Structure−Activity Relationships of the Antimalarial Agent Artemisinin. 6. The Development of Predictive In Vitro Potency Models Using CoMFA and HQSAR Methodologies

Mitchell A. Avery,*[,†,‡] Maria Alvim-Gaston,[†,§] Carlos R. Rodrigues,[∥] Eliezer J. Barreiro,[∥] Fred E. Cohen,[⊥] Yogesh A. Sabnis,[†] and John R. Woolfrey[†,∞]

Department of Medicinal Chemistry, School of Pharmacy, Thad Cochran National Center for Natural Products Research, and Department of Chemistry, University of Mississippi, University, Mississippi 38677, Laboratório de Avaliação e Síntese de Substâncias Bioativas (LASSBio), Departamento de Fármacos, Faculdade de Farmácia, UFRJ, RJ, Brazil, 21944-910, and Department of Cellular Molecular Pharmacology, University of California, San Francisco, California 94143-0446

Received January 17, 2001

Artemisinin (**1**) is a unique sesquiterpene peroxide occurring as a constituent of *Artemisia annua* L. Because of the effectiveness of Artemisinin in the treatment of drug-resistant *Plasmodium falciparum* and its rapid clearance of cerebral malaria, development of clinically useful semisynthetic drugs for severe and complicated malaria (artemether, artesunate) was prompt. However, recent reports of fatal neurotoxicity in animals with dihydroartemisinin derivatives such as artemether have spawned a renewed effort to develop nontoxic analogues of artemisinin. In our effort to develop more potent, less neurotoxic agents for the oral treatment of drug-resistant malaria, we utilized comparative molecular field analysis (CoMFA) and hologram QSAR (HQSAR), beginning with a series of 211 artemisinin analogues with known in vitro antimalarial activity. CoMFA models were based on two conformational hypotheses: (a) that the X-ray structure of artemisinin represents the bioactive shape of the molecule or (b) that the hemin-docked conformation is the bioactive form of the drug. In addition, we examined the effect of inclusion or exclusion of racemates in the partial least squares (pls) analysis. Databases derived from the original 211 were split into chiral ($n = 157$), achiral ($n = 34$), and mixed databases ($n = 191$) after leaving out a test set of 20 compounds. HQSAR and CoMFA models were compared in terms of their potential to generate robust QSAR models. The r^2 and q^2 (cross-validated r^2) were used to assess the statistical quality of our models. Another statistical parameter, the ratio of the standard error to the activity range (*s*/AR), was also generated. CoMFA and HQSAR models were developed having statistically excellent properties, which also possessed good predictive ability for test set compounds. The best model was obtained when racemates were excluded from QSAR analysis. Thus, CoMFA of the $n = 157$ database gave excellent predictions with outstanding statistical properties. HQSAR did an outstanding job in statistical analysis and also handled predictions well.

Introduction

The effectiveness of artemisinin (**1**) and its derivatives as antimalarial drugs for the treatment of multi-drug-resistant *P. falciparum* has received considerable attention in recent years.[1−6] More often than not, the focus of these studies has been to demonstrate antimalarial efficacy in vitro for new structural classes[7−11] or modification of the natural product architecture.[12−15] A rational approach for the discovery of a pharmaceutically acceptable, economically viable, peroxide-based antimalarial awaits development of a global mechanism of action model for organic peroxides[16−24] and/or a predictive quantitative structure−activity relationship (QSAR) based pharmacophore model.[25−29] With the advent of parallel synthesis methods and technology, we might expect the number of antimalarial artemisinin analogues to be tested to grow dramatically. Combinatorial methods could also be envisioned as a semirational approach to this above discovery strategy. One method of orchestrating these strategies is to make use of quantitative structure−activity relationship (QSAR) models for the rapid prediction and virtual prescreening of antimalarial activity.[30] Our efforts to furnish new iterations of QSAR in this area have been chronicled,[3,29−33] leading to highly bioactive, nonneurotoxic[29] derivatives of artemisinin.[34]

Traditional QSAR studies have been used since the early 1970s to predict activities of untested molecules. Currently, 3D-QSAR techniques such as comparative molecular field analysis (CoMFA)[35,36] have been employed to build QSAR models for a wide range of applications.[37−39] CoMFA analysis involves the alignment of molecules in a structurally and pharmacologically reasonable manner on the basis of the assumption that each compound acts via a common binding site. In

* To whom correspondence should be addressed. Phone (662)-915-588. Fax (662)-915-5638. E-mail: mavery@olemiss.edu.
† Department of Medicinal Chemistry, University of Mississippi.
‡ Department of Chemistry, University of Mississippi.
§ Present address: Eli Lilly and Co., Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, IN 46285.
∥ Departamento de Fármacos, UFRJ.
⊥ Department of Cellular Molecular Pharmacology, University of California.
∞ Present address: COR Therapeutics Inc., 256 E.Grand Ave., South San Francisco, CA 94080.

10.1021/jm0100234 CCC: $22.00
Published on Web 12/08/2001

this method, we should have some knowledge or hypothesis regarding active conformations of the molecules under study as a prerequisite for structural alignment.

Development of a 3D-QSAR model representing the in vitro antimalarial activity of a database of artemisinin analogues can be a complex process. The in vitro potency of a drug is dependent on structural features that affect the sequence of events leading to accumulation of drug at its site of action as well as the mode of action itself. Peroxidic antimalarial activity might be disrupted by unpredictable metabolic events,[40−43] premature deactivation by iron(II) salts affected by its accumulation into the cell,[23,44−49] or altered by the molecular details of radical formation.[19,24] However, by judicious selection of analogues and careful model construction (alignment, conformational predictions), a 3D-QSAR study can lead to a predictive model providing useful SAR information. Using CoMFA, we have developed artemisinin pharmacophores based on two distinct conformational hypotheses. The determination of the "active" conformation that each compound will retain is a critical issue. Because of the lack of structural data supporting a specific "active" conformation, we might assume that a compound is active in an energetically minimized conformation. In addition, since an X-ray structure of the parent molecule is available,[50] this can serve as a template upon which other analogues can be overlaid.

We constructed the second model taking in consideration Meshnick's mechanistic hypothesis involving complexation of artemisinin to hemin, leading to the generation of bioactive radicals.[16] Here, the docking interaction with hemin alters the analogue conformation, and the final structure included in the creation of the pharmacophore represents the result of a force field calculation incorporating the influence of the hemin molecule.[51]

We selected over 100 analogues from our laboratory and a number from the literature to assemble the 211 artemisinin analogues for the database.[3] A test set of 20 molecules was removed from the original 211 that evenly spanned the antimalarial activity range, as well as the structure diversity of our database. The database was separated into a training set (191 compounds) and a test set (20 compounds) to determine the external predictivity of the resulting model. Further studies were carried out to examine the effect of chirality on the analysis. A database with only chiral compounds was produced from the original 191 by omitting racemates, where $n = 157$ (chiral only), and the test set size was 15. The pooled racemates were also subjected to QSAR analysis with a training set of $n = 34$ (racemates) and test set size of 5.

Hologram quantitative structure−activity relationships (HQSAR), a 2D-QSAR protocol (Tripos Associates, Inc.), eliminates the need for determination of 3D structure, putative binding conformations, and molecular alignment. In HQSAR, each molecule in the database is broken down into a series of unique structural fragments, which are arranged to form a molecular hologram.[52] Unlike other fragment-based fingerprinting methods, HQSAR encodes all possible molecular fragments (linear, branched, and overlapping). Supplementary 3D information, such as hybridization and chirality, may also be encoded in the molecular holograms. HQSAR models have been compared to other QSAR techniques, such as CoMFA, and have demonstrated similar predictive capability. The HQSAR program highlights substructural features in sets of molecules that are relevant to biological activity.

In this study, well-characterized training sets of 191, 157, or 34 artemisinin analogues with in vitro antimalarial activity were examined utilizing CoMFA and HQSAR to generate predictive QSAR models, providing a useful guideline for the design of more potent new artemisinin analogues.

Biological Data

A total of 211 artemisinin analogues were collected in which several different ring systems were represented. All of the analogues were either peroxides or trioxanes (tables are available as Supporting Information), which should act via similar mechanisms of action,[3,7,14,15,53−66] or their corresponding biologically inert reduced oxides (e.g. 1-deoxyartemisinin). Another important criterion was that the selected compounds must have been tested using the same assay method, i.e., in vitro against the chloroquine-resistant, mefloquine-sensitive *P. falciparum* W-2 clone,[67] and have a reported control activity for artemisinin. This approach is mandatory for a bioassay method that utilizes parasitized red blood cells in which there can be interday and interlaboratory variation in the IC_{50} for artemisinin. Variation is intrinsic to this in vitro assay system in which exact degrees of parasitemia of the red blood cells are not readily controllable from day to day. Therefore, utilization of relative activity (RA) for quantitative comparisons was essential. The RA was calculated from the experimentally derived control IC_{50} of artemisinin (IC_{50} values in ng/mL) divided by the IC_{50} of the analogue and corrected for molecular weight (eq 1).

$$\log \mathrm{RA} = \log[(IC_{50}\ \text{of artemisinin}/IC_{50}\ \text{of the analog}) \times (\text{MW of the analog}/\text{MW of the artemisinin})] \quad (1)$$

Before inclusion into the spreadsheets, the RA was converted to the log RA. In some cases to be discussed below, the RA was adjusted for racemates in the database for which it had been assumed that only one enantiomer was bioactive. In these cases, the RA was multiplied by 2 before taking the log.

