3D QSAR Investigations on Antimalarial Naphthylisoquinoline Alkaloids by Comparative Molecular Similarity Indices Analysis (CoMSIA), Based on Different Alignment Approaches†

Gerhard Bringmann* and Christian Rummey

Institut für Organische Chemie, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany

Received July 26, 2002

3D QSAR models based on the CoMSIA descriptor fields were established using a diverse data set of 53 antimalarial biaryl compounds (tested in vitro against a chloroquine-resistant strain of *Plasmodium falciparum*), consisting mainly of naphthylisoquinoline alkaloids, but also including phenylanthraquinone structures and naphthylindenes. For the alignment, two commercially available automated approaches, FLEXS and GASP, were compared; initially none of them succeeded in treating the important phenomenon of axial chirality correctly, but after some manual refinement of the alignments initially obtained, the best overall model, based on a modified FLEXS alignment, showed a q^2 (cross-validated r^2) of 0.818 (eight components), using only the hydrophobic and the H-bond donor and acceptor fields. Using a test set of five compounds the model showed a squared multiple correlation coefficient for the test set (predictive r^2) of 0.578. The analysis of the 3D contour maps permitted interesting conclusions about the effects of particular functional groups on the biological activity and will now guide the design of novel, hopefully even more active compounds.

INTRODUCTION

Naphthylisoquinoline alkaloids,[1] among them dioncopeltine A[2] (**1a**, see Figure 1) and dioncophylline C[3] (**3a**), constitute a group of structurally remarkable and biosynthetically unprecedented acetogenic[4] natural products, originally isolated from tropical lianas of the plant families Ancistrocladaceae and Dioncophyllaceae.[5,6] These alkaloids exhibit an interesting spectrum of promising bioactivities, including antileishmanial,[7] antitrypanosomal,[8] fungicidal,[9] and (for some dimeric representatives) also anti-HIV[10] properties. Of even higher importance is their good antiplasmodial activity, both in vitro[11,12] and in vivo;[13] this finding is of particular significance since malaria, as the most important infectious tropical disease, is the cause for over 1 000 000 death cases per year.[14] The increasing resistance against established antimalarial drugs[15] emphasizes the demand for new lead structures and makes naphthylisoquinoline alkaloids attractive synthetic targets. The dominating structural element of these natural products is the usually rotationally hindered and thus stereogenic biaryl axis as e.g. in dioncopeltine A (**1a**, natural), preventing an interconversion with its likewise available, but as yet only synthetic, unnatural atropo-diastereomer **2a**.

Using the "lactone method",[16,17] a novel synthetic pathway to even highly hindered biaryl systems, we have synthesized more than 20 such alkaloids in their authentic, natural forms[17−19] but have also prepared first structural analogues. This, besides our continued isolation work,[20] gave us the possibility of testing a broad variety of such compounds against *Plasmodium falciparum* in vitro and, for the most promising representatives, also in vivo.

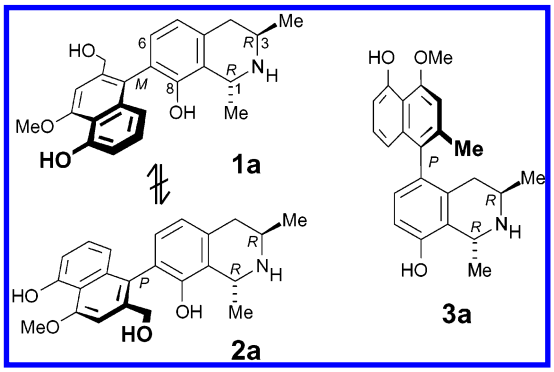

Figure 1. Two of the most active antiplasmodial naphthylisoquinoline alkaloids and part of our data set, dioncopeltine A (**1a**) and dioncophylline C (**3a**), both initially isolated from the tropical liana *Triphyophyllum peltatum*;[2,3] the as yet merely synthetic atropodiastereomer **2a** cannot be converted into **1a** at room temperature.

