Antimalarial Activity of Molecules Interfering with *Plasmodium falciparum* Phospholipid Metabolism. Structure−Activity Relationship Analysis[†]

Michèle Calas,*,[‡] Gérard Cordina,[‡] Jacques Bompart,[§] Mohamed Ben Bari,[‡,||] Taïb Jei,[||] Marie L. Ancelin,[⊥] and Henri Vial[⊥]

Laboratoire des Aminoacides, Peptides et Proteines, CNRS, UMR 5810, and CNRS, UMR 5539, CP 107,
Université de Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex 5, France, Laboratoire de Chimie Organique
Pharmaceutique, Faculté de Pharmacie, Université de Montpellier I, 15, Avenue C. Flahault, 34060 Montpellier Cedex, France,
and Département de Chimie, Faculté des Sciences Ben M'Sik, Casablanca, Maroc

A series of 80 compounds, primary, secondary, and tertiary amines and quaternary ammonium and bisammonium salts, most of them synthesized as potential choline or ethanolamine analogs, were tested against the *in vitro* growth of *Plasmodium falciparum*, the human malaria parasite. They were active over the $10^{-3}-10^{-8}$ M concentration range. A structure−activity relationship study was carried out using autocorrelation vectors as structural descriptors, and multidimensional analysis. Principal component analysis, ascending hierarchical classification, and stepwise discriminant analysis showed that both the size and shape of the molecule were essential for antimalarial potency, making the lipophilicity and electronegativity distribution in the molecular space essential. Using the autocorrelogram describing the molecular shape and the electronegativity distribution on the molecular graph, 98% of the molecules were correctly classified either as poorly active or active with only three explanatory variables. The most active compounds were quaternary ammoniums salts whose nitrogen atom had only one long lipophilic chain of 11 or 12 methylene groups (E5, E6, E10, E13, E20, E21, E22, E23, F4, F8), or the bisammoniums whose polar heads were linked by linear alkyl chains of 10 to 12 carbon atoms (G4, G23). The hydroxyethyl group of choline was not very beneficial, whereas the charge and substitutions of nitrogen (aimed at increasing lipophilicity) were essential for optimal interactions. A crude topographic model of the ligand (choline) binding site was thus drawn up.

Introduction

Malaria is the most widespread parasitic illness affecting intertropical zones, with around two billion sufferers, *i.e.*, slightly more than a third of the world population. *Plasmodium falciparum* is the parasite responsible for most malaria cases (80%), which often prove fatal. Resistance of *P. falciparum* to current antimalarials is now a serious problem throughout the world. The emergence of resistance is rapid for drugs possessing similar action mechanisms. Cross-resistance to currently used antimalarials that are not structurally related is much more worrysome since the range of antimalarials is far too small. New pharmacological models are thus needed to select molecules having an original mode of action.[1]

The asexual proliferation of *Plasmodium* within human erythrocytes is associated with the neosynthesis of a huge quantity of membranes, leading to an increase of approximately 600% in the erythrocytic phospholipid (PL) content. Around 85% of parasite PL consists of phosphatidylcholine (PC) and phosphatidylethanolamine (PE). The intracellular parasite has its own enzymatic machinery to synthesize the bulk of PL from the polar heads (mostly choline, ethanolamine and serine) and fatty acids drawn from the plasma.[2] On the other hand, the biosynthetic metabolism is completely missing in the mature mammalian erythrocytic host.[3]

Vial *et al.*[2] have identified the parasite metabolic pathways used for PL biosynthesis, a vital metabolic function. They notably demonstrated that one of the limiting steps in *de novo* PC biosynthesis is choline transport into the infected erythrocyte.[4] These studies have led to the development of a new pharmaceutical approach to the treatment of malaria involving interference with the phospholipidic metabolism by competition or substitution of PL polar head analogs.[5,6]

The present study compares the *in vitro* antimalarial activity of a quite large number of compounds analogous to ethanolamine or choline (most of them synthesized) against the human malarial parasite. Structure−activity relationships (SAR) determined by multidimensional methods, performed to optimize the antimalarial activities and to design new active and specific compounds, highlighted the essential features concerning the molecular requirements for antimalarial potency and the shape of the pharmacological target leading to a topographic model of the ligand binding site.

Studied Molecules

The 80 molecules used in the SAR are described in Tables 1, and their *in vitro* antimalarial activities against *P. falciparum* are also indicated. These molecules were classified into eight groups: compounds from series A (Table 1a), B (Table 1b), and C (Table 1c)

[†] Abbreviations: DF, discriminant function; HAC, hierarchical ascendant classification; PCA, principal component analysis; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PL, phospholipid; QSAR, quantitative structure−activity relationship; SDA, stepwise discriminant analysis.
  * Author to whom correspondence should be addressed.
[‡] Laboratoire des Aminoacides, Peptides et Proteines.
[§] Laboratoire de Chimie Organique Pharmaceutique.
[||] Departement de Chimie.
[⊥] CNRS, UMR 5539, CP 107.
[⊗] Abstract published in *Advance ACS Abstracts,* October 1, 1997.

Table 1. Structures and *in Vitro* Antimalarial Activities against *P. falciparum* $(IC_{50})^a$

a) Primary amines (Series A).

NAME	COMPOUNDS	IC_{50} (µM)
A0	$HO-(CH_2)_3-NH_2$	800^a
A1	$HO-(CH_2)_5-NH_2$	$>2000^a$
A2	$HS-CH_2-CH_2-NH_2$	200
A3	$HO-CH_2-CH(CH_3)-NH_2$	300^a
A4	$HO-CH(CH_3)-CH_2-NH_2$	800^a
A5	$(HO-CH_2)_2CH-NH_2$	50^a
A6	$HO-CH_2-CH(C_2H_5)-NH_2$	80^a
A7	$HO-CH_2-C(CH_3)_2-NH_2$	900
A8	$(HO-CH_2)_2C(C_2H_5)-NH_2$	600
A9	$H_2N-(CH_2)_5-NH_2$	2000
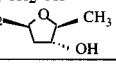
b) Secondary amines R_1- NH - R_2 (Series B).

NAME	R_1	R_2	IC_{50} (µM)
B1	$HO-CH_2-CH_2$	$HO-CH_2-CH_2$	1100
B2	$HO-CH_2-CH_2$	$C_6H_5-CH_2$	40
*B3	$HO-CH_2-CH_2$	$C_{12}H_{25}$	0.51
*B4	$HO-CH_2-CH(C_2H_5)$	$C_{12}H_{25}$	0.61
B5	$-CH_2-CH(OH)-CH_2-CH_2-$		>1000
B6	$-CH_2-CH(OH)-CH_2-CH_2-CH_2-$		>1000
*B7	$CH_3-O-CH_2-CH_2$	$C_{12}H_{25}$	2.1
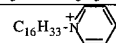
c) Tertiary amines R_1- N (R_2) - R_3 (Series C).

NAME	R_1	R_2	R_3	IC_{50}(µM)
C1	$HO-CH_2-CH_2$	$-(CH_2)_2-$		50^a
C2	$HO-CH_2-CH_2$	$-(CH_2)_4-$		420
C3	$HO-CH_2-CH_2$	$-(CH_2)_5-$		350
C5	$HO-CH_2-CH_2$	CH_3	CH_3	1300
C6	$HO-CH_2-CH_2$	C_2H_5	C_2H_5	400^a
C7	$HO-CH_2-CH_2$	$(CH_3)_2CH$	$(CH_3)_2CH$	450^a
*C8	$HO-CH_2-CH_2$	CH_3	$C_{12}H_{25}$	1.3
C9	$pCl\Phi OCH_2C(=O)O(CH_2)_2$	CH_3	CH_3	500
C4	$HO-CH_2-CH_2-N\underset{\smile}{\quad}N-CH_2-CH_2-OH$			1200
C12	$(CH_3)_2N - (CH_2)_6 - N(CH_3)_2$			130
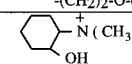
d) Series D compounds.

