents.

2. Materials and methods

2.1. Materials

HPLC grade methanol, toluene, physiological saline and RPMI-1640 medium were purchased from Fisher Scientific (Fair Lawn, NJ, USA). Dulbecco's Phosphate Buffered Saline and dimethyldichlorosilane were from Sigma (St. Louis, MO, USA). MP1 (3-mL, 15 mg) solid phase extraction (SPE) cartridge was obtained from Varian (Lake Forest, CA, USA). Silanized HPLC vials (1.5 mL) were supplied by Waters (Milford, MA, USA). Human type A^+ RBCs was obtained from Mississippi Blood Services (Jackson, MS, USA).

2.2. Liquid chromatography–mass spectrometry

The HPLC system consisted of a Waters Alliance 2695 Separation Module with a 996 PDA detector. HPLC was performed at room temperature using a Diamondbond C_{18} column (particle size 3 μm, 150 × 4.6 mm, Zirchrom, Anoka, MN, USA), with methanol as a mobile phase at a flow rate of 0.3 mL min^{-1}. An aliquot (50 μL) of the samples were injected, and the run time was 35 min.

Masses were acquired on a Waters Micromass ZQ Mass Spectrometry using positive electro-spray ionization (ESI). Data acquisition and analyses were carried out using MassLynx version 4.0 software. Nitrogen was used as the nebulizing gas (50 L h^{-1}) and the desolvation gas (300 L h^{-1}) with a desolvation temperature at 200 °C. The source temperature was set at 100 °C. The capillary voltage was set at 3.0 kV and cone voltage at 60 V. Artemisinin and select derivatives were quantified in the positive mode using selected ion monitoring (SIM) for their respective molecular masses and/or adducts.

2.3. Non-parasitized and parasitized erythrocytes

The D_6 strain of *P. falciparum* was cultivated in human type A^+ RBCs and sub-cultured every day. Culture suspensions were prepared in 10% malaria complete RPMI-1640 medium in T-25 Corning tissue culture flasks and gassed using a gas mixture comprising 90% nitrogen, 5% oxygen, and 5% carbon dioxide. The flasks were then incubated at 37 °C for 42–44 h to reach trophozoite stage, the metabolically active form of the malarial parasite within the RBCs. Prior to each experiment, the parasitemia was determined by using a thin blood film stained with Diff-Quick Solutions I and II and was counted by an Olympus BH-2 light microscope. The parasitemia was maintained between 6% and 7% in all studies.

2.4. Uptake by non-parasitized and parasitized erythrocytes

Fresh non-parasitized RBCs were washed twice with Dulbecco's phosphate-buffered saline and centrifuged at 10,000 *g* for 5 min. The white blood cells and platelets were removed with a glass transfer pipette and discarded. Parasitized RBCs were centrifuged, and the culture medium was also discarded. The methanol from 700 μL of artemisinin or select derivatives was evaporated under dry-stream nitrogen, and the drug was reconstituted in 10.0 mL of physiological saline. An aliquot (1 mL) of drug solution was further diluted with 3.7 mL of saline in a 25-mL conical flask along with 2.3 mL of non-parasitized or parasitized RBCs. For the blank, 2.3 mL of saline was added in place of non-parasitized or parasitized RBCs. The glassware was treated with 5% (v/v) dimethyldichlorosilane in toluene to minimize the possibility of glass-binding prior to use. The final drug concentration was 3.55 μM, with a RBC hematocrit of 33%. Samples were incubated at 37 °C in a shaker water bath for 2 h. At the end of the incubation period, the contents were transferred to centrifuge tubes and spun at 10,000 *g* for 5 min to separate the supernatant and pellets. The supernatant was carefully withdrawn and transferred to a MP1 SPE cartridge preconditioned by sequential washing with methanol and distilled water (1.0 mL of each). After application of the samples, the compounds were eluted by washing the cartridges with methanol (2 × 0.5 mL) under reduced pressure. The eluents were evaporated to dryness under a stream of nitrogen at 40 °C. The residue was reconstituted in 950 μL of methanol and 50 μL of internal standard (IS, dihydro-artemisinin was used as the IS of artemisinin, and artemisinin was used as the IS of derivatives) in methanol at a concentration of 100 μg mL^{-1}. An aliquot (50 μL) of samples were injected into the LC–MS.