Some preliminary information has been obtained about the antimalarial mode of action of naphthylisoquinoline alkaloids—i.e., information about its phase-sensitivity,[21,22] the activity against erythrocytic[23] and exoerythrocytic[24] forms, an impairment of heme degradation and polymerization, and about their interaction with parasitic redox enzymes.[25] Since nothing is as yet known about a specific target, and because of the stereochemically demanding structures of the alkaloids, a 3D QSAR analysis seemed rewarding to find more active compounds with hopefully simplified structures.

For this purpose, the use of molecular fields as descriptors for the correlation of biological activities with 3D chemical structures has become an efficient tool frequently applied, especially after the introduction of the CoMSIA (comparative molecular similarity indices analysis) technique,[26] an improved modification of the CoMFA method.[27]

The crucial step of such an analysis is the molecular alignment. Many compounds of our data set, including the

* Corresponding author phone: (+49)931-888-5323; fax: (+49)931-888-4755; e-mail: bringman@chemie.uni-wuerzburg.de.

† Acetogenic Isoquinoline Alkaloids, part 154. For part 153, see reference 76.

10.1021/ci025570s CCC: $25.00
Published on Web 01/08/2003

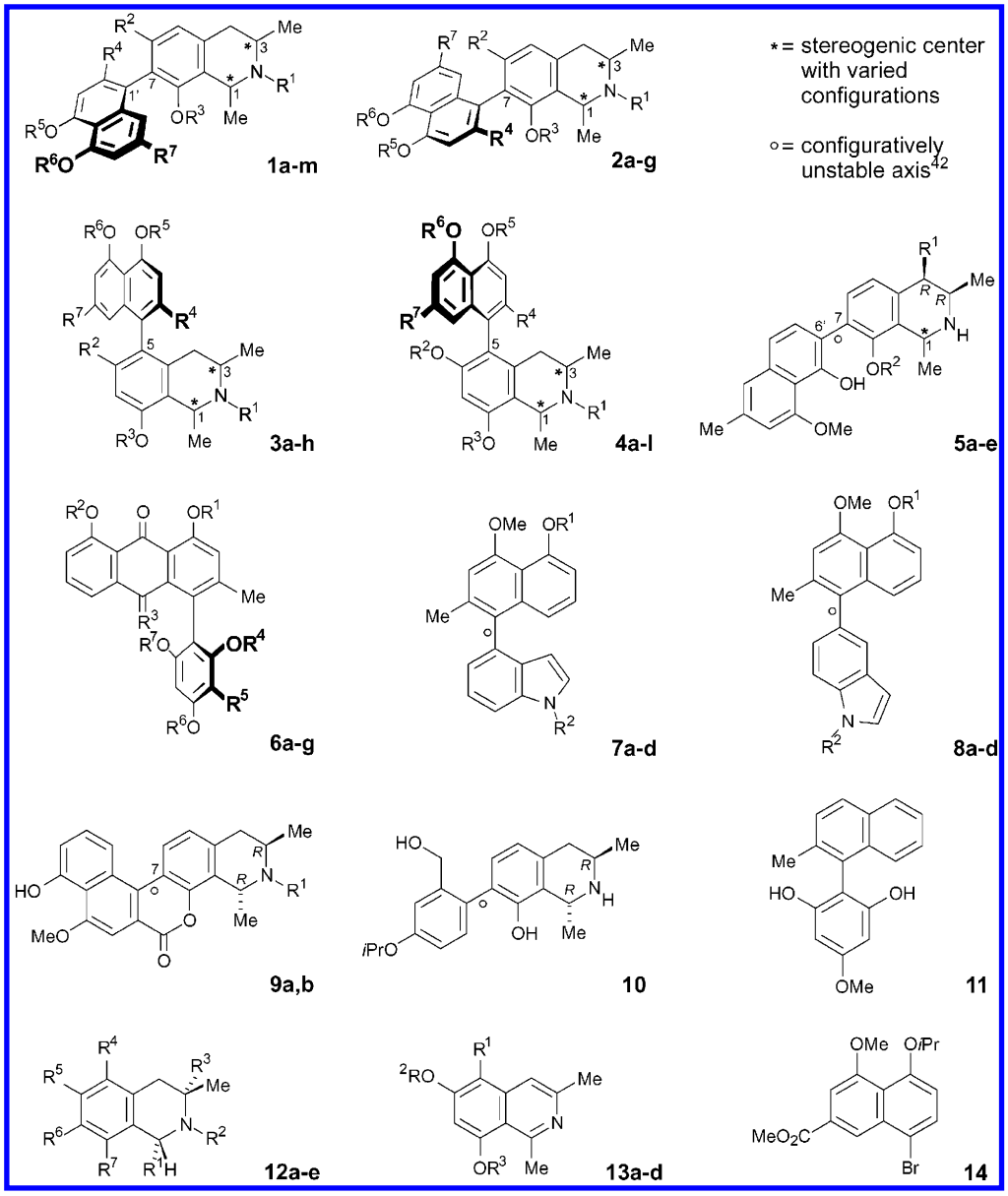

Figure 2. The compounds of the data set, consisting of 11 groups and three single structures (see Tables 1 and 2): naphthylisoquinoline alkaloids (groups **1**−**5**, 39 compounds), knipholone and derivatives (group **6**), synthetic naphthylindenes (groups **7** and **8**), lactone-bridged naphthylisoquinolines (group **9**), the phenylnaphthalene **11**, and "monomers" **12**-**14**.