NAME	COMPOUNDS	IC_{50} (µM)
D1	$(CH_3)_3 N^+-CH_2-CH_2-Cl$, Cl^-	2000
D2	$(CH_3)_3 N^+-CH_2-CH_2-O-PO_3{}^{2-}$	1200^c
D3	$(CH_3)_3 N^+-CH_2-CO_2{}^-$	>8000
D4	$(CH_3)_3 N^+-CH_2-CH_2-O-CO-CH_3$, Cl^-	$>8000^c$
D5	$(CH_3)_3 C-CH_2-CH_2-OH$	>2000
D6	$(CH_3)_3N^+-CH_2\underset{\text{''}OH}{\overset{O}{\diagdown}}CH_3$, Cl^-	>200
D7	$(CH_3)_2 Se^+-CH_2-CH_2-OH$, Cl^-	420
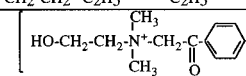
e) Quaternary ammonium salts: $R_2 - \underset{R_3}{\overset{R_1}{N^+}} - R_4$, X^- (Series E)

NAME	R_1	R_2	R_3	R_4	X^-	IC_{50} (µM)
E1	C_2H_5	C_2H_5	C_2H_5	C_2H_5	Br^-	700
*E2	CH_3	CH_3	CH_3	C_7H_{15}	Br^-	300
*E3	CH_3	CH_3	CH_3	C_9H_{19}	Br^-	1.6
E4	CH_3	CH_3	CH_3	$C_{10}H_{21}$	Br^-	0.7^b
*E5	CH_3	CH_3	CH_3	$C_{11}H_{23}$	Br^-	0.5
E6	CH_3	CH_3	CH_3	$C_{12}H_{25}$	Br^-	0.5^b
E7	CH_3	CH_3	CH_3	$C_{14}H_{29}$	Br^-	0.9^b
E8	CH_3	CH_3	CH_3	$C_{16}H_{33}$	Br^-	0.8^b
E9	CH_3	CH_3	CH_3	$C_{18}H_{37}$	Cl^-	2.1
*E10	C_2H_5	C_2H_5	C_2H_5	$C_{12}H_{25}$	Br^-	0.064
*E13	C_3H_7	C_3H_7	C_3H_7	$C_{12}H_{25}$	Br^-	0.033
*E20	CH_3	CH_3	C_2H_5	$C_{12}H_{25}$	Br^-	0.11
*E21	CH_3	CH_3	C_3H_7	$C_{12}H_{25}$	I^-	0.15
*E22	CH_3	CH_3	C_4H_9	$C_{12}H_{25}$	I^-	0.26
*E23	CH_3	CH_3	$Br-(CH_2)_2$	$C_{12}H_{25}$	Br^-	0.21
*E30	CH_3	CH_3	$C_{12}H_{25}$	$C_{12}H_{25}$	Br^-	0.7
E40	CH_3	CH_3	$C_6H_5-CH_2$	$C_{14}H_{29}$	Cl^-	1
E41	CH_3	CH_3	$C_6H_5-CH_2$	$C_{18}H_{37}$	Cl^-	0.7
E50	$C_{16}H_{33}-\overset{+}{N}\diagdown$				Br^-	0.7

f) Quaternary ammonium salts: $R_2 - \underset{R_3}{\overset{R_1}{N^+}} - CH_2- CH_2-OH$, X^- (Series F):

NAME	R_1	R_2	R_3	X^-	IC_{50} (µM)
*F1	CH_3	CH_3	C_2H_5	I^-	>300
*F2	CH_3	CH_3	C_5H_{11}	Br^-	81.3
*F3	CH_3	CH_3	$C_{10}H_{21}$	Br^-	0.97
*F4	CH_3	CH_3	$C_{12}H_{25}$	Br^-	0.48
*F5	CH_3	CH_3	$C_{14}H_{29}$	Br^-	0.6
*F6	CH_3	CH_3	$C_{18}H_{37}$	Br^-	1.3
*F7	CH_3	$C_{12}H_{25}$	$C_{12}H_{25}$	Br^-	0.84
*F8	C_2H_5	C_2H_5	$C_{12}H_{25}$	Br^-	0.34
*F9	C_2H_5	C_2H_5	$C_{14}H_{29}$	Br^-	0.55
*F10	C_2H_5	C_2H_5	$C_{18}H_{37}$	Br^-	0.62
*F11	C_2H_5	$C_{12}H_{25}$	$C_{12}H_{25}$	Br^-	1.5
*F12	CH_3	CH_3	$\Phi-CH_2$	Br^-	112
*F13	CH_3	CH_3	$\Phi-(CH_2)_2$	Br^-	110
F14	CH_3	CH_3	$\Phi-CH_2-CH(CH_3)$	Br^-	250
*F15	CH_3	CH_3	$(cycl.)C_6H_{11}$	Br^-	120
*F16	CH_3	$-(CH_2)_4-$		I^-	230
*F18	CH_3	$-(CH_2)_2-O-(CH_2)_2-$		I^-	4830
F19	$\diagup\diagdown\overset{+}{N}(CH_3)_3$ OH			Br^-	700

g) Bisquaternary ammonium salts $R_2 - \underset{R_3}{\overset{R_1}{N^+}} - (CH_2)_n - \underset{R_3}{\overset{R_1}{N^+}} - R_2$, $2X^-$ (Series G and H):

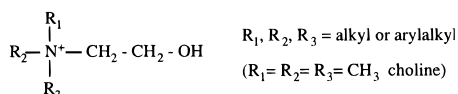| NAME	R_1	R_2	R_3	n	X^-	IC_{50} (µM)
G1	CH_3	CH_3	CH_3	6	Br^-	700
*G2	CH_3	CH_3	CH_3	8	Br^-	12
G3	CH_3	CH_3	CH_3	10	Br^-	1.7^b
*G4	CH_3	CH_3	CH_3	12	Br^-	0.09
G20	CH_3	$-(CH_2)_4-$		5	$C_4H_5O_6{}^-$	650
*G23	CH_3	$-(CH_2)_4-$		10	Br^-	0.15
*H0	$HO-CH_2-CH_2$	C_2H_5	C_2H_5	5	Br^-	1260
*H3	$HO-CH_2-CH_2$	C_2H_5	C_2H_5	10	Br^-	2
HC3	$\left[HO-CH_2-CH_2-\underset{CH_3}{\overset{CH_3}{N^+}}-CH_2-\underset{O}{\overset{}{C}}-\diagup\diagdown \right]_2$				Br^-	4^b

^{a} Compounds marked with an asterisk were synthesized in our laboratory. The activity of compounds marked b or c has already been reported in refs 6 and 5, respectively. Values for activities of compounds marked c are probably too high because of partial hydrolysis to choline during experiment.

correspond to primary, secondary, and tertiary amines, respectively. Series D (Table 1d) includes compounds possessing a trimethyl quaternary ammonium as in choline, but with various substitutions of the hydroxyl group. Series D also contains two choline analogs in which the nitrogen atom was replaced by a noncharged atom, a carbon (D5), or another charged atom, selenium (D7). Series E (Table 1e) and F (Table 1f) were composed of monoammonium quaternary salts, while series G and H (Table 1g) contained quaternary bisammonium salts. Series E to H differ by the N substitutions, which can be summarized as follows:
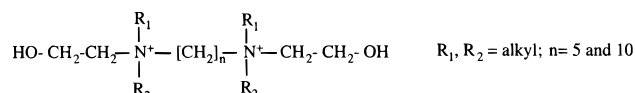
Series E:

$$R_2 - \overset{\overset{\displaystyle R_1}{|}}{\underset{\underset{\displaystyle R_3}{|}}{N^+}} - R_4 \qquad R_1, R_2, R_3, R_4 = \text{alkyl or arylalkyl}$$

Series F:

$$R_2 - \overset{\overset{\displaystyle R_1}{|}}{\underset{\underset{\displaystyle R_3}{|}}{N^+}} - CH_2 - CH_2 - OH \qquad \begin{array}{l} R_1, R_2, R_3 = \text{alkyl or arylalkyl} \\ (R_1 = R_2 = R_3 = CH_3 \ \text{choline}) \end{array}$$

Series G:

$$R_2 - \overset{\overset{\displaystyle R_1}{|}}{\underset{\underset{\displaystyle R_3}{|}}{N^+}} - [CH_2]_n - \overset{\overset{\displaystyle R_1}{|}}{\underset{\underset{\displaystyle R_3}{|}}{N^+}} - R_2 \qquad R_1, R_2, R_3 = \text{alkyl or cycloalkyl; n= 5 to 12}$$

Series H:

$$HO - CH_2 - CH_2 - \overset{\overset{\displaystyle R_1}{|}}{\underset{\underset{\displaystyle R_2}{|}}{N^+}} - [CH_2]_n - \overset{\overset{\displaystyle R_1}{|}}{\underset{\underset{\displaystyle R_2}{|}}{N^+}} - CH_2 - CH_2 - OH \qquad R_1, R_2 = \text{alkyl; n= 5 and 10}$$
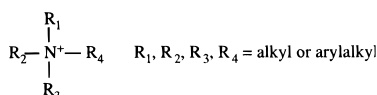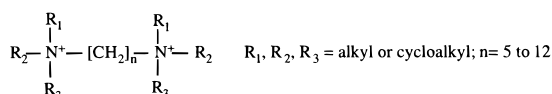
Molecular Descriptor

For comparative studies on a large number of compounds, it is desirable to be able to quantitatively manipulate data as complex as the molecular structure, taking into account parameters as diverse as molecule geometry, lipophilicity, the ability to give or receive hydrogen bonds, etc. In this context, the autocorrelation technique developed by Broto *et al.*[7] was deemed suitable for our study. Since in our case all the molecules contained at least one nitrogen atom with various substituents, the activity of drugs was more likely related to the nature and geometry of nitrogen atom substituents rather than to an overall property of molecules. The Broto autocorrelation technique describes a molecule as a number of properties distributed in a logical space (molecular graph). This is possible since most physicochemical properties can be assigned more or less directly to individual atoms. Comparing molecules is the same as comparing property distributions on their molecular graphs. Each molecule is represented by a descriptor derived from a mathematical function. After transformation, this descriptor appears as a vector, which expresses the distribution of a property, generally a physical one, over the molecular graph. An autocorrelation vector is calculated for each property (**P**) considered. Each component of this vector (P_n) is calculated from the property $P(x_i)$ intrinsic to each atom x that constitutes the molecule. The number of components (N) equals the number of bonds separating the most distant atoms + 1. Considering a property $P(x)$ distributed among the atoms of the graph, the component of order n, P_n ($0 \le n \le N - 1$) is equal to the sum

of the products $P(x_i)P(x_j)$ for all of the atoms x_i and x_j separated by n bonds. Commonly considered properties are connectivity, molecular geometry, electronegativity, van der Waals volume, lipophilic contribution, etc.