2.5. Development of a Multiple Linear Regression model to predict the percent uptake of artemisinin and its derivatives to parasitized RBCs

Multiple Linear Regression (MLR) which demonstrates great ease of implementation and the interpretability of resulting equations was the statistical method of choice for building the QSAR model. The forward-stepping variant of MLR was utilized, starting with the selection of a single variable which contributes most to the model based on its highest *F*-statistics or lowest *p*-value. At each step, MLR alters the model from the previous step by adding predictor variables and terminating the search when a statistically significant model has been obtained.[9] In the present study, the percentage uptake of artemisinin and selected derivatives by parasitized red blood cells (RBCs) was used as the dependent variable. To examine the predictive ability of the QSAR model, four artemisinin derivatives (**1**, **5**, **10**, and **19**) were selected from 23 compounds at random to construct the external test set, and the remaining 19 artemisinin derivatives comprised the training set that was employed to calibrate the QSAR model. Figure 1 depicts all the structures used in this study, and their biological activities (uptake percent of artemisinin and its derivatives in parasitized RBC) are mentioned in Table 2. 3-D structures of all the compounds were built in Sybyl 7.1 (Tripos Inc., St. Louis, MO). Compounds **4**, **12** and **15** were modeled as charged molecule. The structures were refined using 10,000 steps of steepest descent to a RMS convergence of 0.01 Kcal/mol Å using Tripos force field and the Gasteiger–Huckel partial charges method. The minimized structures were transferred to the Chemistry Development Kit (CDK) version 0.78, which is a freely available open-source Java library for the calculation of 2D as well as 3D descriptors. The suite of descriptors calculated by CDK include geometric, topological, electronic and hybrid descriptors.[10] Strike 1.5 (Schrodinger, LLC, New York, NY, 2005) was used for building as well as statistical validation of the MLR model.

In all, 174 descriptors were calculated for each structure in the data set. A reduced set of 75 descriptors was obtained after the removal of constant, near constant and highly inter-correlated ($R^2 > 0.8$) descriptors as they contain nearly identical information. Selection of the best model was made based on a variety of statistical parameters, including the squared correlation coefficient (R^2),