QSAR Method

CoMFA. We created a 3D database (Sybyl 6.5, Tripos Associates), using the reported crystal structure for artemisinin as a template.[50] To best describe the aromatic side chains of some analogues, as well as the aliphatic artemisinin backbone, Gasteiger−Hückel charges were calculated for each of the compounds. The structures were minimized by applying the Tripos force field, producing structures close to that of the artemisinin X-ray structure. Because the majority of the compounds were tetracyclic peroxides such as artemisinin itself, little if any difference from the X-ray structure could be found upon minimization (gas phase). Side chains at C-9 and C-3 were able to achieve extended (staggered) conformations upon repeated minimizations. Manual realignment of side chains was conducted,

always leading to higher energy local minima. Structures with more flexible side chains were minimized manually, followed by Monte Carlo search methods (Macromodel v5), confirming that global minima had been reached for the tetracyclic analogues. For more flexible peroxides included in the tables, such as Table 4 (*seco*-artemisinins), extensive NMR, modeling, and crystallographic work had already been done to establish a plausible conformation for inclusion into the database.[65,68] The other flexible systems included in the database had crystallographic structure determinations that agreed with solution NMR and modeling data.[69,11]

In the development of the 3D-QSAR models, choice of analogue conformation was important in providing a realistic pharmacophore. The active conformation of each analogue might be approximated if a crystal structure of artemisinin interacting with hemin were available. Since acquisition of these data is not feasible because hemin reacts rapidly with artemisinin,[70] different approaches to select the "active" conformation were used.

First, if the interaction of a peroxidic drug with hemin is a relatively rapid and structurally insensitive step in a potentially complex sequence of events, we could ignore its contribution to define the potency of this class of molecules. In this case, it may be possible to use the E_{min} (minimum energy) structures to provide a realistic 3D-QSAR. Alternatively, by modeling the manner in which artemisinin interacts with hemin, a conformation can be predicted in which interaction between flexible side chains and hemin are minimized and focus is placed on association of the peroxide bridge with hemin. In this instance, the absolute energy of the peroxidic ligand might be slightly higher than the lowest energy conformation determined by molecular mechanics methods.

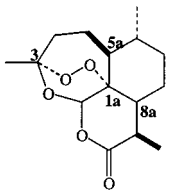
Current studies point toward the interaction of the peroxy moiety with the hemin iron(II) core as a major component of the antimalarial mode of action. This artemisinin−hemin association has been modeled,[51] and it has been proposed that activation of artemisinin proceeds through a radical intermediate occurring at a point near the minimum intermolecular energy of interaction.[71−76] In an attempt to create a model that might reflect the active conformation of each analogue more accurately, we used the Sybyl Dock protocol to approximate analogue−hemin binding. Dock calculates the fields of nonbonded electrostatic and Lennard-Jones steric interaction between points on a lattice (by default 0.25 Å spacing) and the docking site. Associating each atom of the ligand with the nearest lattice point and summing the fields over all ligand atoms approximate the ligand−site interaction.

Initially, an energetically favorable interaction was achieved by manual manipulation of the ligand with respect to hemin. Consideration of a large number of artemisinin analogues required an assumption that the peroxide docks to the hemin iron in the manner identified by Shukla et al.[51] Docking was complete when the minimal combined electrostatic and steric docking energies were found.

To take into consideration the effect the metalloporphyrin molecule plays in the active conformation of the ligand, the ligand structural conformation was minimized with respect to hemin using the Tripos force field. In the minimization process, a small amount of translation occurred, the ligand arriving at a conformational minimum when a rms gradient energy difference of ≤ 0.05 kcal/mol was achieved. Ligands within the database were realigned on the basis that a common orientation of the peroxy group with the Fe(II) core of hemin is likely a mechanistic bottleneck and would, therefore, be more important than a minor differences in gas-phase docking energies. Figure 1 shows the overlap of the compounds.

In some cases, substituents at C-9 and C-3 had a significant impact on the docking interaction to an extent where the critically important contribution of the peroxide bridge interacting with the hemin iron atom was diminished. To reduce the significance of nonperoxide bridge interaction with hemin, analogue substituents at C-3 or C-9 were rotated to a "local minima" energy conformation so that the substituent did not play the primary role in the docking interaction. The torsional angle was modified in the docking process from dihedral C10−C9−C16−C17 = 174° to give dihedral C10−C9−C16−C17 = 58° (Figure 2). The "local minima" are within energy differences reasonably attainable under the test conditions. Although some analogues are not in their lowest energy conformation as calculated in gas-phase mode, we believe that this process leads to useful binding information, potentially mimicking the actual active conformation.

All 211 minimized analogues were aligned with the assumption that the core ring structure and peroxide bridge are most meaningful with respect to biological activity. Though a number of flexible and static alignments were tried, the best model was obtained by fitting atoms C-1a, C-3, C-5a, and C-8a.



The deoxyartemisinin analogues included in our database do not contain the peroxide bridge needed for antimalarial activity.[55] Therefore, the best fit was accomplished using the atoms that hold the peroxide or oxide bridge in close alignment, atoms C-1a and C-3. Overlap of the 211 molecules can be seen in Figure 1.

HQSAR. Molecular fingerprinting (hologram) is a method that represents a compound as a unique string of numbers or "bins". The bins represent all of the unique fragments included within a particular molecule and are assigned by a cyclic redundancy check (CRC) algorithm (Figure 3). Thus, the hologram includes information on the quantity and type of each fragment in each molecule. The hashing of more than one type of fragment into the same bin can sometimes lead to poor correlation statistics in the pls analysis, and for this reason, a number of holograms with differing bin lengths are created for each molecule. The hologram length that leads to the best pls analysis is used in development of the pharmacophore.

A number of parameters may be adjusted to optimize the HQSAR model. By increasing fragment size, the

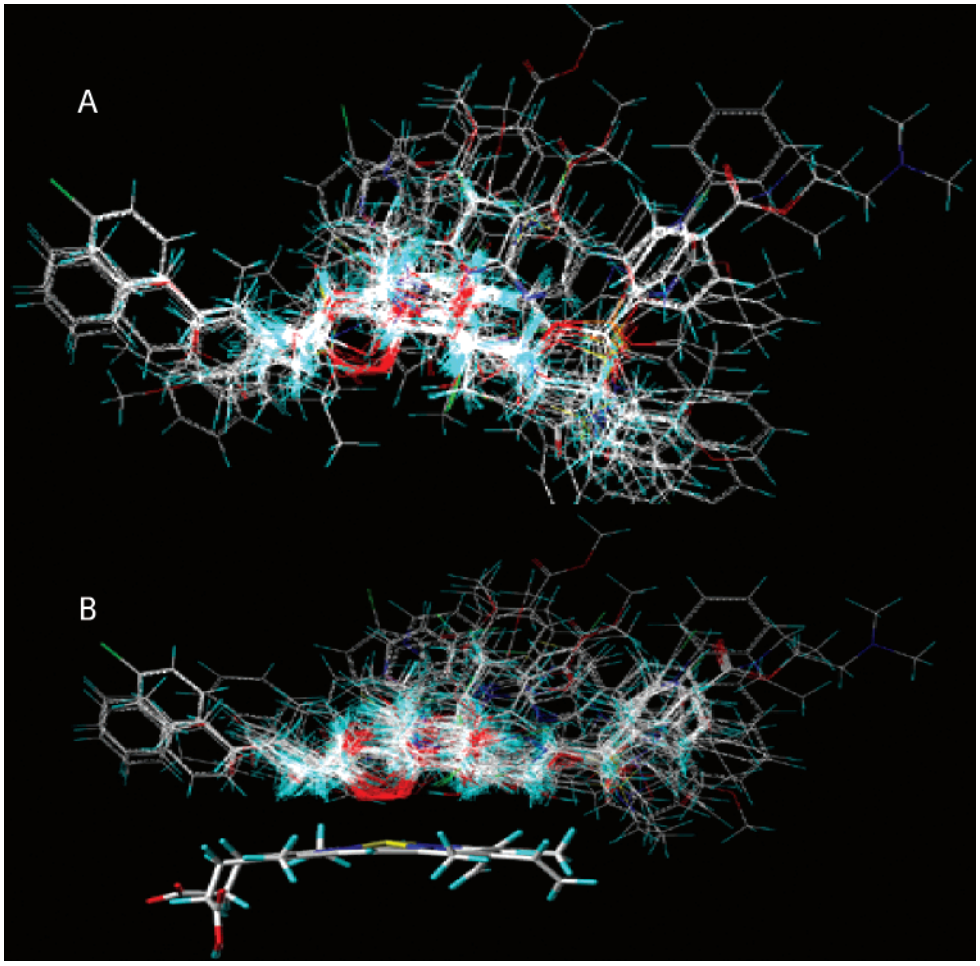


Figure 1. Alignment of 211 artemisinin analogues used in the CoMFA model development: (A) standard alignment; (B) alignment based on proposed docking conformation.