In this study, three different vectors were used to describe each molecule according to the following: (1) molecule shape [$P(x) = 1$ irrespective of the atom]; (2) lipophilicity distribution on the molecular graph ($P(x)$ = atomic contribution to the logarithm of the partition coefficient in the octanol–water system[8]); (3) the electronegativity distribution on the molecular graph [$P(x)$ = electronegativity from Pauling values]. The number of components of each vector is proportional to the molecule length. In this study, 24 components were necessary to describe the longest compounds (F7, F11), but the 11th and higher components were nil for 34% of the compounds. We therefore limited the number of components to be taken into account to 10.

Statistical Analysis of Structure–Activity Relationships

This mathematical analysis was aimed at verifying whether the biological response, here the antimalarial activity (IC_{50}) was a function of the parameters used to describe the molecule, in our case the autocorrelation vectors. This study does not give any Hansch type mathematical relationships (*i.e.*, biological response as a function of structural parameters), but rather ascertains the extent to which structurally related molecules provide similar biological results. We thus performed three different analyses, namely principal component analysis (PCA), hierarchical ascendant classification (HAC), and stepwise discriminant analysis (SDA).

Principal Component Analysis (PCA). The molecular descriptor used for this study required a high number of intercorrelated variables. In this multidimensional space, the population of studied molecules forms a distribution of points from which direct analysis is very difficult. We therefore subjected the autocorrelation vector matrices to principal component analysis.[9] The objective of this analysis was to obtain a representation of N objects initially described by K variables, using a reduced number of useful variables. The objects are then located in a system defined by these new axes. The "first principal inertia axis" of a point distribution is a straight line from which the distance of N points related to this line is minimal. The "second principal inertia axis" is perpendicular to the first one and defined so that the distance of the N points related to this second axis is minimal, and so on. In many cases, most of the information can be described by the first two or three axes. Then projection of the point distribution over the plane defined by these "principal axes" taken two by two allows a visual representation of the information contained in the initial data.

There are two major areas of interest for PCA use. First, projecting individuals on the principal factorial plane enables graphical visualization of structural similarities between molecules. Indeed, their proximity on the graph are dependent on their similarities. Besides, PCA also replaces multiple components of the vector, which are strongly intercorrelated, by new variables that are linear combinations of the previous ones, but orthogonal to each other and thus in a

convenient form for mathematical analysis, *e.g.*, hierarchical ascending classification or discriminant analysis.

Hierarchical Ascendant Classification (HAC). This classification technique reveals groups into which molecules have structural similarities (without taking the biological activity variable into account). This classification is carried out using variables obtained by PCA processing of the autocorrelation vectors. HAC groups compounds into batches. The method used was Ward's (second order moment method) using Euclidian metric and an associative algorithm called "reciprocal neighbours".[10] HAC was used to check whether molecules from the same group had similar activities, *i.e.*, to determine whether the molecular descriptor used before PCA analysis could explain activity.

Stepwise Discriminant Analysis (SDA). In this method,[11] compounds characterized by the new variables issued from PCA, were separated into two classes, "active" and "inactive". The variables were linearly combined to obtain a discriminant function (DF), so that the compounds belong to one of the two classes when the function is greater than a defined value, and to the other class in the opposite case. In other words, this method checks whether molecules in the previously defined space (PCA graph) are clustered as active or inactive. An axis, called a "discriminant axis", could thus be found in this space, on which the projection of all compounds best separates the two classes, allowing the greatest discrimination between active and inactive compounds. The discrimination is significant if the number of uncorrectly classified individuals is minimal for each class. The SDA can then be used as a predictive tool to avoid synthesizing molecules whose projection on this axis would correspond to the inactive class. To ensure the validity of SDA, it is necessary to know the position of the discriminatory axis from the factorial axes obtained by PCA, which requires determining the DF coefficients. In the discriminant analysis iterative steps, each variable (factorial axes of PCA) is investigated individually, beginning with the one having the highest discriminant power and stopping when the effect of a supplementary variable is no longer significant. After DF determination, the results were validated by comparison with those obtained after random inversion of active and poorly active compounds. This validation was carried out 10 times.

Results

Drug Effect on *P. falciparum* Growth *in Vitro*. The active concentrations (expressed as IC_{50}) of the 80 compounds against the human malaria-causing parasite, *P. falciparum*, are indicated in Table 1. For all compounds whose IC_{50} was lower than 1 mM, complete inhibition of parasite growth was obtained over 1.5–2.5 log scale concentration, indicating that the biological effect is very likely related to the drug affinity for its target (probably the choline carrier of the infected erythrocyte[2,5]) and to the concentration of the two partners according to the law of mass action.

All the compounds were found to be lethal to *P. falciparum in vitro* in a dose-dependent manner with an IC_{50} ranging from >8 mM to as low as 30 nM. Amine compounds in groups A, B, and C (except those N-substituted with a long alkyl chain, see below) exhibited far lower antimalarial activities (IC_{50} ranging from 10^{-3} to 10^{-5} M) than compounds containing one (group E or F) or two (group G, H) quaternary ammonium groups, with IC_{50} from 10^{-4} to 10^{-8} M.

Compounds of group D, possessing a trimethyl quaternary ammonium as in choline but with various substitutions of the hydroxyl group were very weakly active (IC_{50} > 1 mM), regardless of the net molecular charge of the compound at physiological pH [positive (D1, D4), negative (D2), or neutral (D3)]. In addition, replacing the quaternary ammonium by a tetrasubstituted carbon, forming a topologically similar molecule (D5), led to very low antimalarial activity indicating that a positive charge (as in a quaternary ammonium salt) is required for the activity. This feature is confirmed by the higher activity of the choline analog D7 in which a charged selenium atom replaces a charged nitrogen atom.

Compounds exhibiting the highest antimalarial activity were in groups E and F, *i.e.*, compounds that possess one quaternary ammonium, and in G and H group, which possess two quaternary ammonium groups. Concerning the monoquaternary ammonium salts, increasing the length of this alkyl chain from 2 to 18 carbon atoms led to a decrease in the IC_{50} whether they contain a hydroxyethyl group (series F) or not (series E). Hence, with alkyl chains containing seven carbon atoms or less (E1, E2, F1, F2), the IC_{50} values were higher than 80 μM. In group F, the presence of a ring structure with six carbon atoms (aromatic or not) in the short alkyl chain did not greatly modify this pattern (F12–F15, F19). With alkyl chains containing at least nine carbon atoms, the IC_{50} ranged from 2.1 to as low as 0.033 μM (E3–E23; F3–F6 and F8–F10). Interestingly, the antimalarial activity peaked with an alkyl chain of 12 methylene groups and the IC_{50} plateaued at 0.5–1 μM beyond. Further addition of methylene groups did not further increase the antimalarial activity. The long alkyl chain thus appeared to determine antimalarial activity and the presence of an aromatic cycle did not greatly modify this activity (compare E7/E40 or E9/E41) even when the nitrogen itself was engaged in an aromatic cycle (E8/E50). Besides, for the same alkyl chain length, compounds belonging to group E appeared as effective as or more than those of group F (E4/F3; E6, E10, and E13/F4 and F8; E7/F5 and F9; E9/F6 and F10). The presence of a hydroxyethyl group on the nitrogen atom is therefore not an essential condition for high antimalarial activity.

For the active compounds of series E and F with long alkyl chains, we attempted to bulk up the nitrogen group by substituting methyl groups. In series E, replacement of the three methyl substituents by three ethyls or three propyls significantly increased efficacy, indicating that an increase in lipophilicity around nitrogen is quite benificial for antimalarial potency (E6/E10/E13). In series F, *N,N*-diethyl substitution did not modify efficacy, as compared to *N,N*-dimethyl, (F4/F8; F5/F9; F6/F10). The presence of a large atom, *e.g.*, a bromine in the proximity of nitrogen, did not improve the activity (E23/E20 and E21). Moreover, the presence of a second long alkyl chain (12 carbon atoms) on the nitrogen atom did not improve the IC_{50} compared to a methyl (E6/E30 or F4/F7), and it was sometimes even unfavorable (F8/F11).