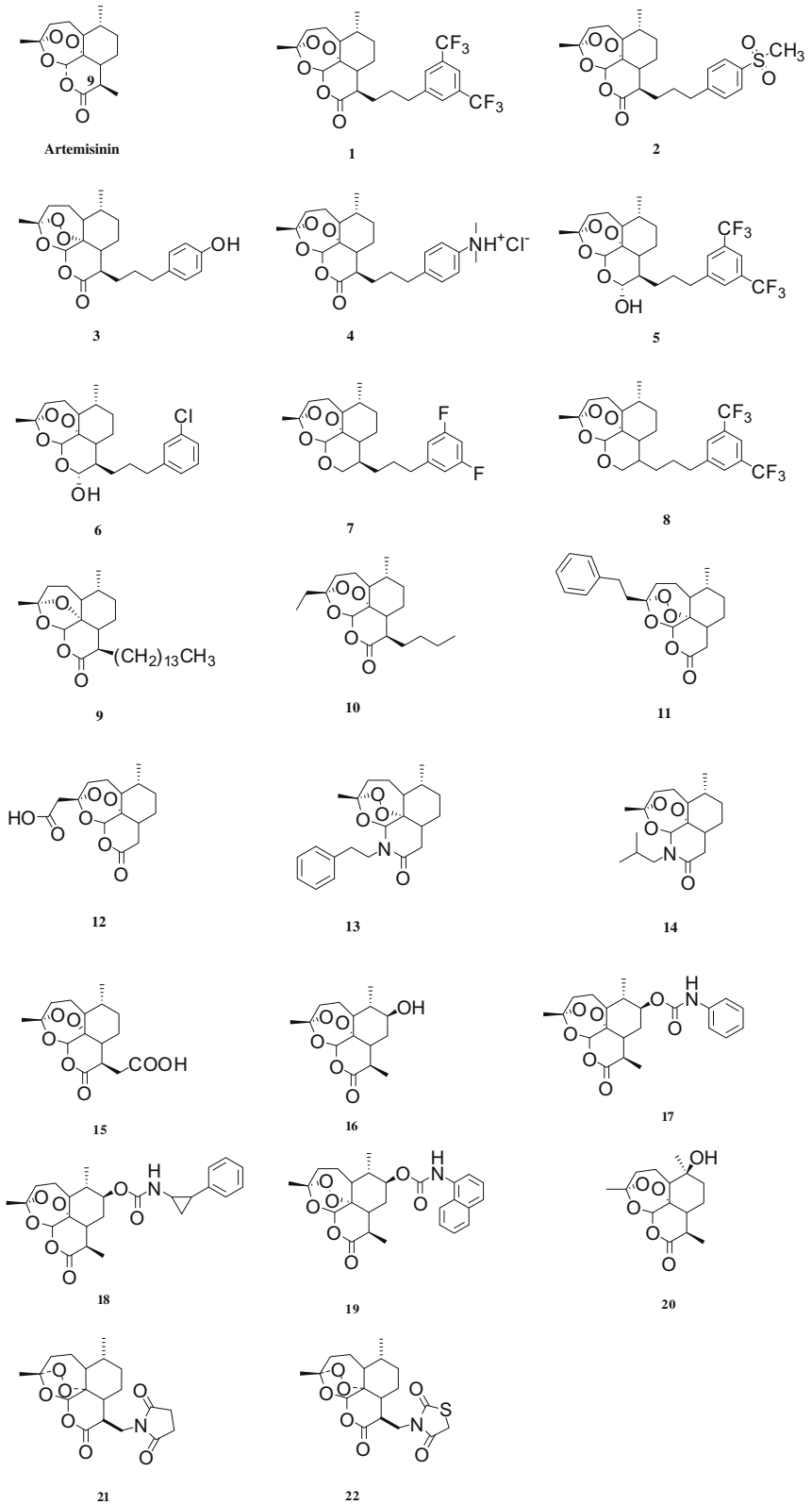

Figure 1. Structure of artemisinin and its derivatives used in study.

the standard deviation (S.D.), Fisher's value *F* and *t* value of greater than four for all descriptors in the regression model. Statistical validation of the generated models was carried out using the 'leave-one-out' (LOO) procedure of the cross validation method and predictions for the external test set. Analysis of residuals from the regression equations was used to identify outliers.

3. Results and discussion

The retention time of artemisinin and its derivatives is presented in Table 1. Calibration curves constructed for these compounds were linear with correlation coefficients greater than 0.99, as shown in Table 1. The mean intra- and inter-day accuracy

Table 1
Retention time and calibration curve of artemisinin and its derivatives

Compound	Retention time (min)	Calibration curve	Correlation coefficient
Artemisinin	18.62	$y = 1.6996x + 7.1505$	0.996
1	6.80	$y = 1.3336x + 0.8951$	0.995
2	5.94	$y = 0.6112x + 2.5464$	0.994
3	5.98	$y = 1.3737x - 0.1507$	0.991
4	6.16	$y = 1.1550x - 0.4983$	0.994
5	7.58	$y = 0.6283x + 5.3280$	0.991
6	6.53	$y = 0.6985x + 2.1063$	0.999
7	6.15	$y = 0.2332x + 0.6236$	0.999
8	10.83	$y = 0.8068x + 3.5364$	0.994
9	11.19	$y = 3.3788x - 1.3813$	0.995
10	23.39	$y = 2.4521 - 1.1907$	0.992
11	12.31	$y = 1.2421 + 0.9863$	0.996
12	8.64	$y = 1.3625x - 0.4186$	0.995
13	8.20	$y = 5.3449x - 3.3645$	0.999
14	12.47	$y = 1.1613x - 0.0195$	0.992
15	9.08	$y = 1.1822x + 0.1012$	0.997
16	9.16	$y = 1.3152x - 1.7484$	0.996
17	9.88	$y = 1.5559x - 1.6950$	0.985
18	7.33	$y = 2.3257x - 4.3899$	0.999
19	9.75	$y = 0.4226x + 3.8057$	0.998
20	9.41	$y = 1.0254x - 1.7968$	0.992
21	12.32	$y = 5.6287x - 0.2399$	0.995
22	17.35	$y = 3.2272x - 3.4672$	0.997

Note: y is the concentration of compound to be detected, and $x = \frac{\text{concentration of IS}}{\text{peak area of IS}} \times$ peak area of compound to be detected.

Table 2
The uptake percent of artemisinin and its derivatives by non-parasitized and parasitized red blood cells (RBCs) with a hematocrit of 33%

Compound	Parasitized RBCs	Non-parasitized RBCs
Artemisinin	61.07	34.68
1	72.09	41.07
2	61.90	38.53
3	63.09	39.01
4	51.92	32.82
5	71.76	40.43
6	71.21	40.13
7	70.91	40.01
8	72.00	42.11
9	73.00	45.09
10	62.50	39.11
11	62.34	38.96
12	54.51	34.03
13	61.80	38.43
14	61.51	35.17
15	54.23	33.92
16	53.90	33.10
17	51.04	32.18
18	51.63	32.31
19	51.68	32.63
20	53.89	32.99
21	66.06	36.57
22	57.69	33.94

Samples were incubated for 2 h at 37 °C with a concentration of 3.55 μM.

ranges from 97% to 104% with a value of less than 2% in precision. The limit of detection and extraction recovery was less than 10 ng and larger than 90%, respectively.