naphthylisoquinoline alkaloids, contain rotationally hindered biaryl partial structures, whose axial chirality can be an important factor for bioactivity (see e.g. **1a** and **2a** in Table 6 or ref 28)—like chirality in general.[29] Even in the case of stable axial chirality, the remaining rotational mobility at the axis makes this central linkage between the two molecular portions a flexible element. Thus, the compounds dealt with in this paper occupy a large conformational space, so that the structures are difficult to align.

For the fully automated, flexible alignment,[30,31] numerous approaches have been developed to generate, align, and evaluate conformers of drug-size molecules, such as MIMUMBA[32] and, more recently, DISTILL.[33] Even though significant progress has been made in this field, there is possibly no alignment procedure that can fully replace the intuition of a human brain and the chemical knowledge of the system. Thus, every result automatically achieved has to be individually checked prior to the actual molecular field calculation.

In this work we compared two different approaches to compute an alignment for highly flexible structures. Both were found to require only slight manual refinement to provide valuable results. The first program was FLEXS, introduced by Lemmen et al.[34,35] It uses a fragment-based algorithm initially developed for the popular docking program FLEXX.[36,37] The second procedure was GASP ("**G**enetic **A**lgorithm **S**imilarity **P**rogram"), developed by Jones et al.[38] This program uses a genetic algorithm to produce multiple conformers, which are then checked for molecular similarities for superimposition. Both programs are implemented in the molecular modeling package SYBYL.[33,39]

MATERIAL AND METHODS

The chosen data set includes approximately 40 naphthylisoquinoline alkaloids with different coupling positions (divided into groups **1**−**5**, see Figure 2 and Table 1), the phenylanthraquinone knipholone (**6b**),[40] and four derivatives

Table 1. Naphthylisoquinoline Alkaloids in Groups **1**−**5** (for Basic Structures See Figure 2)