The two ammonium groups of bisquaternary ammonium salts (series G and H) were separated by a hydrophobic chain whose number of methylene groups varied. Series H compounds, as those of series F, contained the choline *N*-(2-hydroxyethyl) group. For alkyl chains containing up to 10 carbon atoms, antimalarial activity was not improved in either group as compared to monoquaternary ammonium compounds (E4/G3 or F3/H3) and was even unfavorable for short alkyl chains (F2/H0). By contrast, for 12 methylene groups, antimalarial activity was improved by the presence of an additional quaternary ammonium. Indeed, G4 was found to be 5.5-fold more efficient than E6. This means that the two polar heads of the bisquaternary ammonium compounds bind two domains (two anionic sites or two high electron density regions) of the target, separated by a distance corresponding to about 12 methylene groups.

For bisammonium salts, the same efficacy pattern was observed as for monoammonium salts, *i.e.*, the longer the hydrophobic alkyl chain, the better the antimalarial activity (G1/G2/G3/G4; G20/G23, and H0/H3). Interestingly, HC3, whose distance between the two nitrogens roughly corresponded to 12 methylene groups (as in G4), had lower antimalarial activity than G4. In this bisquaternary ammonium series, the presence of pyrrolidinium as cationic head significantly improved the activity (G23/G3 or H3).

Finally, 11 compounds (A6, E4−E6, E10, E30, F3, F8, G1, G3, and H0) tested *in vitro* against parasites resistant to current antimalarial drugs, especially chloroquine, were found to be as active as against sensitive parasites, demonstrating their efficacy in the polypharmacoresistance field (not shown).

Multidimensional Analysis of the Structure−Activity Relationship. Considering the high number of compounds for which we measured the antimalarial activity, we performed a SAR study using multidimensional methods to determine more precisely which molecular variables could be involved in the induction of the antimalarial activity. The general molecular characteristics for interference with the pharmacological target could then be drawn up, and the antimalarial activity of new molecules belonging to these series could be predicted before performing synthesis.

Such analyses had to be performed with compounds that share the same mechanism of action, *i.e.*, the same pharmacological target. Most tertiary amines and all quaternary ammonium salts act *via* specific blockage of *de novo* PC biosynthesis by inhibiting choline entry into malaria-infected erythrocytes.[2,5] This indicates that the choline carrier in infected erythrocytes was quite likely the pharmacological target of these analogs. Fifty-one compounds were retained for the following analysis on the basis of belonging to a common family whose members possess a positively charged nitrogen atom with at least three substituents. They include compounds of series C (C2, C3, C5−C7, Table 1c) and all compounds of series E to H (Table 1e−g). Group D was eliminated from the analysis because it contained molecules other than amines or ammonium salts (D5, D7) or because it included choline metabolites (D2, D3, D4).

Principal Component Analysis (PCA). Three different autocorrelation vectors (of 10 components)

were first generated to widely describe our compounds, namely "lipophilicity", "electronegativity", and "shape" vector. It was striking that each PCA analysis obtained using each of these autocorrelation vectors gave very similar results for these 51 compounds (data not shown), indicating that, on each molecular graph, the distributions of the three atomic properties (shape, lipophilicity, and electronegativity) were very similar. Indeed, the homogeneity of the considered molecules series was so high that the components of the three vectors must be closely related. In the present study, only the PCA using both the shape vector and electronegativity vector is reported, but the same conclusions could be drawn using lipophilicity instead of electronegativity (data not shown).

The 51 × 20 matrice was submitted to PCA (Table 3). The first two axes retained 97.66% of the initial information, while information retained by the five first axes accounted for up to 99.7%. We thus considered these five axes as sufficient to properly describe the structural characteristics of the molecules, and they were used for the hierarchical ascendant classification and discriminant analysis (see below). Table 4 shows the correlation between the 20 initial variables (autocorrelation "shape" vector, VS0−VS9, and autocorrelation "electronegativity" vector, VE0−VE9) and the five principal components. Axis 1 was strongly correlated with all vector components, it is a "size" axis. Axis 2 was positively correlated with the forth and fifth components of each vector, and negatively with the two last ones. It discriminated compounds mainly according to their "form" (linear or branched). It took also in account the electonegativity distribution on the molecular graph. Indeed, correlation coefficients between this axis and "electronegativity" vector components (VE) were always, at equal rank, superior to coefficients between this axis and "shape" vector components (VS). The three other principal axes, less correlated with the variables (below 0.20), were also typological axes, but their significance was more difficult to establish.

Figure 1 depicts a projection of the 51 compounds on the principal factorial plane (axis 1/axis 2). This plane, which contains about 97% of the information, provides a representation of the original point distribution without significant distortion. This illustrates the significance given to the axes. Axis 1 discriminates small molecules, as C2, C5, F1, and E1 located on the right of the axis, from large molecules, as F11, F7, E41, or HC3 which are on the left. Axis 2 distinguishes linear molecules (bottom) from branched ones (top). Very linear molecules, *e.g.*, F7 or E30 are situated negatively on this axis, whereas HC3 or F14, more branched, have positive coordinates. Besides, molecules possessing an oxygen atom, more electronegative than a carbon atom, were, at similar size and ramification, situated more positively on this axis than their carbonated homologs (compare F4/E21,F8/E10 or F6/E9).

The 51 compounds were then divided into two classes according to their *in vitro* antimalarial activity. The IC$_{50}$ of 1.5 μM was chosen as the threshold in order to obtain two equally weighed groups. Hence, 26 molecules whose IC$_{50}$ against *P. falciparum* was lower or equal to 1.5 μM were classified as active (underlined in Figure 1), while 25 molecules with IC$_{50}$ > 1.5 μM were classified as poorly active. Axis 1 principally discrimi-