In the ultraviolet region, artemisinin and most of its derivatives absorb between 210 and 220 nm but possess very poor extinction coefficients, resulting in reduced sensitivity. Therefore, pre- and post-column derivatization methods have been used to convert artemisinin into a UV active compound that absorbs with a large extinction coefficient at longer wavelengths.[11,12] A major disadvantage of this method is interference with other constitutes in the extracts at the absorption wavelength of derivatized compounds. High-performance liquid chromatography employing various detection methods such as chemiluminescent, reductive electrochemical detection and evaporative light-scattering detector are available for the quantization of artemisinin.[13–15] These methods were not applied to some artemisinin derivatives. One of the most sensitive methods developed is the radio-labeling of the artemisinin compounds using ^{3}H or ^{14}C isotopes.[8,16] Although, the detection of radiolabelled compounds is extremely sensitive, it is very expensive and time consuming. LC–MS is known to be a powerful separation and detection technique in a large number of analytical fields, and particularly for the detection of drugs in biological fluids. High sensitivity and selectivity of mass spectrometry present several advantages.

Based on the previous results from our research group[8], an incubation time of 2 h and hematocrit of 33% were employed in the present investigation. The uptake percent of artemisinin and derivatives into the non-parasitized RBCs ranged between 35% and 45%, while that into parasitized RBCs was between 51% and 72%, as shown in Table 2. The antimalarial activity of the artemisinin class of compounds is considered to be due in part to the presence of the C–O–O–C endoperoxide bridge.[17] Derivative **9** showed the highest percent uptake into the infected RBCs, however it was a mono-oxo derivative, that is, it lacked the peroxide bridge responsible for parasiticidal activity. Derivatives **1** and **5** through **8** exhibited about 70% uptake into the parasitized RBCs and possessed C9-aryl alkyl groups along with halogen atoms on the benzene rings. Derivatives **2** through **4** belonging to this group showed considerably lower uptake, suggesting that the presence of SO_2CH_3 or a hydroxyl group on the benzene ring hindered the uptake. Figure 2 shows the scatter plot of uptake percent of the derivatives into the parasitized RBCs versus their molecular weights. Derivatives with molecular weights below 325 exhibited an average of 54% uptake, except artemisinin (molecular weight of 282), with 61% uptake. Derivatives with molecular weights ranging from 325 to 405 showed an average uptake of about 62%. Other artemisinin derivatives with molecular weights greater than 500 exhibited above 70% uptake into parasitized RBCs.

RBCs are the host cells for malaria parasites, and any effect of the drug on red cell membranes might be relevant for its in vivo effects.[18] Parasitized RBCs are considerably modified, since the parasites evolve various stratagems in order to secure an adequate traffic of solutes for their survival.[19,20] The parasitized erythrocyte exhibits a substantial increase in its permeability to high molecular weight solutes as compared to the uninfected erythrocyte. Much of this increase in permeability can be attribute to a single type of permeation pathway with characteristics quite distinct from those of the host erythrocytes. The parasitized RBC membrane has alterations in the distribution of phospholipids between the inner and outer sheaths and may contain parasite-produced proteins, clefts and pores which may serve as integral membrane transporters of drugs like artemisinin. Artemisinin and its derivatives can thus be transported into the RBC cytosol with these proteins as carriers. Significantly higher partitioning ratios of artemisinin and derivatives in parasitized RBCs than that in non-parasitized RBCs confirmed the selective accumulation of these antimalarial compounds in parasitized RBCs. Similar observations achieved by Asawamahasakda et al. showed that tritiated dihydro-artemisinin was taken up by infected RBC membranes but not by non-parasitized RBCs.[18]

To further our understanding of structure property relationships of the compounds, we developed 2D-QSAR model using the descriptors set from the CDK kit to predict the percent uptake of artemisinin and select analogues into parasitized RBC by the forward-stepping MLR method. We began with a pool of 174 descriptors ranging from topological, geometrical, constitutional, hybrid and electronic classes as implemented in CDK. This set was further