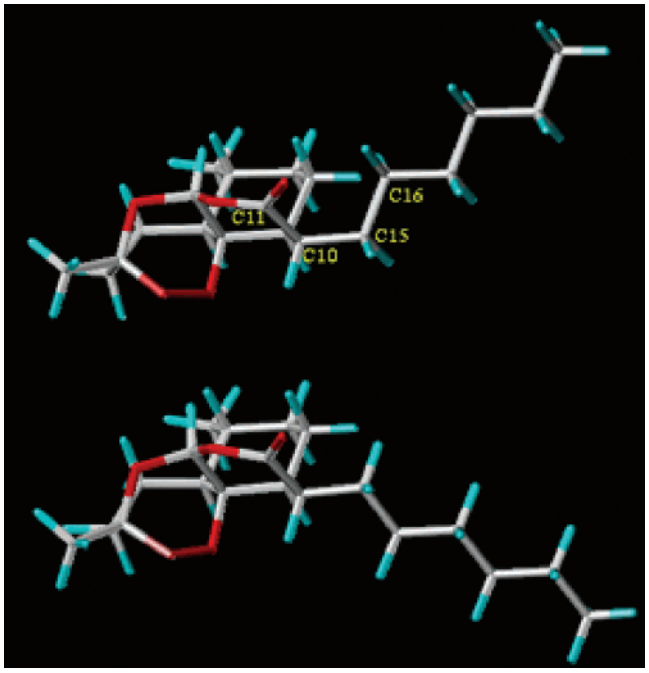


Figure 2. Torsional angle modified in the docking process.

number of fragments may be reduced and a pseudo-distance parameter created, though we also concurrently limit the number of similar fragments found in each analogue. Allowing consideration of distinct atom types, bond types, atom hybridization, and stereocenters, as well as observing hydrogens, which are, by default, not considered, can augment the nature of the fragments assigned by the CRC.

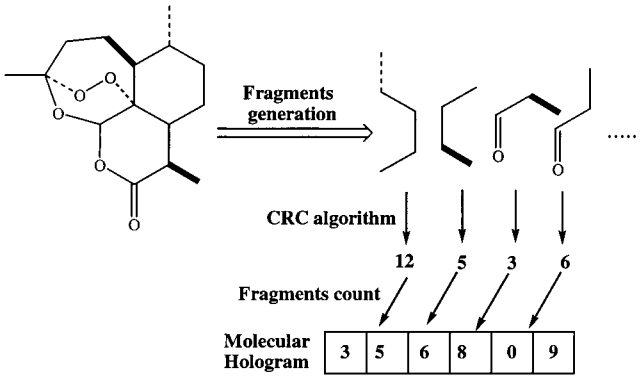


Figure 3. Generation of the molecular hologram.

QSAR Descriptors

Calculation of CoMFA Descriptors. CoMFA involves the measurement of the electrostatic and steric fields around a template molecule and the relationship of these measurements to the molecule's biological activity. These measurements are taken at regular intervals throughout the lattice using a probe atom of designated size and charge. In our study, lattice spacing

Table 1. Statistical Results of CoMFA pls Analyses

entry	no. of compds	probe[a]	q^2	r^2	S	F	optimum component
1	191	2 Å/C.3	0.72	0.88	0.53	169.85	6
2	191	2 Å/H	0.72	0.82	0.54	162.18	4
3	191 (dock)	2 Å/C.3	0.72	0.85	0.60	171.45	6
4	191 (chiral)	2 Å/C.3	0.72	0.88	0.53	171.35	4
5	157[b]	2 Å/C.3	0.73	0.96	0.35	315.63	9
6	157	2 Å/H	0.70	0.95	0.39	249.85	4
7	157 (dock)	2 Å/C.3	0.68	0.95	0.37	278.80	8
8	34 (racemic)	2 Å/C.3	0.24	0.92	0.32	54.76	1
9	34 (racemic)	2 Å/H	0.26	0.91	0.35	45.23	1

[a] Lattice spacing (in angstroms), probe atom type (sp^3-C or proton). [b] The 191 database without 34 racemates.

was 2 Å and either an sp^3 carbon or hydrogen with a probe charge of +1 served as the probe atom.

Calculation of HQSAR Descriptors. The molecular hologram representation (HQSAR) package in Sybyl 6.5 was used, converting molecules in the database into counts of their constituent fragments. All linear, branched, and overlapping substructure fragments in the size range four to seven atoms were generated for each molecule. The exact model produced by HQSAR is dependent on the hologram length (53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353, 401 bins), but we also considered the information contained in the generated fragments. The particular nature of substructure fragments generated by HQSAR and, consequently, the information contained in the resultant molecular holograms can be altered through the settings of some parameters.

In our study, 191 artemisinin analogues were assigned fragments by the CRC protocol using the default fragment size and fragment distinction parameters. Various pls analyses were performed using a variety of fragment size and fragment distinction selections.

3D-QSAR Analyses

Multivariate statistical analysis using pls produced predictive QSAR models of the database to correlate variation of in vitro antimalarial activity with variation in the descriptors described in the previous sections. The predictive power of the model was determined by using "leave-one-out" (LOO) cross-validation and by the use of a test set of 20 structurally and biologically diverse artemisinin analogues excluded from the model creation (Table 1). A cross-validated r^2 (q^2), obtained as a result of this analysis, served as a quantitative measure of the predictive ability of the final QSAR models. The q^2 value is a statistical indication of how well a model can predict the activity of members left out of the model formation. In contrast, the conventional r^2 is simply a reflection of how well the fit equation reproduces input values.[39] The selection of the "best" QSAR model was chosen on the basis of a combination of q^2, r^2, predictive-r^2 (pr^2), and s. With this approach, the optimal model would approach 3 as s approached 0. This number could be viewed as $[(q^2 + r^2 + pr^2) - s]$, in which the perfect model would equal 3.0 (Table 8). Another aspect of measuring model goodness-of-fit is to examine the ratio of the standard error to activity range (s/AR). This ratio should generally be <10% for good QSAR models.

The standard and docked ($n = 191$) aligned databases were analyzed using CoMFA/pls. A number of analogues in the database were racemates (34), though the database itself was constructed using the enantiomer having the absolute configuration corresponding to (+)-artemisinin (**1**). Analysis using CoMFA/pls of the nonracemic compounds ($n = 157$) was done in an effort to find a model that might be able to predict the activities of each individual enantiomer, taking into consideration that only the total activity of the mixture is experimentally available.

Table 2. Influence of Various Fragment-type Parameters for the 191 Compound Database Using Fragment Size Default (4−7)

statistical parameters	q^2	r^2	S	s/AR	optimum components
none[a]	0.69	0.79	0.70	9.90	7
Con[b]	0.73	0.87	0.57	9.50	9
Con−H[c]	0.65	0.77	0.74	8.10	5
Con−H−Chi[d]	0.76	0.89	0.51	9.40	10

[a] For all cases, the atom and bond flags are turned on. [b] Connectivity flag is on. [c] Hydrogen and Connectivity are on. [d] Chirality option is combined with Con−H.

Table 3. Influence of Various Fragment Size Parameters for the 191 Compound Database Using Atom/Bond/Connectivity/Hydrogen/Chirality as Fragment Type

fragment length	q^2	r^2	S	s/AR	optimum components
2−5	0.73	0.61	0.80	9.40	6
3−6	0.66	0.77	0.73	10.40	6
5−8	0.67	0.76	0.74	9.00	6
6−9	0.68	0.80	0.68	8.80	5
7−10	0.69	0.77	0.74	8.80	4

Considering structural features of artemisinin analogues, the first HQSAR model was generated by use of various fragment-type parameters: atoms, bonds, connections, hydrogens, and chirality. Table 3 summarizes the results from the pls analyses associated with the 191 database, as well as the various fragment parameters used to improve the statistical analysis. The best fragment type, fragment length, and optimum number of the components from pls analyses were chosen on the basis of the highest q^2.

Results and Discussion

As shown in Table 1 (entry 1), the pls analysis of the $n = 191$ (2 Å/C.3) standard alignment database indicated a moderate r^2 (0.88) value, good predictive q^2 (0.72), and a moderate standard error (0.53). Use of the proton probe atom gave nearly identical results (entry 2).

In the literature, many peroxidic racemates (34 compounds) have been tested for in vitro antimalarial activity. The artemisinin-like enantiomer of each pair was included in the $n = 191$ database under the assumption that this was the active isomer. In an attempt to better understand the contribution of each enantiomer, these compounds were removed from the database, and a third model ($n = 191-34$ racemates $= 157$) was developed. The notable predictivity of the model was indicated by a high q^2 (0.73) value for the default probe atom and lattice size (entry 5). Judging by the standard error of 0.35, the lowest among the valid models, r^2 (0.96), and F test values (315.63), the model reproduced the input data somewhat better than the $n = 191$ precursor. The proton probe atom did not give substantially different results from the default probe (entry 6).