Table 2. Physicochemical and Spectral Data for Synthesized Compounds

name	yield (%)	mp (°C)	^{1}H NMR (δ in ppm)
B3, HBr	66	142−4	0.88 (m, 3H); 1.27 (m, 18H); 1.55 (m, 2H); 2.65 (m, 6H); 3.65 (t, 2H)^a
B4, HBr	75	41	0.90 (m, 6H); 1.27 (m, 20H); 2.12 (s, 2H); 2.60 (m, 3H); 3.50 (m, 3H)^a
B7, HBr	75	180	0.90 (t, 3H); 1.30 (m, 18H); 1.65 (m, 2H); 3.15 (m, 2H); 3.25 (m, 2H); 3.90 (t, 2H); 3.45 (s, 3H); 7.10 (m, 2H)^a
C8, HBr	73	63	0.91 (m, 3H); 1.29 (m, 18H); 1.48 (m, 2H); 2.20 (s, 3H); 2.70 (m, 6H); 3.50 (t, 2H)^a
E2	71	>280	0.87 (m, 3H); 1.34 (m, 8H); 1.75 (m, 2H); 3.50 (s, 9H); 3.60 (m, 2H)^a
E3	50	212	0.90 (m, 3H); 1.32 (m, 12H); 1.75 (m, 2H); 3.45 (s, 9H); 3.60 (m, 2H)^b
E5	12	115	1.02 (m, 3H); 1.42 (m, 16H); 1.80 (m, 2H); 3.25 (s, 9H); 3.40 (m, 2H)^b
E10	14	84	0.85 (t, 3H); 1.40 (m, 27H); 1.68 (m, 2H); 3.22 (m, 8H)^a
E13	28	56	1.08 (m, 12H); 1.28 (m, 18H); 1.75 (m, 8H); 3.42 (t, 8H)^a
E20	89	188	0.87 (m, 3H); 1.25 (m, 18H); 1.40 (m, 3H); 1.70 (m, 2H); 3.40 (s, 6H); 3.58 (m, 2H); 3.72 (q, 2H)^a
E21	77	70	1.00 (m, 6H); 1.22 (m, 18H); 1.70 (m, 4H); 3.38 (s, 6H); 3.55 (m, 4H)^a
E22	90	62	1.00 (m, 6H); 1.27 (m, 20H); 1.72 (m, 4H); 3.42 (s, 6H); 3.60 (m, 4H)^a
E23	90	174	1.30 (m, 24H); 1.62 (m, 4H); 3.35 (m, 16H); 3.80 (m, 4H); 5.30 (t, 2H)^a
E30	24	170	0.88 (m, 6H); 1.27 (m, 36H); 1.87 (m, 4H); 2.87 (s, 6H); 3.12 (m, 4H)^a
F1	84	72−80	1.25 (t, 3H); 3.15 (s, 6H); 3.50 (m, 4H); 3.88 (m, 2H); 4.92 (t, 1H)^a
F2	55	50	0.90 (t, 3H); 1.26 (m, 4H); 1.75 (m, 2H); 3.41 (s, 6H); 3.60 (m, 2H); 3.80 (m, 2H); 4.15 (m, 2H); 5.20 (m, 1H)^c
F3	77	125−30	0.91 (t, 3H); 1.28 (m, 14H); 1.75 (m, 2H); 3.41 (s, 6H); 3.60 (m, 2H); 3.80 (m, 2H); 4.17 (m, 2H); 5.18 (m, 1H)^c
F4	70	198	0.90 (t, 3H); 1.32 (m, 18H); 1.75 (m, 2H); 3.41 (s, 6H); 3.60 (m, 2H); 3.80 (m, 2H); 4.15 (m, 2H); 5.10 (m, 1H)^c
F5	77	204	0.90 (t, 3H); 1.30 (m, 22H); 1.75 (m, 2H); 3.41 (s, 6H); 3.60 (m, 2H); 3.80 (m, 2H); 4.17 (m, 2H); 5.00 (m, 1H)^a
F6	67	204	0.90 (t, 3H); 1.27 (m, 30H); 1.75 (m, 2H); 3.41 (s, 6H); 3.60 (m, 2H); 3.80 (m, 2H); 4.15 (m, 2H); 5.05 (m, 1H)^a
F7	75	66−68	0.90 (t, 6H); 1.28 (m, 36H); 1.80 (m, 4H); 3.45 (s, 3H); 3.80 (m, 6H); 4.17 (m, 2H); 5.00 (m, 1H)^a
F8	60	135	0.90 (m, 3H); 1.30 (m, 24H); 1.70 (m, 2H); 3.54 (m, 8H); 4.28 (m, 2H); 5.15 (m, 1H)^a
F9	38	65	1.05 (m, 3H); 1.45 (m, 28H); 1.90 (m, 2H); 3.50 (m, 8H); 4.13 (m, 2H)^b
F10	45	136	0.88 (m, 3H); 1.30 (m, 36H); 1.75 (m, 2H); 3.50 (m, 8H); 4.25 (m, 2H); 5.20 (m, 1H)^a
F11	72	98−99	0.92 (t, 6H); 1.34 (m, 39H); 1.75 (m, 4H); 3.60 (m, 8H); 4.16 (m, 2H); 5.00 (m, 1H)^a
F12	66	80−85	3.17 (s, 6H); 3.55 (m, 4H); 3.98 (m, 2H); 5.95 (t, 1H); 7.70 (m, 5H)^c
F13	70	100	3.35 (s, 8H); 3.65 (m, 4H); 3.90 (m, 2H); 5.40 (t, 1H); 7.42 (m, 5H)^c
F15	71	130	1.30 (m, 6H); 1.70 (m, 4H); 3.30 (s, 6H); 3.70 (m, 5H); 5.20 (m, 1H)^c
F16	90	135	2.75 (m, 4H); 3.07 (s, 3H); 3.60 (m, 6H); 3.82 (m, 6H); 5.83 (t, 1H)^c
F18	90	130	3.25 (s, 3H); 3.60 (m, 6H); 3.90 (m, 6H); 5.08 (m, 1H)^c
G2	60	291	1.52 (m, 8H); 1.92 (m, 4H); 3.25 (s, 18H); 3.50 (m, 4H)^a
G4	60	231	1.45 (m, 16H); 1.90 (m, 4H); 3.10 (s, 18H); 3.30 (m, 4H)^a
G23	86	260 degrad	1.40 (s, 12H); 1.70 (m, 4H); 2.10 (m, 8H); 3.05 (s, 6H); 3.45 (m, 12H)^c
H0	19	241	1.40 (m, 14H); 1.85 (m, 4H); 3.50 (m, 16H); 4.25 (m, 4H)^b
H3	77	188−90	1.30 (m, 24H); 1.62 (m, 4H); 3.35 (m, 16H); 3.80 (m, 4H); 5.30 (t, 2H)^a

^{a} In CDCl$_3$. ^{b} In D$_2$O. ^{c} In DMSO-d_6.

Table 3. Principal Component Analysis of Autocorrelation Vectors ("Shape" Vector + "Electronegativity" Vector)

no.	eigen values^a	trace	cumulative %
1	18.99	94.96	94.96
2	0.54	2.70	97.66
3	0.18	0.90	98.55
4	0.13	0.65	99.21
5	0.10	0.49	99.70
6	0.02	0.12	99.82
7	0.02	0.08	99.89
8	0.01	0.04	99.94
9	0.01	0.03	99.96
10	0.00	0.02	99.98
11	0.00	0.01	99.99
12	0.00	0.01	100.00
13	0.00	0.00	100.00
14	0.00	0.00	100.00
15	0.00	0.00	100.00
16	0.00	0.00	100.00
17	0.00	0.00	100.00
18	0.00	0.00	100.00
19	0.00	0.00	100.00
20	0.00	0.00	100.00

^{a} Eigen values are given in decreasing order. The first five components contain 99.7% of the initial information.

Table 4. Correlations between Variables and Principal Components^a

	axis 1	axis 2	axis 3	axis 4	axis 5
VS0	−0.98	−0.15	−0.07	0.06	0.12
VS1	−0.98	−0.10	−0.11	0.02	0.11
VS2	−0.97	0.08	−0.18	−0.14	0.03
VS3	−0.96	0.23	0.08	0.00	0.10
VS4	−0.98	0.15	0.10	0.09	0.07
VS5	−0.99	−0.02	−0.03	0.12	−0.00
VS6	−0.97	−0.17	−0.07	0.12	−0.01
VS7	−0.98	−0.21	−0.01	0.05	−0.01
VS8	−0.97	−0.20	0.07	−0.05	0.00
VS9	−0.96	−0.22	0.14	−0.10	0.05
VE0	−1.00	0.02	−0.04	0.02	0.02
VE1	−0.99	0.03	−0.11	−0.02	0.01
VE2	−0.95	0.21	−0.15	−0.16	−0.05
VE3	−0.94	0.31	0.08	−0.01	0.03
VE4	−0.95	0.28	0.11	0.04	−0.04
VE5	−0.99	0.10	−0.00	0.08	−0.07
VE6	−0.98	0.01	−0.05	0.05	−0.16
VE7	−0.99	−0.07	0.01	0.00	−0.12
VE8	−0.98	−0.12	0.08	−0.08	−0.06
VE9	−0.97	−0.14	0.15	−0.12	−0.02

^{a} VS, shape vector; VE, electronegativity vector.

nates active from poorly active compounds. Small molecules, on the right of this axis, are mostly poorly active, whereas bulkier molecules, on the left, are almost all active. The most active compounds whose IC$_{50}$ was <0.3 μM (*E10*, *E13*, *E20−E23*, *F4*, *F8*, *G4*, and *G23*) are all situated in a restricted area of the plane.

Hierarchical Ascending Classification (HAC). Successive clustering of the 51 compounds according to their structural similarities in the PCA space provided a tree which highlighted three distinct groups:

Cluster 1: C2, C3, C5, C6, C7, E1, E2, E3, E4, E5, E6, F1, F2, F3, F12, F13, F15, F16, F18, F19, G1, G2

Cluster 2: E7, E8, E9, *E10*, *E20*, *E21*, *E22*, *E23*, F4, F5, F8, F14, G3, *G4*, G20

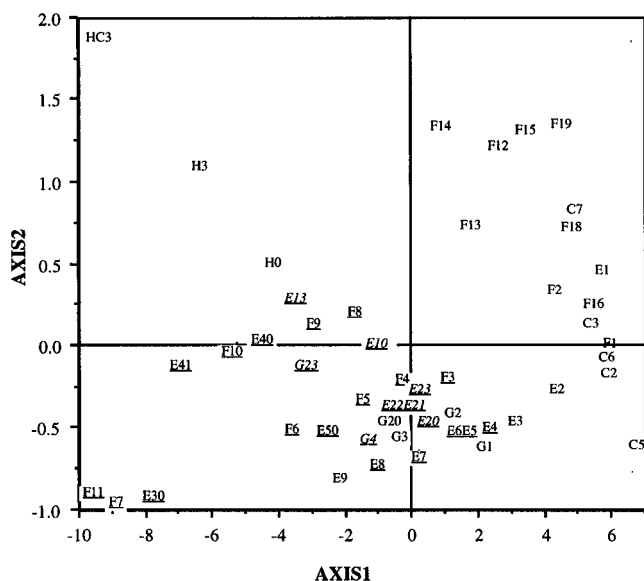**Cluster 3**: *E13*, E30, E40, E41, E50, F6, F7, F9, F10, F11, *G23*, H0, H3, HC3

Figure 1. Projection of the 51 compounds on the principal factorial plane. Antimalarial active compounds ($IC_{50} \leq 1.5$ μM) are underlined, and the most active ($IC_{50} < 0.3$ μM) are in italics.

Figure 2. Percentage of correctly classified compounds from SDA according to selected variables. This figure shows the results for active compounds ($IC_{50} < 1.5$ μM) ($\triangle$), poorly active ($IC_{50} > 5$ μM) ($\bigcirc$), and all 46 compounds ($\blacksquare$).