no.	trivial name	IC_{50}[a]	cc[b]	R^1	R^2	R^3	R^4	R^5	R^6	R^7
1a	dioncopeltine A[11,12]	1.880	1*R*/3*R*	H	H	H	CH_2OH	Me	H	H
1b	ancistrocladisine	−0.236	3*S*[c]		OMe	Me	Me	Me	Me	H
1c	*cis*−1,2-dihydroancistrocladisine	−0.237	1*R*/3S	H	OMe	Me	Me	Me	Me	H
1d	*trans*−1,2-dihydroancistrocladisine	−0.265	1S/3S	H	OMe	Me	Me	Me	Me	H
1e	ancistrocongoline D[d,43]	−0.687	1*R*/3*R*	H	OH	Me	Me	Me	Me	H
1f	ancistrogriffine A[d]	0.729	1*S*/3*S*	H	OH	Me	H	Me	Me	Me
1 g	ancistrobertsonine D[44]	−0.639	1*R*/3*S*	H	OH	Me	Me	Me	Me	H
1h	*N*-benzyldioncopeltine A	0.299	1*R*/3*R*	Bn[e]	H	H	CH_2OH	Me	H	H
1i	*N*−5′-*O*-dibenzyldioncopeltine A	−0.001	1*R*/3*R*	Bn	H	Bn	CH_2OH	Me	H	H
1j	habropetaline A[45]	1.896	1*R*/3*R*	H	H	H	CH_2OH	Me	Me	H
1k	dioncophylline A[11,1]	0.419	1*R*/3*R*	H	H	H	Me	Me	Me	H
1l	5′-*O*-demethyldioncophylline A[46]	0.381	1*R*/3*R*	H	H	H	Me	Me	H	H
1m	4-*O*−5-*O*-didemethyldioncophylline A	−0.343	1*R*/3*R*	H	H	H	Me	H	H	H
2a	7-*epi*-dioncopeltine A	0.538	1*R*/3*R*	H	H	H	CH_2OH	Me	H	H
2b	ancistrogriffine C[d]	−0.138	1*S*/3*S*	H	OMe	H	H	H	Me	Me
2c	*N*-benzyl-7-*epi*-dioncopeltine A	0.257	1*R*/3*R*	Bn	H	H	CH_2OH	Me	H	H
2d	7-*epi*-dioncophylline A[47]	0.516	1*R*/3*R*	H	H	H	Me	Me	Me	H
2e	*N*-methyl-7-*epi*-dioncophylline A	−0.172	1*R*/3*R*	Me	H	H	Me	Me	Me	H
2f	dioncophylline D[f,48]	0.797	1*R*/3*R*	H	H	H	H	H	Me	Me
2 g	*d*	−0.369	1*S*/3*S*	Me	OMe	Me	H	Me	Me	Me
3a	dioncophylline C[47]	1.782	1*R*/3*R*	H	H	H	Me	Me	H	H
3b	hamateine	−0.261	*g*		H	Me	Me	Me	Me	H
3c	ancistrocongoline C[d]	−0.853	1*R*/3*R*	Me	OH	Me	H	Me	Me	Me
3d	ancistrolikokine B[49]	0.226	1*S*/3*R*	H	OH	Me	H	Me	H	Me
3e	ancistrolikokine C[49]	−0.356	1*R*/3*R*	Me	OH	Me	H	Me	H	Me
3f	ancistrobertsonine B[44]	−0.868	1*R*/3*S*	Me	OMe	Me	Me	Me	Me	H
3 g	*N*-methyldioncophylline C	0.453	1*R*/3*R*	Me	H	H	Me	Me	H	H
3h	korupensamine B[50]	0.936	1*R*/3*R*	H	OH	H	H	H	Me	Me
4a	ancistrocladine[11,12]	−0.542	1*S*/3*S*	H	H	Me	Me	Me	Me	H
4b	*N*-methylancistrocladine	−0.329	1*S*/3*S*	Me	H	Me	Me	Me	Me	H
4c	ancistrocongoline A[43]	0.300	1*R*/3*R*	Me	H	H	H	H	Me	Me
4d	ancistrocongoline B[d,43]	0.423	1*R*/3*R*	H	Me	H	H	Me	Me	Me
4e	ancistroealaine A[7]	−0.401	3*S*[c]	-	Me	Me	H	Me	Me	Me
4f	ancistroealaine B[7]	−0.102	1*S*/3*S*	H	Me	Me	H	H	Me	Me
4g	ancistrolikokine A[49]	0.464	1*R*/3*R*	Me	Me	H	H	H	Me	Me
4h	ancistrolikokine D	−0.305	3*S*[c]		H	Me	H	Me	H	Me
4i	ancistrobertsonine A	−1.051	1*S*/3*S*	Me	H	Me	H	Me	Me	Me
4j	ancistrobertsonine C[44]	−0.347	1*R*/3*S*	Me	Me	Me	H	Me	Me	Me
4k	ancistrocladeine	−0.277	*g*		H	Me	Me	Me	Me	H
4l	korupensamine A[50]	0.738	1*R*/3*R*	H	H	H	H	H	Me	Me
5a	dioncophyllinol B[48]	1.074	1*R*	OH	H					
5b	1-*epi*-dioncophylline B[48]	0.370	1*S*	H	H					
5c	8-*O*-methyl-1-*epi*-dioncophylline B[48]	1.431	1*S*	H	Me					
5d	dioncophylline B[11,12,48]	0.636	1*R*	H	H					
5e	8-*O*-methyldioncophyllinol B[48]	−0.221	1*R*	OH	Me					