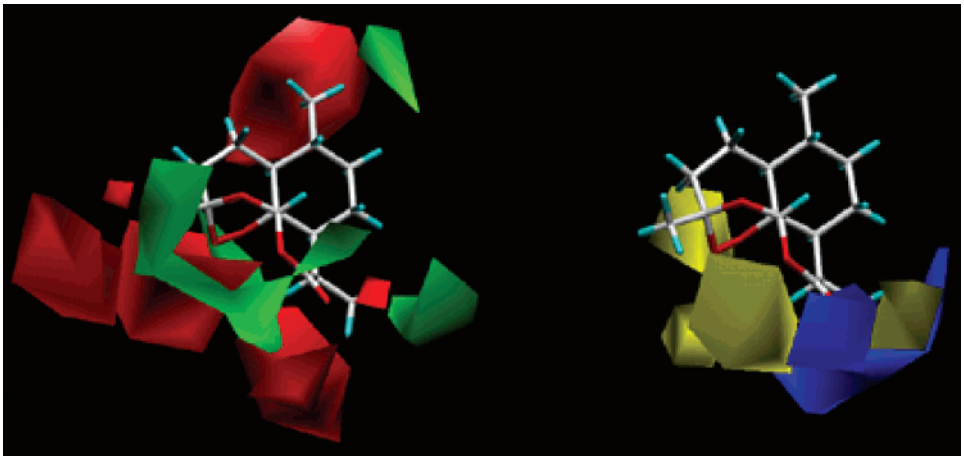


Figure 4. CoMFA contour maps about (+)-artemisinin for the standard alignment database (n = 191, 2 Å/C.3). In the steric contour map to the left, green contours indicate areas where steric bulk is predicted to increase antimalarial activity, while red contours indicate regions where steric bulk is predicted to decrease activity. The electrostatic contour map on the right displays yellow polyhedra where partial negative charge is correlated with antimalarial activity; the blue polyhedra indicating a relationship between partial positive charge and activity.

Prediction of the in vitro activity for the racemic pair of compounds leads to a less clear interpretation of the data. Is only one enantiomer active? If so, should the biological activity value (RA) be "*doubled*", assuming only half of the assayed material is active? If the enantiomers are unequally active, how should the data be treated? The model in which the RAs were doubled before taking the log [or 191 (Chiral), entry 4] was also used to predict the activities of the test set, providing some insight into the answers to these questions.

To aid in visualization, Figure 4 shows the electrostatic and steric maps for the pls analysis with n = 191, 2 Å/C.3. In the steric map, red contours correspond to regions in space where steric bulk would be predicted to decrease antimalarial activity. The red regions appear around the endoperoxide and other areas on the α-face of the molecule, suggesting that increased steric bulk on the underside of the molecule would be detrimental to the activity. Conversely, green contours represent areas around the template molecule, where an increase in antimalarial activity due to increased steric bulk is anticipated.

In Figure 4, green polyhedra surround the regions near the C-9 methyl, C-3 methyl, and the 7β-H of artemisinin, suggesting that greater steric bulk in these areas would increase activity. It is interesting to note how an active analogue, for example, compound **70** (9-butyl), fits within the contours compared to the natural product artemisinin or an inactive analogue. Also, an increase in lipophilicity resulting from removal of the carbonyl at C-10 is generally favorable, as seen in Table 3. Compound **70** (log RA = 1.32) is simply the reduced product of the less active lactone **8** (log RA = 0.17).

Within the CoMFA electrostatic map (Figure 4), yellow contours are displayed in areas where partial negative charge is associated with increased activity of the database analogues. Yellow contours are visible near the peroxide bridge, supporting the important role the peroxide plays in activity, and near O-11 (or N-11). Blue contours indicate areas within the lattice where electropositive properties of a molecule are predicted to increase activity. These regions include the partial positive charges associated with hydrogen atoms bound to carbon, and can be correlated with lipophilic interaction. A broad band of blue extends from the C-9 methyl toward the carbonyl at C-10. Since no analogues bearing electropositive groups in this area of the dataset were included in the analysis, the blue contour may be an accidental byproduct of the dataset.

The CoMFA protocol was repeated for the n = 191 dock-aligned database, and pls analysis indicated that the model (entry 3) was comparable to that found with the standard alignment. Overall, the dock alignment led to a model with comparable r^2 (0.85) and F-test (171.45) but higher standard error (0.60).

Contour maps for the n = 191 2 Å/C.3 dock model (Table 1), representing electrostatic and steric volumes deemed important to antimalarial activity, can be seen in Figure 5. Some major visual differences are apparent between the dock and standard models. For example, in the dock alignment rule models, the contribution of the yellow contours, displayed in areas where negative charge is associated with increased activity, is diminished (Figure 5) relative to the standard alignment model (Figure 4). Perhaps this difference can be explained by the possibility that groups that interfere with hemin binding are negatively correlated with potency. In the dock models, where the side chains have been artificially moved out of the region occupied by hemin, the yellow contours are much smaller.

In the steric map, green contours indicate areas where steric bulk is predicted to increase antimalarial activity (Figure 5). The green contours in the dock model are much stronger than those in the standard models and are incidentally along the periphery of the artemisinin skeleton corresponding to the C-9, N-11, and C-3 substituent space.

There are also variations in the electrostatic contours between the docked and standard models. For example, in the standard model, yellow contours appear near the C-6 methyl but are absent in the docked model. This point is supported by the increase in activity from arteether (**60**) to the more negative charged difluoro analogue **181**. In addition, in the electrostatic CoMFA maps (Figure 4), the yellow contours surrounding the peroxide spanned slightly different areas of the trioxane/

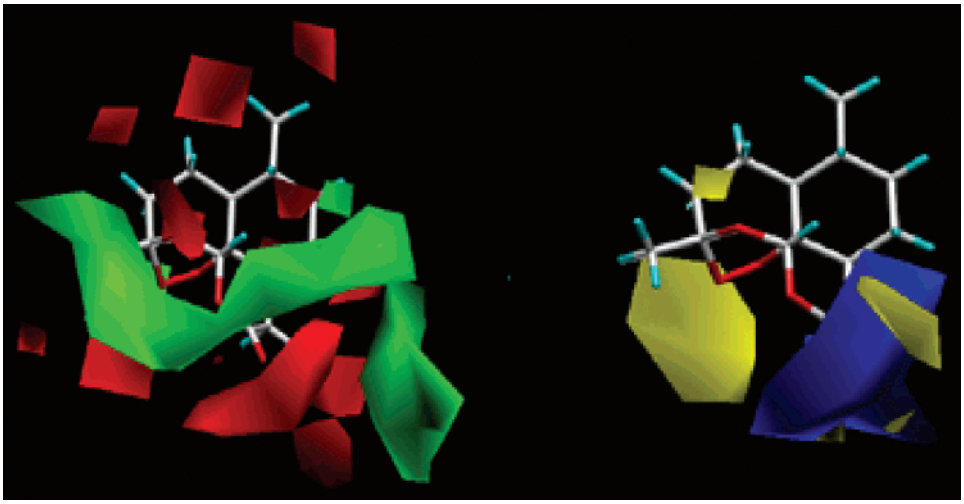


Figure 5. CoMFA contour maps about (+)-artemisinin for the dock aligned database (n = 191, 2 Å/C.3). In the steric contour map to the left, green contours indicate areas where steric bulk is predicted to increase antimalarial activity, while red contours indicate regions where steric bulk is predicted to decrease activity. The electrostatic contour map on the right displays yellow polyhedra where partial negative charge is correlated with antimalarial activity; the blue polyhedra indicating a relationship between partial positive charge and activity.

Table 4. Influence of Fragment-type Parameters on the r^2, q^2, Optimum Number of Components, Standard Error of Estimate (s) after pls Analysis for 157 Database Using Fragment Size Default (4−7), and Atom/Bond/Connectivity/Hydrogen/Chirality as Fragment Type

fragment length (fragment type)	q^2	r^2	S	optimum components
4−7 (Con−H−Chi)	0.72	0.86	0.62	6

pyran oxygen atoms. Nonetheless, the yellow contour just below the peroxide, assumes a greater magnitude than the standard model. Further, the exact location of the contours about the peroxy group is somewhat different in the standard model, with the contours tracking the location of the lone pair electrons of the peroxide (Figure 5 versus Figure 4).

In the first HQSAR approach, we have used default values for the fragment type and fragment size variables (4−7). The best model was found when the "atoms", "bonds", "hydrogen", "chirality", and "connectivity" options were used as fragment types, with 10 components being optimum. The pls analyses for HQSAR are summarized in the Table 2.

The best model using Con−H−Chi (q^2 = 0.76) had default fragment lengths (4−7). This model was used to investigate if different fragment lengths could improve the statistical parameters. Fragment size parameters control the minimum and maximum length of fragments to be included in the hologram fingerprint. The HQSAR results for the different fragment sizes are summarized in Table 3.

Choice of the default value for fragment lengths above led fortuitously to the better statistical result as a general erosion of statistical quality was seen with other fragment lengths in Table 3.

HQSAR was also applied to the n = 157 database (chiral model). The new HQSAR model was built using atom/bond/connectivity/hydrogen/chirality as fragment type and 4−7 as fragment size (Table 4). In this final case, we see little loss in fidelity between the 191 and the 157 datasets for q^2 or r^2 and only a slight difference in s value (∼0.1).