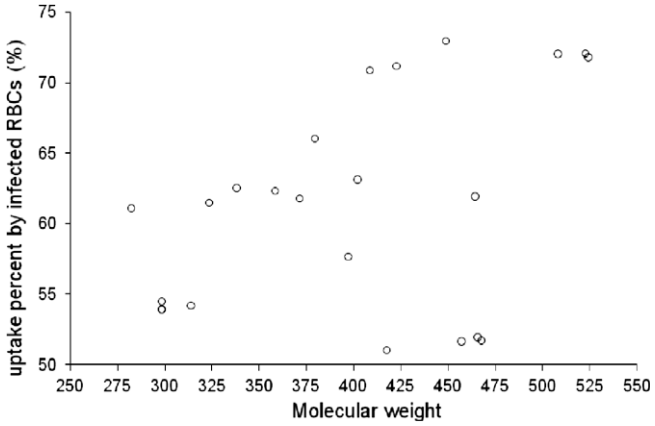

Figure 2. Uptake percent-molecular weight profiles of artemisinin and its derivatives incubated in parasitized red blood cells for 2 h at 37 °C with a concentration of 3.55 μM.

curtailed down to 75 (discussed in method section). A data set of 24 compounds with percent uptake in parasitized RBCs as property was split into the training and test set of 19 and 4 compounds, respectively. The initial model with 19 compounds in the training set resulted in a R^2 value of 0.63 with a poor F, q^2 and standard deviation value as depicted in Table 3 as model 1. The analysis of residuals identified compounds **4**, 6 and **21** as outliers as shown in Figure 3a. After deleting these compounds from the training set, MLR was carried out again and the statistics of the resultant model improved drastically as shown in Table 3 as model 2. The R^2 for the training set was 0.95 with standard deviations of 1.83. Figure 3b shows the plot for experimental versus calculated percent uptake of artemisinin and its derivates into parasitized RBCs after removal of outliers. Root mean square deviation (RMSD) for the training set was 1.59. All calculated t-values for the individual descriptors were significant with low p values which confirm the significance of each of the descriptors. The F statistic (on 3 and 13 degrees of freedom) for this model was 71.0 with a p value of 6.59×10^{-08}. The model obtained was further validated by calculating the cross validated R^2 (q^2) values obtained using the leave one out cross validation method. This is the measure of the predictive power of a regression equation. The q^2 value for the best regression model was 0.91, which is suggestive of a robust model. Finally, the randomized R^2 was calculated to ensure that the linear models were not due to chance correlations. The dependent variable for the training set was scrambled and linear models were built with the randomized dependent variables. Poor results of scrambling as depicted by the R^2 value of 0.20 and standard deviation of 7.09 suggest a statistically significant QSAR relationship with the real dependent variable.[21] The R^2 for the external set was 0.98 with SD of 3.77 and RMSD of 3.43 which displayed a good predictive ability of the developed model.

The best linear model consisted of the three descriptors tabulated in Table 4. A positive value of the regression coefficient suggests that the indicated descriptor contributes positively to the value of percent uptake, whereas a negative value is indicative of a lower percentage uptake of compounds into the parasitized RBCs.

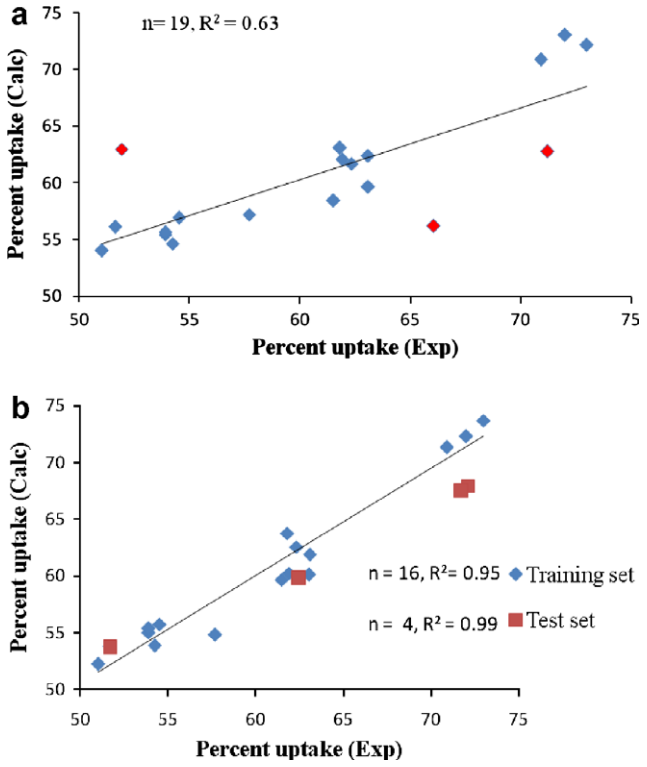

Figure 3. Correlation between experimental and calculated percent uptake into parasitized RBCs (a) with outliers (red) and (b) after removing outliers for training set.