[a] -log(IC_{50}). [b] Chiral center(s). [c] Double bond between C1 and N. [d] Test set compounds. [e] Bn = benzyl. [f] Axis freely rotating at room temperature.[48] [g] Isoquinoline part fully unsaturated.

thereof (Table 2), naphthylindenes that had been predicted to be active in earlier studies[41] (groups **7** and **8**), and synthetic precursors such as lactone-bridged naphthylisoquinolines **9**, but also simple isoquinolines, **12** and **13**, and the naphthalene derivative **14**. Since these molecules differ largely in size, difficulties in aligning the data set resulted. Therefore we treated the much smaller "monomer"-type structures of groups **12**, **13**, and **14** separately.

Biological Data. The preparation of the compounds studied in this work, by isolation or by chemical synthesis, has been described in detail elsewhere.[16,17,52] Antiplasmodial activity was determined using the K1 strain (resistant to chloroquine and pyrimethamine), for which a modification[53] of the [^{3}H]-hypoxanthine incorporation assay[54] was used. Thus, infected human red blood cells were exposed to serial drug dilutions in microtiter plates for 48 h at 37 °C in a gas mixture with reduced oxygen and elevated CO_2. After addition of [^{3}H]-hypoxanthine to each well and further incubation for 24 h the wells were harvested on glass fiber filters and counted in a liquid scintillation counter. From the sigmoidal inhibition curves the IC_{50} values were calculated. The assays were run in duplicate and repeated at least once. For the QSAR calculations -log(IC_{50}) values were used, with the IC_{50} values expressed in nmol/mL.

Structure Preparation. Structure minimization and CoMSIA studies were performed on Silicon Graphics OCTANE R10000 and O2 R5000 computers using the molecular modeling package SYBYL 6.7.1.[33] All structures were minimized with the Tripos force field, partial atomic charges were added using the Gasteiger-Hückel method.[55,56] FLEXS uses CORINA to generate ring conformers (see below), while GASP does not alter closed ring systems. Therefore tetrahydroisoquinoline conformers were manually checked for unfavorable conformations after force-field minimization.

Alignment Rules. The molecular alignment is by far the most critical parameter in a molecular field analysis. The outstanding structural element of naphthylisoquinoline alkaloids is the biaryl axis, which, however, cannot be used directly for aligning because of the existence of different coupling positions both in the tetrahydroisoquinoline portion