Validation of the 3D-QSAR Models

CoMFA Validation. The q^2 value of a CoMFA model, together with other statistical information from the pls analysis, provides information on the predictive capability of the model. In this study we have generated CoMFA models that describe the pharmacophore either with or without the involvement of hemin, both of which provide good q^2 values. Selection of the model that most accurately depicts reality is not trivial, since many variables are inherent in the cell-culture bioassay results. However, it may be possible to distinguish between models by the appropriate choice of test molecules. If one model were more consistent in predicting the activities of novel compounds, then perhaps it would be more realistic and therefore preferable over the other model in drug design efforts.

The log RA predicted by CoMFA (standard alignment) and HQSAR are highly consistent with the experimental data and in good agreement with each other. Though the CoMFA (dock alignment) model leads to a pharmacophore with good predictive statistical results (acceptable r^2 and q^2), it is less accurate in predicting the activities of the 20 compound test-set. Further, there does not exist sufficient data regarding the biological activity for each enantiomer in the racemic analogues used. To help define the antimalarial contribution made by each enantiomer, a model was developed in which 34 racemates were placed as single (+)-artemisinin-like enantiomers in the original database of 191 compounds. By doing this it was implied that the opposite enantiomers are as active as their counterpart, and an optically pure compound is not necessary for antimalarial activity. These "mixed chiral and achiral" models have been discussed as 191 (2 Å/C.3) and 191 dock. Databases without the racemates, or "chiral" models 157 (2 Å/C.3) and 157 dock, were also included in our study. Finally, adjustment to an "achiral" model was attempted in which the relative activities of the racemates in the 191 (2 Å/C.3) model were multiplied by 2 (before taking the log), leading to model 191 (chiral) (2 Å/C.3). Finally, all of the 39 racemates were excised from the 211 database and split into a n = 34 training set with a test

Table 5. Prediction Values for the Test Set in the QSAR Models

		predicted log RA							
compd no.	log RA	157 HQSAR	191 CoMFA (2 Å/C.3)	191 CoMFA (2 Å/H)	191 CoMFA (dock)	191 CoMFA (chiral)[a]	157 CoMFA (2 Å/C.3)	157 CoMFA (2 Å/H)	157 CoMFA (dock)
9	−0.32	0.58	0.62	0.35	−0.29	0.52	0.92	0.75	1.10
14	0.86	0.20	0.49	0.17	0.47	0.44	0.23	0.62	−0.10
40	0.78	0.86	0.60	0.51	0.81	0.64	1.03	0.88	1.15
52	−4.00	−4.44	−3.47	−4.30	−4.12	−3.44	−3.16	−2.80	−2.56
67	−0.77	−0.05	0.12	−0.05	0.23	0.07	0.80	0.08	1.12
81	1.79	0.99	0.83	0.43	0.79	0.78	0.73	0.49	0.65
119	−0.12	−1.13	0.60	0.06	0.19	0.58	0.42	0.02	−0.04
124	0.19	−0.83	0.66	0.36	0.69	0.92	0.18	0.13	0.25
129	−0.04	−0.08	−0.09	0.30	0.34	−0.11	0.00	−0.13	0.00
141	1.47	0.75	0.59	0.37	−0.63	0.63	0.92	0.79	0.57
143	−4.00	−3.23	−2.62	−3.08	−2.72	−2.68	−3.02	−3.60	−3.67
180	−1.24	−1.24	1.15	0.62	0.91	0.37	−0.66	−0.91	−0.24
186	−4.00	−4.09	−4.21	−3.27	−3.85	−4.10	−4.23	−4.26	−3.78
199	0.31	0.58	0.86	0.21	0.76	0.85	1.29	−0.01	−1.38
204	1.07	1.54	0.30	0.25	0.42	0.47	1.13	1.44	0.43
112	−0.97		−1.38	−1.30	−1.27	−1.08			
115	−0.56		0.19	−0.08	0.42	0.50			
148	0.95		0.65	0.83	0.91	0.97			
160	−0.29		0.05	−0.05	0.09	0.34			
166	0.62		0.38	0.35	0.52	0.67			

[a] The RA for the racemates in the model was multiplied by 2.

Table 6. Predictions of CoMFA Test Set ($n = 5$) log RA for Racemic Database ($n = 34$)

compd no.	log RA	34 racemic (2 Å/C.3)	34 racemic (2 Å/H)
112	−1.27	−1.63	−1.81
115	−0.86	−1.33	−0.60
148	0.65	0.58	0.79
160	−0.59	−0.06	−0.25
166	0.32	0.54	0.39

set of 5. In each case, data was subjected to the statistical pls analysis (Table 1). Relative antimalarial activities were predicted for the test set using these databases, as shown in Tables 5 and 6.

The prediction r^2 was obtained by simple regression analysis of the test set log RA values versus the predicted values for each model, as shown in Table 7. Comparing the performance of the models in predicting the test set data (Table 5) indicated the following four models: the chiral CoMFA model 191 (2 Å/C.3), which includes corrections for the racemates, with an r^2 of 0.84; both chiral CoMFA models 157 (H and C.3 probe), with $r^2 = 0.86$ and 0.89; and finally, the HQSAR model, which, with $r^2 = 0.89$, reflected experimentally derived antimalarial activities best. It is interesting to note that the best HQSAR model incorporated chirality coding (Con−H−Chi).

The exceptionally poor q^2 value for the racemic dataset (Table 1, entries 8, 9) was easily explained by examining the residual values for the pls analysis. There were several instances of residuals with values over +2 or −2. How this CoMFA model did so well in prediction is a matter of coincidence. It is clear that the model probably should not be used for predictions with the low q^2 value at the outset. This poor performance of racemates should be a clear warning to others carrying out QSAR studies that one cannot mix racemates with chiral compounds. If it is know unequivocally that only one of the enantiomers of the racemate is responsible for the compound's activity, then provision can be made to adjust the activity value accordingly.

It is tempting to conclude that chirality does seem to be important for antimalarial activity in this class of drugs. While this is inconsistent with isolated reports demonstrating that the in vitro antimalarial effects of selected enantiomers gave the same potencies,[19,20,77] it is also consistent with in vitro studies involving extensive numbers of enantiomeric pairs of trioxanes.[78] While the chirality issue is currently unresolved, one can envision that changes in metabolism and distribution from enantiomer to enantiomer would lead to unpredictable and variable activities for enantiomeric pairs. In order for this model to be sensitive to these complex issues, it will be necessary for numerous enantiomeric pairs to be added to the database, and an alignment rule will need to be devised.

As we can see, in the contour map generated from the "dock-aligned" database, interaction with the hemin iron is a predominant component of the dock-minimized model. It follows that the predictivity of the 157-docked model might favor compounds in which this interaction is influenced by the dock-minimization process.

HQSAR Validation. Because the structure and molecular properties of a molecule can be encoded within a 2D fingerprint (including biological activity), we should be able to predict the activity of a molecule from its fingerprint. To visualize this principle, Figure 6 shows the most important fragments of artemisinin (**1**) labeled by the results of the final pls analysis.

Twenty compounds were excluded from the database to serve as test compounds to evaluate the predictive

Table 7. Standard Error in Model Predictions of Test Set (r^2)[a]

model	HQSAR	191 (2 Å/C.3)	191 (2 Å/H)	191 (dock)	191 (chiral)	157 (2 Å/C.3)	157 (2 Å/H)	157 (dock)	34 (2 Å/C.3)	34 (2 Å/H)
r^2	0.89	0.77	0.81	0.74	0.84	0.86	0.89	0.72	0.87	0.90

[a] Regression analysis of actual log RA vs predicted log RA.

Table 8. Model Selection Based on $[(r^2 + q^2 + pr^2) - s]$[a]

model	HQSAR	191 (2 Å/C.3)	191 (2 Å/H)	191 (dock)	191 (chiral)	157 (2 Å/C.3)	157 (2 Å/H)	157 (dock)	34 (2 Å/C.3)	34 (2 Å/H)
$[(r^2 + q^2 + pr^2) - s]$	2.03	1.84	1.81	1.71	1.91	2.20	2.15	1.98	1.71	1.72

[a] r^2, q^2, and s from Table 1; pr^2 from Table 7.

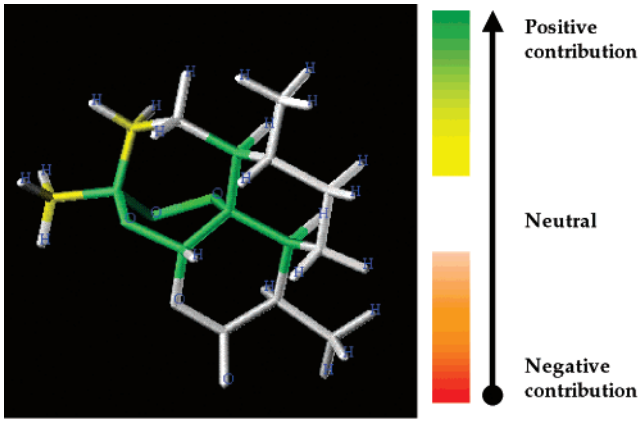


Figure 6. The HQSAR contribution map for artemisinin.

ability of the present QSAR models. Table 5 summarizes a comparison of the experimental log RA values and the predicted log RA values following the CoMFA and HQSAR results.