In this regard, nHBAcc and Wlambda2.unity decreases, whereas DPSA-3 increases the percentage uptake in parasitized RBCs. The correlation matrix for the independent variables is depicted in Table 5. The matrix suggests that the variables are orthogonal to each other. The minimum and maximum values for all three descriptors and dependent variables are mentioned in Table 6. The nHBAcc descriptor stands for the number of hydrogen bond acceptors, and it corresponds to the number of nitrogen (N) and oxygen (O) atoms in the molecule. Here, any O and N atoms where the formal charge is non-positive (i.e., formal charge $\leqslant 0$) except an aromatic ether oxygen (i.e., an ether oxygen adjacent to at least one aromatic carbon) as well as an oxygen that was adjacent to a nitrogen and any nitrogen that was adjacent to an oxygen was considered a hydrogen bond acceptor.[10] Hydrogen bonding is an important parameter for permeability. Furthermore, to look for the trends in the variation of the percent uptake into parasitized RBCs by calculated descriptors value, we divided the datasets into the groups of the higher (>70%) and the lower percent uptakes (<60%). In this regard, the average values for the number of hydrogen bond acceptors for compounds having higher and lower percent uptake were 4.5 and 7, respectively. For instance, derivatives **8** and **9** (Fig. 1), having a higher uptakes, possess nHBAcc values of 4 whereas compounds **18** and **19**, with low percent uptake in parasitized RBCs, have a value of 8. So, this descriptor suggested that an increase

Table 3
MLR regression statistics

Model	N	S.D	R^2	F	P	LOO		Randomization	
						q^2	RMS	r^2	S.D.
1	19	4.87	0.63	8.7 (3.0, 15.0)	$1.41e^{-03}$	0.51	5.11	0.16	7.38
2	16	1.83	0.95	71 (3.0, 12.0)	$6.59e^{-08}$	0.91	2.07	0.2	7.09

Table 4
Best MLR model and *t*-values

Variable	Coefficient	Standard error	*t*-value
Intercept	90.117	3.4852	25.8568
Wlambda2.unity	−5.2747	1.0349	5.0967
nHBAcc	−4.5735	0.3369	13.5727
DPSA-3	0.19465	0.0317	6.1487

Wlambda2.unity: directional whim size descriptor of the second eigenvalue λ_2 of the weighted covariance matrix of atomic coordinates; nHBAcc: number of hydrogen bond acceptor; DPSA-3: difference in the charged partial surface area of a molecule.

Table 5
The correlation matrix of independent variables of MLR equation

	Wlambda2.unity	nHBAcc	DPSA-3
Wlambda2.unity	1.0000		
nHBAcc	−0.1312	1.0000	
DPSA-3	0.3333	0.1616	1.0000

Table 6
Minimum and maximum values for the descriptors and the dependent variable used in the development of MLR model

	Dependent variable	nHBAcc	DPSA-3	Wlambda2.unity
Maximum	73	8	95.2751	3.4730
Minimum	51.04	4	32.7466	1.4773

in the number of hydrogen bond acceptor decreased the percent uptake in parasitized RBC.

The second descriptor DPSA-3 enumerates the difference in the charged partial surface area of a molecule calculated using the formula

$$\text{DPSA-3} = \text{PPSA-3} - \text{PNSA-3},$$

where PPSA-3 and PNSA-3 are the partial atomic charge weighted positive and negative surface area, respectively. PPSA-3 and PNSA-3 are the sum of the product of solvent-accessible surface area and the partial charge for all positive and negative charged atoms, respectively.[12] This descriptor merges the information on the atomic contributions to the solvent-accessible surface area of the molecule with partial atomic charge of atoms and, thus, utilizes the shape and electronic information to characterize the molecules. The charged partial surface area descriptors are useful in predicting the variance of the structure property distribution.[22] Higher values of this descriptor were suggestive of the increased percent uptake. The average difference in the charged partial surface areas of the group with a high percent uptake (>70%) into RBCs was 81, whereas the average for the group with the lower percent uptake (<60%) into RBCs was 56.4. Therefore, an increase in the value of this descriptor proportionally increased the percent uptake.