The group distribution notably depends on the coordinates of axis 1 (the axis of minimum inertia), *i.e.*, roughly according to the molecular size. Cluster 1 contains the smallest molecules, tertiary amines (series C) which do not have bulky substituents, quaternary ammonium salts (series E and F) whose nitrogen substituents contain less than 12 carbon atoms, except E6, and two quaternary bisammonium compounds (G1, G2) whose nitrogen atoms are separated by an alkyl chain of six or eight methylene groups. These molecules were generally poorly active ($IC_{50} > 1.5$ μM) except for four compounds out of 22 (E4 0.7 μM, E5 0.5 μM, E6 0.5 μM, and F3 0.97 μM).

Clusters 2 and 3 contain the largest molecules. Most of them were active except for E9 (2.2 μM), G3 (1.7 μM), H3 (2 μM), and HC3 (4 μM), whose IC_{50} was very close to the activity threshold (1.5 μM) fixed to discriminate compounds, and for F14 (250 μM), G20 (650 μM), and H0 (1260 μM). Cluster 2 contains less bulky molecules as compared to cluster 3. It also contains the majority of the most active compounds (six out of eight with IC_{50} lower than 0.3 μM). The molecular descriptors used for PCA were thus pertinent for explaining the antimalarial activity.

Stepwise Discriminant Analysis (SDA). The compounds, characterized by the five independent variables obtained from PCA (coordinates on axes 1−5), were separated as described above in two classes according to their activity. For rigorous stepwise discriminant analysis, we eliminated from this SDA five compounds (E3, E9, G3, H3, HC3) whose IC_{50} values were too close to the limit fixed for class activity discrimination between poorly active and active compounds (1.5 μM < IC_{50} < 5 μM).

The remaining 46 compounds were then submitted to a discriminant analysis. This allowed a stepwise selection of axis 1, axis 2, axis 4, axis 5, and lastly axis 3, as variables explaining activity. Figure 2 shows the percentage of correctly classified individuals according to the number of variables introduced by SDA. In the first step, axis 1 (which discriminates according to
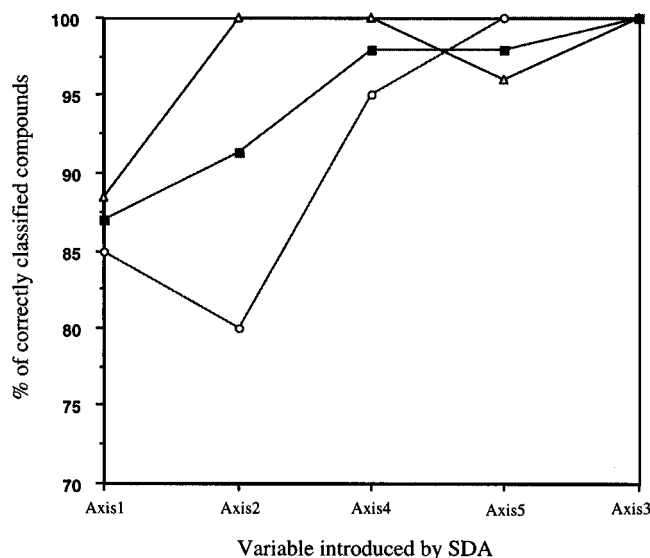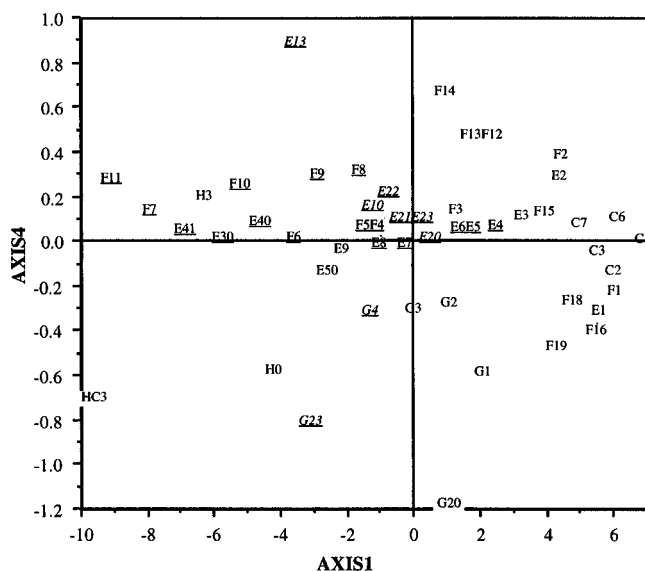

Figure 3. Projection of the 51 compounds on the factorial plane axis 1/axis 4. Antimalarial active compounds ($IC_{50} \leq 1.5$ μM) are underlined, and the most active ($IC_{50} < 0.3$ μM) are in italics.

molecular size) correctly classified 88.5% of the active compounds and 85% of poorly active molecules. In addition to axis 1, introduction of axis 2, which discriminates according mainly to the linearity of the molecules, but also to the presence of an heteroatom, such as an oxygen, allowed a faultless classification of active compounds, and only four inactive molecules out of 20 (G1, G2, G20 and H0) were predicted as being active. The additional introduction of axis 4 increased the overall classification, since 45 compounds over 46 were correctly predicted. The significance of axis 4 (which only represents 0.65% of the variance, see Table 3) was difficult to determine precisely, but could separate symmetrical molecules (*i.e.*, quaternary bisammonium salts: G and H series) from very asymmetrical molecules, such as E13, F12−F14 (see Figure 3). Introduction of axis 5 did not modify the overall prediction. Last, addition of axis 3 allowed to perfectly classify all

Table 5. Stepwise Discriminant Analysis[a]

	coefficients of the discriminant function (DF)	Fischer
axis 1	−0.70	42
axis 2	−3.65	31
axis 4	3.64	26
	0.77	

[a] For each compound DF = −0.70(axis 1) − 3.65(axis 2) + 3.64(axis 4) + 0.77. If DF > 0, the compound is predicted to be active; if DF < 0, the compound is predicted to be inactive.

the compounds. Hence, three variables (axes 1, 2 and 4) were sufficient for significant discrimination (98%). Table 5 indicates the coefficients of the discriminant function (DF) for these three variables. A compound was predicted to be active if the DF, applied to its coordinates on axes 1, 2, and 4, led to a positive value; it was considered inactive otherwise. In the present case, where it depends on three variables, DF is the equation of a plane in the space of these three variables (axis 1, axis 2, and axis 4). This plane perpendicular to the "discriminant axis" divides the space in two areas, each corresponding to one class (active or inactive compounds). In this way, only one compound out of 46 was incorrectly predicted as active: G2 (12 μM). However, the DF of this compound was very close to zero, *i.e.*, just at the boundary of the discriminatory limit (data not shown). This quite low number of uncorrectly classified molecules highlights the high significance of discriminant analysis.

Discussion

In the past 20 years, there has been a very impressive increase in our understanding of the biochemistry and molecular biology of malaria parasites, with attention focused on specific parasite molecules that are keys to the parasite life cycle. Directed pharmacology research has involved identification and characterization of targets that can be specifically affected pharmacologically. Membrane biogenesis accompanying parasite growth requires parasite-driven lipid synthesis because the erythrocyte lacks this capacity.[2,3] This knowledge has led to development of a novel antimalarial pharmacological approach involving analogs of PL polar head groups acting by substitution,[6] competition,[5,12] or using unnatural fatty acids.[13] The most promising results were obtained with choline transport blockers, probably mediated by the native (though altered) carrier of the erythrocyte membrane. This target is easily accessible from the extracellular space and constitutes the rate-limiting step for the supply of choline to the parasite for PC synthesis.[4]

A structure−activity relationship analysis aims at determining pharmacophore requirements for optimal inhibition of the malarial parasite growth. In the present case, multidimensional analyses have shown that shape, electronegativity, and lipophilicity properties of the compound were clearly related to the *in vitro* antimalarial activity against the human parasite *P. falciparum*. PCA, HAC, and SDA analyses of these variables enabled satisfactory discrimination (about 98%) between active and poorly active compounds. Therefore the general features for a compound to be active could be drawn up. Provided that a compound belongs to one of the series defined above, this method is able to quickly and successfully predict its antimalarial activity before synthesis.

Structural variable description by an autocorrelation vector is an efficient way to take into account not only the molecular form but also physicochemical property distribution on the molecular graph. In our case, the homogeneity of the molecule series was such that the components of the three vectors were roughly related, and positive correlations highlighted the importance of the "lipophilicity" and "electronegativity" distribution on the molecule, in QSAR studies. Total interactions between compound and target clearly involved two components, with an ionic interaction as a predominant prerequisite, together with a desirable hydrophobic interaction.

Overall, these results also roughly indicate the structure of the pharmacological target (likely the choline carrier) in the *Plasmodium*-infected erythrocyte. Probably, the cationic molecules meet size and local packing requirements to fit into an anionic (or electron-rich) pocket of the proposed model. The anionic site appears to be large enough to accommodate a nitrogen atom bearing two methyls together with one butyl group (E22) or a nitrogen atom bearing three ethyls or three propyl groups (E10 or E13). When steric hindrance is further increased, *e.g.*, presence of two dodecyl groups, one of the dodecyl chains will be forced to extend out ot the anionic pocket which is energetically more unfavorable, thus leading to a decrease in affinity (E30, F11). H bonding is probably not essentiel for enabling the cationic portion of active molecules to come within the anionic pocket, since long active molecules devoid of a hydroxyethyl group (series E or G) have the same activity as those possessing one (series F or H).