Conclusions

In CoMFA, a suitable sampling of the steric and electrostatic fields around a set of aligned structures provides the information required for the development of a predictive pharmacophore model. Although many CoMFA models are statistically excellent and offer good predictive performance, they are inherently limited by the need to align the database molecules correctly within 3D space. Because experimental evidence about ligand−receptor binding conformations is frequently lacking, the bioactive conformation must be postulated on the basis of information about the receptor binding site and/or the common conformational space accessible to different known ligands. The global minimum-energy conformation is commonly selected and a considerable amount of time and expertise is required for molecular modeling. This latter process, if performed correctly, can yield useful information about the nature of the receptor site and receptor−ligand interaction. Unfortunately, statistically sound but misleading models can often be constructed from the same datasets using other conformational hypotheses. The resulting models can still be predictive of a test set, but they are not suited to drug-design, because of a false conformational hypothesis.

On the other hand, HQSAR requires only information about the 2D molecular structure, requiring little or no molecular modeling for development of an alignment rule. Generations of descriptors using CoMFA involves time-consuming processes, such as structural alignment, that often can be carried out effectively only by laborious "trial and error" based attempts. In contrast, the generation of molecular holograms as the chemical descriptors in HQSAR takes considerably less time and expertise. It is worthwhile to mention that construction of the regression equation through standard pls analysis takes less time in HQSAR than CoMFA, because the number of descriptors generated is generally far less. Due to dependence of a CoMFA model on molecular conformation and structural alignment, in which small perturbations can become magnified in the final QSAR model, much care must be taken when generating these models to ensure reproducibility. Because the calculation of HQSAR descriptors from counts of substructural molecular fragments is straightforward, model reproducibility is readily achieved in minimal time. On the basis of the information derived from these QSAR models, we have attempted to determine the properties of the molecules in this study that are responsible for antimalarial activity.

The computer aided pharmacophore modeling was used to better define the complex structure−activity relationships of artemisinin and its analogues as antimalarial compounds. Accordingly, with a database of 191 artemisinin analogues, we developed various CoMFA pharmacophore models based on two different hypotheses. First, the bioactive conformation of the analogues corresponds to the minimum energy structures, and second, the bioactive conformation of the analogues is based on their interaction with the proposed molecular target in the parasite, hemin. Also we have examined the assumptions that enantiomers in a racemate have identical activity or those enantiomers have different activities.

In some models, racemates have been included or excluded from the analysis, while in other studies, the analogues are docked to hemin prior to being placed in the molecular database or the analogues are simply minimized and aligned to the template artemisinin. The results initially indicate that the chiral model ($n = 157$) may be best but that the effect of docking on improved performance of these models was not seen. This finding does not invalidate Meshnick's original hypothesis regarding parasitic hemin being the mechanistic trigger for the action of the drug but merely suggests that it is a fast insensitive step late in the overall cascade of events leading to accumulation of drug at the site of heme in the food vacuole. The issue of how peroxide chirality effects activity is worth mention. According to Boukouvalas,[78] for an extensive collection of trioxanes, in some cases enantiomeric pairs of trioxanes have equivalent potency, but more often than not, the pairs have quite different in vitro potencies. Other researchers have concluded that chirality is not a requirement for potency and that both enantiomers of a set should have equivalent antimalarial activity. Regardless of these opposing results, it seems intuitively obvious that the kinetics (and even outcome) of metabolic processes should vary from one enantiomer or diastereomer to the other. Unpredictable and variable metabolism occurring in whole cell assay should be expected to lead to unpredictable antimalarial potencies for enantiomers or diastereomer pairs.

We believe that the best solution to this issue is to construct enantiomeric pairs for a diverse set of peroxides and trioxanes in order to seed the QSARs. If the

hemin hypothesis were correct, the resulting model would reflect hemin docking, differences in metabolism, lipophilicity, and topology along with other factors defining antimalarial potency, such as drug accumulation into the parasitized erythrocyte.[79]

In terms of validation of such a model, a measure of internal consistency is available in the form of the q^2 value. However, the ultimate test of a model is its ability to predict activities for newly reported compounds. The hemin-docked artemisinin pharmacophore is quite complex, as indicated by a predictive character inconsistent with previous trials with these database. On one hand we have a pharmacophore which, judging by its notable statistical data, adequately predicts the activity of compounds, which encompass a range of activities and structural variation. However, neither of the docked models performed as well as the undocked models.

All of the models emerging from CoMFA ($n = 157$, no racemates) and HQSAR ($n = 191$, mixture) were of utilizable quality (high r^2) and exhibited good predictive ability (good q^2). In particular, the chiral database having no racemates ($n = 157$, Table 1, entry 5) provided the best overall statistical and predictive properties with a score of 2.20 (Table 8). Predictions made with CoMFA and HQSAR models on test set compounds were in reasonable agreement with the experimentally determined values. The CoMFA model reveals regions in 3D space around these artemisinin analogues that are important to antimalarial activity.

Acknowledgment. Funding for this work was provided by the National Institute of Allergy and Infectious Diseases (NIAID) and the UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases (TDR). We thank CAPES (Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for its fellowship to C.R.R.