Wlambda2.unity is the directional WHIM (Weighted Holistic Invariant Molecular) size descriptor of the second eigenvalue (λ_2) of the weighted covariance matrix of atomic coordinates that is designed to capture 3D information regarding size. This descriptor accounts for the molecular size along the principal axes.[23] Here, the unit weighting scheme was used wherein no distinctions were made among the atoms of the molecules during the calculation of their values. Considering the view of the molecules **9** and **16** (see Fig. 4), along the principle axis (*X*-axis), an increase in the length of the alkyl chain as in 9 corresponds to a lower descriptor value 1.477 but higher uptake, 73%, compared to 16 (high descriptor value of 3.423 and low percent uptake of 53.9). Hence, this descriptor value was negatively correlated with the percent uptake. Also,

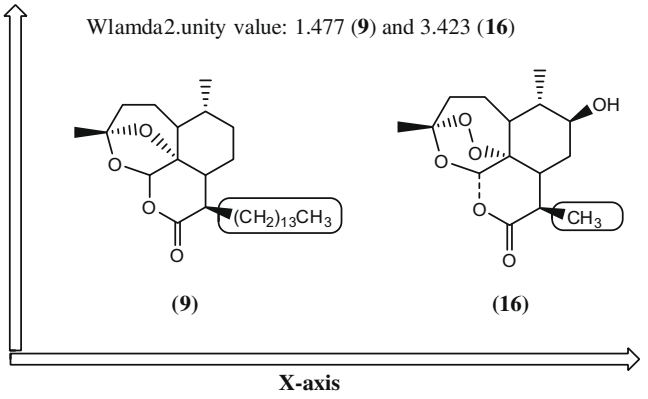

Figure 4. Explanation for wlamda2.unity values for compounds **9** and **16** having percent uptake into RBCs of 73 and 53.9, respectively.

compounds **11**–**14** with no substitution and compounds **16**–**20** having only methyl group at C-9 position exhibited average percent uptake of 55.81 into parasitized RBCs, whereas compounds **5**–**10** with minimum four carbon chain length at C-9 position exhibited an average percent uptake of 70.23. Thus, this descriptor emphasizes that an increase in the alkyl chain length at C-9 position of artemisinin along the *x*-axis led to an increase in the percent uptake into parasitized RBCs. In other words, uptakes of stretched molecules were higher as compared to spherical molecules.

The current MLR equation only correlates the physico-chemical properties of molecules with the percent uptake in parasitized RBCs. It does not necessarily correlate these properties with the parasiticidal activity of this class of compounds. For example, analogue **9** found to have the highest percent uptake in parasitized RBCs, however, it lacks peroxide bridge, a key structural feature, essential for antimalarial activity of this class of compounds[24] which further question its antimalarial potency. The lack of peroxide bridge in derivative **9** might decrease overall polar contribution of this compound (nHBAcc descriptor value of 4) and be responsible for increase in percent uptake into parasitized RBCs. Also, the fourteen carbon chain (C14) at C-9 position of compound 9 may increases the log *P* value of compound **9** which might lead to higher penetration of this compound into RBC membrane. The significance of this QSAR model exists in the sense that exposures of these schizonticidal-staged antimalarials to the parasites occur in RBCs and uptake in RBCs is an essential step for such compounds.

To further validate the developed model, we calculated similar physico-chemical parameters for another sesquiterpene lactone and highly specific inhibitor of sarco/endoplasmic reticulum Ca^{2+}-ATPase (SERCA), thapsigargin (unpublished results), which also lacks an endoperoxide bridge (see Fig. 5). Both artemisinin

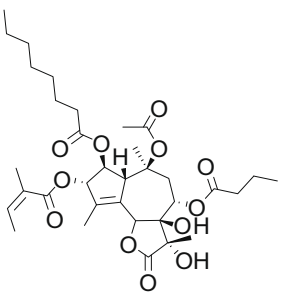

Figure 5. Structure of thapsigargin.

and thapsigargin were thought to target the SERCA of *P. falciparum*.[25] Eckstein-Ludwig et al. explained poor penetration of thapsigargin in to the infected erythrocytes by in vivo labeling studies with the fluorescent derivative of thapsigargin (BODIPY–thapsigargin).[25] To support their experimental observation, we predicted the percent uptake of thapsigargin using developed MLR model. The predicted percent uptake of thapsigargin according to this model was found to be 23.23% which supports the experimental observation made by the authors.

4. Conclusion

A simple, rapid and sensitive analytical method for quantifying artemisinin and its derivatives in non-parasitized or parasitized RBCs was developed and validated. This method combined one-step SPE with ESI^+-SIM detection to achieve the selective and specific quantification of artemisinin and its derivatives. Artemisinin and its derivatives showed higher uptake by parasitized RBCs than non-parasitized RBCs. A MLR model was successfully generated for predicting the percent uptake of artemisinin and its derivatives into parasitized RBCs. The developed model provided useful structure property information to predict the uptake based on the selected descriptors: nHBAcc which illustrated the H-bond accepting capacity of molecules, DPSA-3 which combined the shape and electronic properties of molecules for the prediction of uptake and Wlambda2.unity which emphasized that elongation of molecule at C-9 position favored the percent uptake into parasitized RBCs. This QSAR model is useful in a context that the first step of exposure of these schizonticidal compounds to parasites take place in the RBC compartment and uptake of these compounds is key step for their antimalarial activity. However, the resultant model does not necessarily predict the antimalarial potency of these compounds. The developed model successfully predicted the poor percent uptake of another sesquiterpene lactone, thepsigargin, in agreement to the experimental observation. Currently, this model is useful to predict the uptake of new artemisinin derivatives prior to synthesis with sufficient accuracy.