Although a positive charge is required to guide quaternary ammonium compounds into the vicinity of the target, there may be a balance between two types of interaction with the transporter, one with the anionic site and another with an adjacent hydrophobic domain capable of accomodating polymethylene lipophilic chains, roughly equivalent to 10−12 methylene groups. However, lipophilicity needs a strict spatial distribution. When the alkyl length exceeds the permissible linear dimension chain (11 or 12 methylene groups), energetically less favorable folding of the chain would occur, leading to decreased activity. A possible explanation could be that above 12 methylene groups the alkyl chains could curl up on themselves, leading to expulsion of the quaternary ammonium group out of the anionic site and only the long alkyl chain would be associated in tightly coiled fashion with the hydrophobic region on the target.

When the polymethylene chain contains more than 10 methylene groups, the bisquaternary ammonium compounds were more active than the corresponding monoquaternary compounds (see the 19-fold increase in efficacies between G3 and G4 which possess 10 and 12 methylene groups, respectively). This indicates that two anionic sites separated by the hydrophobic domain are probably present. Up to 10 methylene groups (G3), the alkyl chain probably just barely extends across the hydrophobic domain and deformation of the anionic site structure would be required for envelopment of quaternary groups. Longer polymethylene chains (G4) may furnish larger regions for hydrophobic interaction with the target and allow the quaternary groups to dip freely into the two anionic pockets. Hence, the presence of a

second quaternary ammonium in the molecule probably allows a higher interaction with the active site than would be the case with a monocationic compound.

Finally, the presence in the hydrophobic molecule domain of aromatic rings or the possibility of hydrogen bond formation with the hydrophobic region of the binding site (*e.g.*, through ketone in HC3) does not seem to be critical for the antimalarial effect (*e.g.*, HC3 whose distance between the two quaternary ammonium groups is equivalent to that of G4, although possessing distinct substituents led to much lower activity). This also suggests that the hydrophobic target domain probably does not contain H-bonding donors. Nevertheless, synthesis of additional compounds with different substituents in the hydrophobic domain would be required to draw general conclusions.

In summary, the most potent *in vitro* activities were obtained for *N*-dodecyl-substituted monoammonium compounds, E6, *E10*, *E13*, *E20*−*E23*, F4, F8, and for bisammoniums *G4* and *G23* which have two polar heads separated by a C_{10} or C_{12} alkyl chain. The pharmacological target likely contains an anionic site (or a high electron density region) able to combine a relatively reduced (*N,N,N*-trimethyl, *N,N*-dimethyl-*N*-hydroxyethyl, or *N,N*-diethyl-*N*-hydroxyethyl) positively charged polar head. Analogs that are bulkier than choline (*N,N,N*-triethyl, *N,N,N*-tripropyl groups), *E10* and *E13*, make closer contact with this site, but substrates that are too large (with a second long C_{12} alkyl chain) to enter this space are less well accepted. Clearly, this makes the nitrogen substitution quite subtle and essential for optimal interactions. Adjacent to this site, there is probably an extensive nonpolar region capable of adsorbing only one long polymethylene chain. The efficacies of bisquaternary ammonium salts are consistent with the presence of two anionic sites on the choline carrier separated by a large lipophilic domain.

This study clearly revealed some structural characteristics concerning inhibition of the malaria causing parasite by cationic compounds, analogs of choline. These compounds are inexpensive to produce, stable, and water soluble. Two obligatory steps will now have to be fulfilled to validate this pharmacological model at the preclinical study level. They concern toxicity and levels of activity. Suitable drugs active *in vitro* at concentrations lower than 10 nM could probably be obtained through additional links with the target. The SAR using multidimensional methods performed in the present study would be decisive for predicting active molecules, notably on the basis of their lipohilicity and molecular size. The risk arising with choline analogs might have been their adverse cholinergic side effects. Some ammonium and bisammonium quaternary salts are indeed inhibitors of the high affinity choline transport into synaptosomes.[14,15] We have, however, shown that some characteristics of these two transporters are different enough to obtain specificity and allow the use of these compounds *in vivo* as antimalarial agents with a reasonable safety margin (unpublished results).

Experimental Section

Chemistry. Melting points were determined in capillary tubes on a Thomas-Hoover apparatus and are uncorrected. ^{1}H NMR spectra were recorded on a Brucker AC 250 (250 MHz). Chemical shifts (δ) are expressed in ppm downfield from the internal standart tetramethylsilane (TMS). Elemental analyses were within ±0.4 of calculated values.

Some compounds were commercially available: A0−A9, B1, B2, B5, B6, C1−C7, C12, D1−D6, E1, E4, E6−E9, E40, E41, E50, G1, G3, G20, and HC3 were obtained from Sigma-Aldrich Chemical Co. (St. Louis, MO). D7 was provided by Prof. J. W. Kosh (University of South Carolina), and F14 and F19 by Dr. J.Berthe (Sanofi, Montpellier, France). The other 35 compounds (marked with an asterisk in Table 1) were synthesized in our laboratory, *i.e.*, most of the ammonium quaternary salts.

General Procedure for the Synthesis of Secondary and Tertiary Amines and Ammonium Quaternary Salts. Secondary and tertiary amines and ammonium quaternary salts were generally performed by alkylation of a primary, secondary, or tertiary amine, respectively, using the corresponding alkyl halide in a polar solvent.

***N*-(2-Hydroxyethyl)-1-dodecylamine Bromhydrate (B3), 2-(Dodecylamino)butan-1-ol Bromhydrate (B4), *N*-(2-Methoxyethyl)-1-dodecylamine Bromhydrate (B7), and *N*-(2-Hydroxyethyl)-*N*-methyl-1-dodecylamine Bromhydrate (C8).** 1-Bromododecane (0.1 mol), amine (0.2 mol) (2-hydroxyethylamine for B3, 2-aminobutan-1-ol for B4, 2-(methoxyethyl)amine for B7, or the *N*-methyl-*N*-(2-hydroxyethyl)-amine for C8), and triethylamine (0.2 mol) were added to ethanol (100 mL) and heated under reflux conditions for 12 h. After solvent removal, the residue was dissolved in water and extracted with dichloromethane. The solution was dried, and after solvent removal, the oily residue was dissolved in hexane (100 mL). Dry HBr was then bubbled through the hexane solution and the bromhydrate was isolated.

***N,N,N*-Trimethyl-1-heptanaminium Bromide (E2), *N,N,N*-Trimethyl-1-nonanaminium Bromide (E3), *N,N,N*-Trimethyl-1-undecanaminium Bromide (E5), *N,N,N,N,N,N*-Hexamethyl-1,8-octanediaminium Dibromide (G2), and *N,N,N,N,N,N*-Hexamethyl-1,12-dodecanediaminium Dibromide (G4).** Sodium hydroxide (0.4 mol) was added to a solution of trimethylamine chlorhydrate (0.35 mol) in anhydrous methanol (30 mL). The mixture was filtered and the halogenated derivative (0.2 mol) or dihalogenated derivative (0.1 mol) (1-bromoheptane for E2, 1-bromononane for E3, 1-bromoundecane for E5, 1,8-dibromooctane for G2, or 1,12-dibromododecane for G4) was added to the filtrate. The solution was heated under reflux conditions for 12 h. The solvent was removed, and the salt was recrystallized from a mixture of methanol and diethyl ether.

***N,N,N*-Triethyl-1-dodecanaminium Bromide (E10).** 1-Bromododecane (0.02 mol) was added to a solution of triethylamine (0.03 mol) in anhydrous ethanol (60 mL). The solution was heated under reflux conditions for 12 h; the solvent was removed, and the salt was recrystallized from a mixture of ethanol−acetone−diethyl ether.

***N,N,N*-Tripropyl-1-dodecanaminium Bromide (E13).** 1-Bromododecane (0.02 mol) was added to a solution of tripropylamine (0.02 mol) in anhydrous acetone (50 mL). The solution was heated under reflux conditions for 72 h. The solvent was removed and the salt was recrystallized from a mixture of acetone−diethyl ether.

***N*-Ethyl-*N,N*-dimethyl-1-dodecanaminium Bromide (E20), *N,N*-Dimethyl-*N*-propyl-1-dodecanaminium Bromide (E21), *N*-Butyl-*N,N*-dimethyl-1-dodecanaminium Iodide (E22), *N*-(2-Bromoethyl)-*N,N*-dimethyl-1-dodecanaminium Bromide (E23), and *N*-dodecyl-*N,N*-dimethyl-1-dodecanaminium Bromide (E30).** *N,N*-Dimethyl-1-dodecylamine (0.23 mol) was added to a solution of halogenated derivative (0.8 mol) (bromoethane for E20, 1-iodopropane for E21, 1-iodobutane for E22, 1,2-dibromoethane for E23, and 1-bromododecane for E30) in anhydrous ethanol (60 mL). The solution was heated under reflux conditions for 24 h. The solvent was removed, and the salt was recrystallized from a mixture of ethanol−diethyl ether or acetone (for E23).