Supporting Information Available: Tables of relative antimalarial activity for the database of 211 compounds is provided. Requests for the mdb files for $n = 191$ (training) and $n = 20$ (test) are available from the corresponding author by e-mail. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Liu, C.; Wang, Y.; Ouyang, F.; Ye, H.; Li, G. Advances in artemisinin research. *Huaxue Jinzhan* **1999**, *11*, 41−48.
(2) Bhattacharya, A. K.; Sharma, R. P. Recent developments on the chemistry and biological activity of artemisinin and related antimalarials−an update. *Heterocycles* **1999**, *51*, 1681−1745.
(3) Avery, M. A.; Alvim-Gaston, M.; Woolfrey, J. R. Synthesis and structure−activity relationships of peroxidic antimalarials based on artemisinin. *Adv. Med. Chem.* **1999**, *4*, 125−217.
(4) Dai, L.-X.; Chen, Y.-Q. Recent major advances in the studies on qinghaosu and related antimalarial agents. *Chemtracts* **1999**, *12*, 687−694.
(5) Haynes, R. K.; Vonwiller, S. C. From Qinghao, Marvelous Herb of Antiquity, to the Antimalarial Trioxane Qinghaosu. Some Remarkable New Chemistry. *Acc. Chem. Res.* **1997**, *30*, 73−79.
(6) Vroman, J. A.; Alvim-Gaston, M.; Avery, M. A. Current progress in the chemistry, medicinal chemistry and drug design of artemisinin based antimalarials. *Curr. Pharm. Des.* **1999**, *5*, 101−138.
(7) Posner, G. H.; O'Dowd, H.; Caferro, T.; Cumming, J. N.; Ploypradith, P.; Xie, S.; Shapiro, T. A. Antimalarial sulfone trioxanes. *Tetrahedron Lett.* **1998**, *39*, 2273−2276.
(8) Posner, G. H.; Oh, C. H.; Gerena, L.; Milhous, W. K. Synthesis and antimalarial activities of structurally simplified 1,2,4-trioxanes related to artemisinin. *Heteroat. Chem.* **1995**, *6*, 105−116.
(9) Dong, Y.; Matile, H.; Chollet, J.; Kaminsky, R.; Wood, J. K.; Vennerstrom, J. L. Synthesis and antimalarial activity of 11 dispiro-1,2,4,5-tetraoxane analogues of WR 148999. 7,8,15,16-Tetraoxadispiro[5.2.5.2]hexadecanes substituted at the 1 and 10 positions with unsaturated and polar functional groups. *J. Med. Chem.* **1999**, *42*, 1477−1480.
(10) Jefford, C. W.; Misra, D.; Rossier, J. C.; Kamalaprija, P.; Burger, U.; Mareda, J.; Bernardinelli, G.; Peters, W.; Robinson, B. L. Cyclopenteno-1,2,4-trioxanes as effective antimalarial surrogates of artemisinin. *Perspect. Med. Chem.* **1993**, 459−472.
(11) Haraldson, C. A.; Karle, J. M.; Freeman, S. G.; Duvadie, R. K.; Avery, M. A. The synthesis of 8,8-disubstituted tricyclic analogues of artemisinin. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2357−2362.
(12) O'Dowd, H.; Ploypradith, P.; Xie, S.; Shapiro, T. A.; Posner, G. H. Antimalarial artemisinin analogues. Synthesis via chemoselective C−C bond formation and preliminary biological evaluation. *Tetrahedron* **1999**, *55*, 3625−3636.
(13) Mekonnen, B.; Ziffer, H. A new route to N-substituted 11-azaartemisinins. *Tetrahedron Lett.* **1997**, *38*, 731−734.
(14) Lin, A. J.; Zikry, A. B.; Kyle, D. E. Antimalarial Activity of New Dihydroartemisinin Derivatives. 7. 4-(p-Substituted phenyl)-4(R or S)-[10(α or β)-dihydroartemisininoxy]butyric Acids. *J. Med. Chem.* **1997**, *40*, 1396−1400.
(15) Avery, M. A.; Mehrotra, S.; Johnson, T. L.; Bonk, J. D.; Vroman, J. A.; Miller, R. Structure−Activity Relationships of the Antimalarial Agent Artemisinin. 5. Analogs of 10-Deoxoartemisinin Substituted at C-3 and C-9. *J. Med. Chem.* **1996**, *39*, 4149−4155.
(16) Meshnick, S. R. The mode of action of antimalarial endoperoxides. *Trans. R. Soc. Trop. Med. Hyg.* **1994**, *88*, 31−32.
(17) Robert, A.; Meunier, B. Is alkylation the main mechanism of action of the antimalarial drug artemisinin? *Chem. Soc. Rev.* **1998**, *27*, 273−274.
(18) Kamchonwongpaisan, S.; Meshnick, S. R. The mode of action of the antimalarial artemisinin and its derivatives. *Gen. Pharmacol.* **1996**, *27*, 587−592.
(19) Cumming, J. N.; Ploypradith, P.; Posner, G. H. Antimalarial activity of artemisinin (qinghaosu) and related trioxanes: mechanism(s) of action. *Adv. Pharmacol. (San Diego)* **1997**, *37*, 253−297.
(20) Jefford, C. W. Peroxidic antimalarials. *Adv. Drug Res.* **1997**, *29*, 271−325.
(21) Pandey, A. V.; Tekwani, B. L.; Singh, R. L.; Chauhan, V. S. Artemisinin, an endoperoxide antimalarial, disrupts the hemoglobin catabolism and heme detoxification systems in malarial parasite. *J. Biol. Chem.* **1999**, *274*, 19383−19388.
(22) Posner, G. H.; Park, S. B.; Gonzalez, L.; Wang, D.; Cumming, J. N.; Klinedinst, D.; Shapiro, T. A.; Bachi, M. D. Evidence for the Importance of High-Valent Fe: O and of a Diketone in the Molecular Mechanism of Action of Antimalarial Trioxane Analogues of Artemisinin. *J. Am. Chem. Soc.* **1996**, *118*, 3537−3538.
(23) Posner, G. H.; Cumming, J. N.; Ploypradith, P.; Oh, C. H. Evidence for Fe(IV):O in the Molecular Mechanism of Action of the Trioxane Antimalarial Artemisinin. *J. Am. Chem. Soc.* **1995**, *117*, 5885−5886.
(24) Avery, M. A.; Fan, P.; Karle, J. M.; Bonk, J. D.; Miller, R.; Goins, D. K. Structure−Activity Relationships of the Antimalarial Agent Artemisinin. 3. Total Synthesis of (+)-13-Carbaartemisinin and Related Tetra- and Tricyclic Structures. *J. Med. Chem.* **1996**, *39*, 1885−1897.
(25) Wu, J.; Chen, K.; Ji, R. QSAR (quantitative structure−activity relationships) study on Qinghaosu (artemisinin) derivatives. *Prog. Nat. Sci.* **1991**, *1*, 72−74.
(26) Tang, Y.; Jiang, H. L.; Chen, K. X.; Ji, R. Y. QSAR study of artemisinin (Qinghaosu) derivatives using neural network method. *Indian J. Chem. Sect. B: Org. Chem. Incl. Med. Chem.* **1996**, *35B*, 325−332.
(27) Jiang, H.-L.; Chen, K.-X.; Wang, H.-W.; Tang, Y.; Chen, J.-Z.; Ji, R.-Y. 3D-QSAR study on ether and ester analogues of artemisinin with comparative molecular field analysis. *Zhongguo Yaoli Xuebao* **1994**, *15*, 481−487.
(28) Nguyen-Cong, V.; Van Dang, G.; Rode, B. M. Using multivariate adaptive regression splines to QSAR studies of dihydroartemisinin derivatives. *Eur. J. Med. Chem.* **1996**, *31*, 797−803.
(29) Avery, M. A.; McLean, G.; Edwards, G.; Ager, A. Structure−activity relationships of peroxide-based artemisinin antimalarials. *Biol. Act. Nat. Prod.* **2000**, 121−132.
(30) Lundstedt, T.; Clementi, S.; Cruciani, G.; Pastor, M.; Kattaneh, N.; Andersson, P.; Linusson, A.; Sjostrom, M.; Wold, S.; Norden, B. Chapter 12. Intellegent Combinatorial Libraries. *Computer-Assisted Lead Finding and Optimization*; VHCA/Wiley-VCH: Basel, 1997.
(31) Woolfrey, J. R.; Rodrigues, C. R.; Alvim-Gaston, M.; Barreiro, E. J.; Cohen, F. E.; Avery, M. A. Hologram QSAR (HQSAR) of artemisinin analogues. *Book of Abstracts*; 217th ACS National Meeting, Anaheim, CA, March 21−25, 1999, COMP-137.

(32) Alvim-Gaston, M.; Woolfrey, J. R.; Avery, M. A. 3D-QSAR of peroxidic antimalarials based on artemisinin. *Book of Abstracts;* 217th ACS National Meeting, Anaheim, CA, March 21−25, 1999, COMP-138.

(33) Woolfrey, J. R.; Avery, M. A.; Doweyko, A. M. Comparison of 3D quantitative structure−activity relationship methods: analysis of the in vitro antimalarial activity of 154 artemisinin analogues by hypothetical active-site lattice and comparative molecular field analysis. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 165−181.

(34) Meshnick, S. R.; Jefford, C. W.; Posner, G. H.; Avery, M. A.; Peters, W. Second-generation antimalarial endoperoxides. *Parasitol. Today* **1996**, *12*, 79−82.

(35) Cramer III, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). I. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(36) Cramer III, R. D.; Bunce, J. D.; Patterson, D. E.; Frank, I. E. Crossvalidation, Bootstrapping, and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies. *Quant. Struct. Act. Relat. Pharmacol., Chem. Biol.* **1988**, *7*, 18−25.

(37) Doweyko, A. M. Predictive 3D-pharmacophores developed from HASL models. *Book of Abstracts;* 213th ACS National Meeting, San Francisco, CA, April 13−17, 1997, COMP-306.

(38) Tonmunphean, S.; Kokpol, S.; Parasuk, V.; Wolschann, P.; Winger, R. H.; Liedl, K. R.; Rode, B. M. Comparative molecular field analysis of artemisinin derivatives: ab initio versus semiempirical optimized structures. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 397−409.

(39) Kubinyi, H. 3D-QSAR in Drug Design: Theory, Methods and Applications; ESCOM: Leiden, 1993.

(40) Tregova, A.; Holme, A. D.; Bell, W. G.; McLean, W. G.; Edwards, G.; Ward, S. A. The influence of hepatic metabolism on the neurotoxicity of artemisinin derivatives on cultured NB2a neuroblastoma cells. *Br. J. Pharmacol.* **1996**, *119*, 336 P.

(41) Svensson, U. S. H.; Ashton, M. Identification of the human cytochrome P450 enzymes involved in the in vitro metabolism of artemisinin. *Br. J. Clin. Pharmacol.* **1999**, *48*, 528−535.

(42) Leskovac, V.; Theoharides, A. D. Hepatic metabolism of artemisinin drugs−I. Drug metabolism in rat liver microsomes. *Comp. Biochem. Physiol., C: Comput. Pharmacol. Toxicol.* **1991**, *99C*, 383−390.

(43) Grace, J. M.; Skanchy, D. J.; Aguilar, A. J. Metabolism of artelinic acid to dihydroqinghaosu by human liver cytochrome P4503A. *Xenobiotica* **1999**, *29*, 703−717.

(44) Bloodworth, A. J.; Shah, A. Iron(II)-mediated rearrangement of 1,2,4-trioxanes into 1,2-diol monoesters via 1,5-hydrogen transfer. *Tetrahedron Lett.* **1995**, *36*, 7551−7554.

(45) Smith, S. L.; Maggs, J. L.; Edwards, G.; Ward, S. A.; Park, B. K.; McLean, W. G. The role of iron in neurotoxicity: A study of novel antimalarial drugs. *Neurotoxicology* **1998**, *19*, 557−559.

(46) Wu, Y.-L.; Chen, H.-B.; Jiang, K.; Li, Y.; Shan, F.; Wang, D.-Y.; Wang, Y.-F.; Wu, W.-M.; Wu, Y.; Yao, Z.-J.; Yue, Z.-Y.; Zhou, C.-M. Interaction of biomolecules with qinghaosu (artemisinin) and its derivatives in the presence of ferrous ion-an exploration of antimalarial mechanism. *Pure Appl. Chem.* **1999**, *71*, 1139−1142.

(47) Haynes, R. K.; Vonwiller, S. C. The behavior of qinghaosu (Artemisinin) in the presence of non-heme iron(II) and (III). *Tetrahedron Lett.* **1996**, *37*, 257−260.

(48) Hawley, S. R.; Bray, P. G.; Mungthin, M.; Atkinson, J. D.; O'Neill, P. M.; Ward, S. A. Relationship between antimalarial drug activity, accumulation, and inhibition of heme polymerization in Plasmodium falciparum in vitro. *Antimicrob. Agents Chemother.* **1998**, *42*, 682−686.