Acknowledgment

We appreciate the contributions of Mr. John Trott, NCNPR at the University of Mississippi for providing parasitized RBCs.

References and notes

1. UNDP/World Bank/WHO special programme for research training in tropical diseases. Tropical diseases. *Progress in research*, **1989–1990**, 29.
2. Klayman, D. L. *Science* **1985**, *228*, 1049.
3. White, N. J. *Trans. R. Soc. Trop. Med. Hyg.* **1994**, *88*, S3.
4. Price, R.; Van, V. M.; Phaipun, L.; Luxemburger, C.; Simpson, J.; McGready, R.; ter Kuile, F.; Kham, A.; Chongsuphajaisiddhi, T.; White, N. J.; Nosten, F. *Am. J. Trop. Med. Hyg.* **1999**, *60*, 547.
5. Gu, H. M.; Warhurst, D. C.; Peters, W. *Trans. R. Soc. Trop. Med. Hyg.* **1984**, *78*, 265.
6. Meshnick, S. R.; Thomas, A.; Ranz, A.; Xu, C. M.; Pan, H. Z. *Mol. Biochem. Parasitol.* **1991**, *49*, 181.
7. Kamchonwongpaisan, S.; Chandrangam, G.; Avery, M. A.; Yuthavong, Y. *J. Clin. Invest.* **1994**, *93*, 467.
8. Vyas, N.; Avery, B. A.; Avery, M. A.; Wyandt, C. M. *Antimicrob. Agents Chemother.* **2002**, *46*, 105.
9. Darlington, R. B. *Regression and Linear Models*; McGraw-Hill: New York, NY, 1990.
10. Steinbeck, C.; Han, Y.; Kuhn, S.; Horlacher, O.; Luttmann, E.; Willighagen, E. J. Chem. Inf. Comput. Sci.. *43* **2003**, 493.
11. Zhao, S. S. *Analyst* **1987**, *112*, 661.
12. Thomas, C. G.; Ward, S. A.; Edwards, G. *J. Chromatogr.* **1992**, *583*, 131.
13. Green, M. D.; Mount, D. L.; Todd, G. D.; Capomacchia, A. C. *J. Chromatogr.* **1995**, *695*, 237.
14. Navaratnam, V.; Mansor, S. M.; Chin, L. K.; Mordi, M. N.; Asokan, M.; Nair, N. K. *J. Chromatogr. Sect., B: Biomed. Sci. Appl.* **1995**, *669*, 289.
15. Avery, B. A.; Venkatesh, K. K.; Avery, M. A. *J. Chromatogr. Sect., B: Biomed. Sci. Appl.* **1999**, *730*, 71.
16. Asawamahasakda, W.; Ittarat, I.; Pu, Y. M.; Ziffer, H.; Meshnick, S. R. *Antimicrob. Agents Chemother.* **1994**, *38*, 1854.
17. Klayman, D. L. *Science* **1985**, *228*, 1049.
18. Asawamahasakda, W.; Benakis, A.; Meshnick, S. R. *J. Lab. Clin. Med.* **1994**, *123*, 757.
19. Cabantchik, Z. I. *Blood Cells* **1990**, *16*, 421.
20. Ginsburg, H.; Kutner, S.; Krugliak, M.; Cabantchik, Z. I. *Mol. Biochem. Parasitol.* **1985**, *14*, 313.
21. Guha, R.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1440.
22. Stanton, D.; Jurs, P. C. *Anal. Chem.* **1990**, *62*, 2323.
23. Todeschini, R.; Gramatica, P. *Quant. Struct.-Act. Relat.* **1997**, *16*, 113.
24. Haynes, R. K. *Curr. Opin. Infect .Dis.* **2001**, *14*, 719.
25. Eckstein-Ludwig, U.; Webb, R. J.; Van Goetham, I. D. A.; East, J. M.; Lee, A. G.; Kimura, M.; O'Neill, P. M.; Bray, P. G.; Ward, S. A.; Krishna, S. *Nature* **2003**, *424*, 957.