***N*-(2-Hydroxyethyl)-*N,N*-dimethylethanaminium Iodide (F1), *N*-(2-Hydroxyethyl)-*N,N*-dimethyl-1-pentanaminium Bromide (F2), *N*-(2-Hydroxyethyl)-*N,N*-dimethyl-1-decanaminium Bromide (F3), *N*-(2-Hydroxyethyl)-*N,N*-dimethyl-1-dodecanaminium Bromide (F4), *N*-(2-Hydroxyethyl)-*N,N*-dimethyl-1-tetradecanaminium Bromide (F5), *N*-(2-Hydroxyethyl)-*N,N*-dimethyl-1-octade-**

canaminium Bromide (F6), *N*-(2-Hydroxyethyl)-*N*,*N*-dimethyl-*N*-phenylmethylammonium Bromide (F12), *N*-(2-Hydroxyethyl)-*N*,*N*-dimethyl-*N*-(2-phenylethyl)ammonium Bromide (F13), and *N*-(2-hydroxyethyl)-*N*,*N*-dimethyl-*N*-cyclohexylammonium Bromide (F15). 2-(dimethylamino)ethanol (0.12 mol) and the corresponding brominated (or iodinated) derivative (0.11 mol) were added to methanol or ethanol (100 mL) and heated under reflux conditions for 20 h. The solvent was removed and the crude product recrystallized from a mixture of methanol and diethyl ether.

N,*N*-Diethyl-*N*-(hydroxyethyl)-1-dodecanaminium Bromide (F8), *N*,*N*-diethyl-*N*-(hydroxyethyl)-1-tetradecanaminium Bromide (F9), *N*,*N*-diethyl-*N*-(2-hydroxyethyl)-1-octadecanaminium Bromide (F10), *N*,*N*,*N*,*N*-tetraethyl-*N*,*N*-bis(2-hydroxyethyl)-1,5-pentanediaminium Dibromide (H0), and *N*,*N*,*N*,*N*-Tetraethyl-*N*,*N*-bis(2-hydroxyethyl)-1,10-decanediaminium dibRomide (H3). The above compounds were prepared according to a procedure similar to that given above (for compounds F1 *et al.*) using 2-(diethylamino)ethanol instead of 2-(dimethylamino)ethanol. For syntheses of H0 and H3, half the usual quantity of relevant dihalogenated derivative was used.

N-Dodecyl-*N*-methyl-*N*-(2-hydroxyethyl)-1-dodecanaminium Bromide (F7) and *N*-dodecyl-*N*-ethyl-*N*-(2-hydroxyethyl)-1-dodecanaminium Bromide (F11). 2-(methylamino)ethanol (0.1 mol) (for F7) or 2-(ethylamino)ethanol (0.1 mol) (for F11), 1-bromododecane (0.22 mol), and diisopropylethylamine (0.11 mol) were added to ethanol (100 mL) and heated under reflux conditions for 48 h. The solvent was removed, and several crystallizations from a mixture of methanol and diethyl ether eliminated the diisopropylethylamine bromohydrate which first precipitated. The filtrate was evaporated and the resulting product was recrystallized from a mixture of methanol and diethyl ether.

1-Methyl-1-(2-hydroxyethyl)pyrrolidinium Iodide (F16) and 4-(2-Hydroxyethyl)-4-methylmorpholinium (F18). 1-(2-Hydroxyethyl)pyrrolidine (for F16) or 4-(2-hydroxyethyl)morpholine (for F18), with an excess of methyl iodide, was added to methanol and heated under reflux conditions for 15 h. The crude products were recrystallized from a mixture of methanol and diethyl ether.

1,10-Decamethylenebis(1-methyl pyrrolidinium) Dibromide (G23). This compound was prepared according to a procedure similar to that used in the synthesis of H0 and H3, using 1,10-dibromodecane and *N*-methylpyrrolidine.

The synthesis yields, melting points, and NMR spectra of the compounds are reported in Table 2.

Biological Activity. The Nigerian strain of *P. falciparum* was maintained at 37 °C by serial passages in human erythrocytes suspended in Hepes-buffered RPMI 1640 supplemented with 10% AB human serum using the petri-dish candle-jar method.[16]

The antimalarial activity of the compound was measured in microtiter plates according to the Desjardins *et al.* method.[17] The final volume of the incubation medium in each well was 200 μL, and the haematocrit of the *P. falciparum*-infected erythrocyte suspension was 1−2%, with an initial parasitemia (*i.e.*, the percentage of infected erythrocytes) of 0.3−0.5%. In some cases, the drugs were dissolved in DMSO and then further diluted in culture medium so that the final DMSO concentration never exceeded 0.25%. After contact of the drug with the parasite for one parasite cycle (48 h), [^{3}H]hypoxanthine was added for 18 h to assess parasite viability. The results are expressed as the drug concentration resulting in 50% inhibition (IC$_{50}$) of parasite growth and are means of at least two independent experiments performed in triplicate using different drug stock solutions.

Multidimensional Analyses. Molecular structures were generated by the software MAD.[18] The structural descriptors were calculated and the mathematical treatment of data performed by TSAR,[18] on a Hewlett-Packard workstation APOLLO 715.

References

(1) Wernsdorfer, W. H.; Payne, D. The dynamics of drug resistance in *Plasmodium falciparum. Pharmacol. Ther.* **1991**, *50*, 95−121.

(2) Vial, H.; Ancelin, M. L. Malarial Lipids, an overview. In *Subcellular Biochemistry*; Avila, J. L., Harris, J. R., Eds.; Plenum Press: New York, 1992, Vol. 18, pp 259−306.

(3) Van Deenen, L. L. M.; De Gier, J. In *The red blood cell*; Surgenor, G., Ed.; Academic Press: New York, 1975, pp 147−211.

(4) Ancelin, M. L.; Vial, H. J. Regulation of phosphatidylcholine biosynthesis in *Plasmodium*-infected erythrocytes. *Biochim. Biophys. Acta* **1989**, *1001*, 82−89.

(5) Ancelin, M. L.; Vial, H. J. Quaternary ammonium compounds efficiently inhibit *Plasmodium falciparum* growth *in vitro* by impairment of choline transport. *Antimicrob. Agents Chemother.* **1986**, *29*, 814−820.

(6) Vial, H. J.; Thuet, M. J.; Ancelin, M. L.; Philippot, J. R.; Chavis, C. Phospholipid metabolism as a new target for malaria chemotherapy. Mechanism of action of D-2-amino-1-butanol. *Biochem. Pharmacol.* **1984**, *33*, 2761−2770.

(7) Broto, P.; Moreau, G.; Vandycke, C. *Eur. J. Med. Chem.−Chim. Ther.* **1984**, *19*, 61−65.

(8) Broto, P.; Moreau, G.; Vandycke, C. *Eur. J. Med. Chem.−Chim. Ther.* **1984**, *19*, 71−78.

(9) Manly, B. F. J. *Multivariate Statistical Methods: A primer*; Chapman and Hall: New York, 1986.

(10) Ward J. H. *J. Am. Stat. Assoc.* **1963**, *58*, 61−69.

(11) Romeder J. M. Méthode, Programme d'analyse discriminante. Dunod: Paris, 1973.

(12) Ancelin, M. L.; Vial, H. J.; Philippot, J. R. Inhibitors of choline transport into *Plasmodium*-infected erythrocytes are effective antiplasmodial compounds *in vitro. Biochem. Pharmacol.* **1985**, *34*, 4068−4071.

(13) Beaumelle, B. D.; Vial, H. J. Correlation of the efficiency of fatty derivatives in suppressing *Plasmodium falciparum* growth in culture with their inhibitory effect on acyl-CoA synthetase activity. *Mol. Biochem. Parasitol.* **1988**, *28*, 39−42.

(14) Tamaru, M.; Roberts, E. Structure-activity studies on inhibition of choline uptake by a mouse brain synaptosomal preparation: basic data. *Brain Res.* **1988**, *473*, 205−214.

(15) Roberts, E.; Tamaru, M. The ligand binding site of the synaptosomal choline transporter: a provisional model based on inhibition studies. *Neurochem. Res.* **1992**, *17*, 509−528.

(16) Jensen, J. B.; Trager, W. *Plasmodium falciparum* in culture: use of out dated erythrocytes and description of the candle-jar method. *J. Parasitol.* **1977**, *63*, 883−886.

(17) Desjardins, R. E.; Canfield, C. J.; Haynes, J. D.; Chulay, J. D. Quantitative assessement of antimalarial activity *in vitro* by a semiautomated microdilution technique. *Antimicrob. Agents Chemother.* **1979**, *16*, 710−718.

(18) MAD V2.0, TSAR V2.0, Oxford Molecular Ltd., Magdalen Centre, Oxford Science Park, Sandford-on Thames, Oxford, OX4 4GA, England.

JM9701886