(49) Vattanaviboon, P.; Wilairat, P.; Yuthavong, Y. Binding of dihydroartemisinin to hemoglobin H: Role in drug accumulation and host-induced antimalarial ineffectiveness of α-thalassemic erythrocytes. *Mol. Pharmacol.* **1998**, *53*, 492−496.

(50) Lisgarten, J. N.; Potter, B. S.; Bantuzeko, C.; Palmer, R. A. Structure, absolute configuration, and conformation of the antimalarial compound, Artemisinin. *J. Chem. Crystallogr.* **1998**, *28*, 539−543.

(51) Shukla, K. L.; Gund, T. M.; Meshnick, S. R. Molecular modeling studies of the artemisinin (qinghaosu)−hemin interaction: docking between the antimalarial agent and its putative receptor. *J. Mol. Graphics* **1995**, *13*, 215−222.

(52) Heritage, T. W.; Lowis, D. R. Molecular Hologram QSAR. *Rational Drug Design. Novel Methodology and Practical Applications*; American Chemical Society: Washington, DC, 1999; pp 212−225.

(53) Klayman, D. L. Qinghaosu (artemisinin): An antimalarial drug from China. *Science* **1985**, *228*, 1049−1055.

(54) Brossi, A.; Venugopalan, B.; Gerpe, L. D.; Buchs, P.; Luo, X. D.; Milhous, W.; Peters, W. Arteether, a new antimalarial drug: Synthesis and antimalarial properties. *J. Med. Chem.* **1988**, *31*, 645−650.

(55) Avery, M. A.; Gao, F.; Chong, W. K. M.; Mehrotra, S.; Milhous, W. K. Structure−activity relationships of the antimalarial agent artemisinin. 1. Synthesis and comparative molecular field analysis of C-9 analogues of artemisinin and 10-deoxoartemisinin. *J. Med. Chem.* **1993**, *36*, 4264−4275.

(56) Jung, M.; Li, X.; Bustos, D. A.; ElSohly, H. N.; McChesney, J. D.; Milhous, W. K. Synthesis and antimalarial activity of (+)-deoxoartemisinin. *J. Med. Chem.* **1990**, *33*, 1516−1518.

(57) Avery, M. A.; Bonk, J. D.; Chong, W. K. M.; Mehrotra, S.; Miller, R.; Milhous, W.; Goins, D. K.; Venkatesan, S.; Wyandt, C. Structure−Activity Relationships of the Antimalarial Agent Artemisinin. 2. Effect of Heteroatom Substitution at O-11: Synthesis and Bioassay of N−Alkyl-11-aza-9-desmethylartemisinins. *J. Med. Chem.* **1995**, *38*, 5038−5044.

(58) Lin, A. J.; Li, L. Q.; Klayman, D. L.; George, C. F.; Flippen-Anderson, J. L. Antimalarial activity of new water-soluble dihydroartemisinin derivatives. 3. Aromatic amine analogues. *J. Med. Chem.* **1990**, *33*, 2610−2614.

(59) Lin, A. J.; Klayman, D. L.; Milhous, W. K. Antimalarial activity of new water-soluble dihydroartemisinin derivatives. *J. Med. Chem.* **1987**, *30*, 2147−2150.

(60) Jefford, C. W.; Velarde, J. A.; Bernardinelli, G.; Bray, D. H.; Warhurst, D. C.; Milhous, W. K. Synthesis, structure, and antimalarial activity of tricyclic 1,2,4-trioxanes related to artemisinin. *Helv. Chim. Acta* **1993**, *76*, 2775−2788.

(61) Posner, G. H.; Wang, D.; Gonzalez, L.; Tao, X.; Cumming, J. N.; Klinedinst, D.; Shapiro, T. A. Mechanism-based design of simple, symmetrical, easily prepared, potent antimalarial endoperoxides. *Tetrahedron Lett.* **1996**, *37*, 815−818.

(62) Posner, G. H.; Oh, C. H.; Gerena, L.; Milhous, W. K. Extraordinarily potent antimalarial compounds: new, structurally simple, easily synthesized, tricyclic 1,2,4-trioxanes. *J. Med. Chem.* **1992**, *35*, 2459−2467.

(63) Pu, Y. M.; Torok, D. S.; Ziffer, H.; Pan, X.-Q.; Meshnick, S. R. Synthesis and Antimalarial Activities of Several Fluorinated Artemisinin Derivatives. *J. Med. Chem.* **1995**, *38*, 4120−4124.

(64) Hu, Y.; Ziffer, H.; Li, G.; Yeh, H. J. C. Microbial oxidation of the antimalarial drug arteether. *Bioorg. Chem.* **1992**, *20*, 148−154.

(65) Avery, M. A.; Gao, F.; Chong, W. K. M.; Hendrickson, T. F.; Inman, W. D.; Crews, P. Synthesis, conformational analysis, and antimalarial activity of tricyclic analogues of artemisinin. *Tetrahedron* **1994**, *50*, 957−972.

(66) Acton, N.; Karle, J. M.; Miller, R. E. Synthesis and antimalarial activity of some 9-substituted artemisinin derivatives. *J. Med. Chem.* **1993**, *36*, 2552−2557.

(67) Milhous, W. K.; Weatherley, N. F.; Bowdre, J. H.; Desjardins, R. E. In Vitro Activities of and Mechanisms of Resistance to Antifolate Antimalarial Drugs. *Antimicrob. Agents Chemother.* **1985**, *27*, 525−530.

(68) Avery, M. A.; Chong, W. K. M.; Bupp, J. E. Tricyclic analogue of artemisinin: synthesis and antimalarial activity of (+)-4,5-secoartemisinin and (−)-5-nor-4,5-secoartemisinin. *J. Chem. Soc., Chem. Commun.* **1990**, 1487−1489.

(69) Avery, M. A.; Chong, W. K. M.; Detre, G. Synthesis of (+)-8a,9-secoartemisinin and related analogs. *Tetrahedron Lett.* **1990**, *31*, 1799−1802.

(70) Paitayatat, S.; Tarnchompoo, B.; Thebtaranonth, Y.; Yuthavong, Y. Correlation of Antimalarial Activity of Artemisinin Derivatives with Binding Affinity with Ferroprotoporphyrin IX. *J. Med. Chem.* **1997**, *40*, 633−638.

(71) Gu, J.; Chen, K.; Jiang, H.; Leszczynski, J. A model molecule study of the O-centered and the C-centered free radical intermediates of artemisinin. *Theochem* **1999**, *491*, 57−66.

(72) Butler, A. R.; Gilbert, B. C.; Hulme, P.; Irvine, L. R.; Renton, L.; Whitwood, A. C. EPR evidence for the involvement of free radicals in the iron-catalyzed decomposition of qinghaosu (artemisinin) and some derivatives; antimalarial action of some polycyclic endoperoxides. *Free Radical Res.* **1998**, *28*, 471−476.

(73) Wu, W.-M.; Wu, Y.; Wu, Y.-L.; Yao, Z.-J.; Zhou, C.-M.; Li, Y.; Shan, F. Unified Mechanistic Framework for the Fe(II)-Induced Cleavage of Qinghaosu and Derivatives/Analogs. The First Spin-Trapping Evidence for the Previously Postulated Secondary C-4 Radical. *J. Am. Chem. Soc.* **1998**, *120*, 3316−3325.

(74) Wu, W.-M.; Wu, Y.; Wu, Y.-L.; Yao, Z.-J.; Zhou, C.-M.; Li, Y.; Shan, F. Unified Mechanistic Framework for the Fe(II)-Induced Cleavage of Qinghaosu and Derivatives/Analogs. The First Spin-Trapping Evidence for the Previously Postulated Secondary C-4 Radical. [Erratum to document cited in *Chem. Abstr. 128*, 230531]. *J. Am. Chem. Soc.* **1998**, *120*, 13002.

(75) Tonmunphean, S.; Irle, S.; Kokpol, S.; Vudhichai, P.; Wolschann, P. Ab initio and density functional study on singlet and triplet states of artemisinin. *Theochem* **1998**, *454*, 87−90.

(76) Meshnick, S. R.; Yang, Y. Z.; Lima, V.; Kuypers, F.; Kamchonwongpaisan, S.; Yuthavong, Y. Iron-dependent free radical generation from the antimalarial agent artemisinin (qinghaosu). *Antimicrob. Agents Chemother.* **1993**, *37*, 1108−1114.
(77) O'Neill, P. M.; Miller, A.; Bickley, J. F.; Scheinmann, F.; Oh, C. H.; Posner, G. H. Asymmetric syntheses of enantiomeric 3-p-fluorophenyl 1,2,4-trioxane analogues of the antimalarial artemisinin. *Tetrahedron Lett.* **1999**, *40*, 9133−9136.
(78) Boukouvalas, J. Personnal Communication, 2000.
(79) Vyas, N.; Avery, B. A.; Avery, M. A.; Wyandt, C. M. Carrier-Mediated Partitioning of Artemisinin into *Plasmodium falciparum*-Infected Erythrocytes. *Antimicrob. Agents Chemother.* In press.

JM0100234