Table 2. Structure Groups **6**−**14** (Figure 2)

no.	trivial name	-log(IC_{50})	R^1	R^2	R^3	R^4	R^5	R^6	R^7
6a	bulbine-knipholone[a,51]	−0.238	H	H	O	H	Ac	H	Me
6b	knipholone[40]	−0.189	H	H	O	H	Ac	Me	H
6c	knipholone-anthrone	0.450	H	H	H,H	H	Ac	Me	H
6d	4-*O*-demethylknipholone	−0.071	H	H	O	H	Ac	H	H
6e	6-*O*-methylknipholone	−0.374	H	H	O	H	Ac	Me	Me
6f		−0.178	*i* Pr	*i* Pr	O	Me	H	Me	H
6g		−0.121	*i* Pr	*i* Pr	O	Me	H	Me	Ac
7a		0.935	*i* Pr	Tos[b]					
7b		0.375	H	Tos					
7c		−0.080	*i* Pr	H					
7d		−0.093	H	H					
8a		0.473	*i* Pr	Tos					
8b		0.299	H	Tos					
8c		0.036	*i* Pr	H					
8d		−0.060	H	H					
9a	dioncolactone A	−0.085	H						
9b	*N*-benzyldioncolactone A	−0.392	Bn[c]						
10		1.454							
11		−1.070							
12a		0.363	Me	Bn	H	H	H	H	OCH_2OCH_3
12b		0.047	Me	Bn	H	H	H	Br	OCH_2OCH_3
12c		−0.652	H	Bn	Me	H	OMe	H	OMe
12d		−0.321	H	H	Me	Br	OMe	H	OMe
12e		−0.036	H	Bn	Me	Br	OMe	H	OMe
13a		−0.101	H	H	H				
13b		−0.647	Br	Bn	Bn				
13c		−0.857	Ph	H	H				
13d		−0.617	H	Bn	Bn				
14		−0.103							

[a] Test set compound. [b] Tos = *p*-toluenesulfonyl. [c] Bn = benzyl.

and in the naphthalene part of the molecules. On the other hand, the coupling site is the deciding factor for the spatial arrangement of the potential pharmacophoric elements to each other and is therefore important for the biological activity. In this investigation, two commercially available alignment approaches were evaluated, FLEXS and GASP, which both should be able to produce an alignment of highly flexible compounds. It turned out that neither succeeded in treating axial chirality correctly by itself. Despite this drawback, both eventually produced surprisingly good results after some manual refinement.

FLEXS Alignment.[34,35] With FLEXS a "reference ligand" has to be selected, preferably in in its bioactive conformation. The structure of this reference ligand is used as a template for all other compounds in the data set and is kept fixed during the complete alignment procedure, while the remaining structures ("test ligands") are decomposed into molecular fragments at each rotatable bond; for closed, flexible ring systems, FLEXS uses the external program CORINA[57,58] to generate a set of valuable conformations; the same procedure is also used by FLEXX. After splitting the test ligand, the algorithm places a chosen "base fragment" into a suitable region of the reference ligand and then attaches all remaining fragments in a step-by-step manner. A scoring function is applied to rank the different alignments produced (for details on the procedure and the score evaluation, see ref 34).

The most active compound in the data set, dioncopeltine A (**1a**), was selected as the reference ligand. After the input structures had been prepared, the alignment process was fully automatic. In addition to the standard procedure (as recommended by Tripos, Inc.) we used the "OPTIMIZE" command, which enhances the Gaussian overlap volume after the actual alignment. The solutions were then sorted using the "SORT E_TOTAL" command, which is necessary for the reevaluation of the scoring function after optimizing the placements. The internal clustering algorithm was used with given default values to reduce the number of output structures.

As mentioned, the FLEXS algorithm is not capable of retaining the correct configuration of axially chiral compounds. Since biological activities of compounds that differ only by their axial configuration can vary by up to more than one logarithmic unit (e.g. **1a** and **2a**, see Table 6), this led us to separate out incorrectly configured aligned structures manually after the superposition process. Thus, the highest-ranking compound among the correctly configured ones was selected. This alignment is henceforth referred to as "FLEXS alignment".

An analysis of the remaining alignment possibilities proposed by FLEXS showed that there are generally at least three reasonable ways for naphthylisoquinoline alkaloids to be aligned with each other, independent of the axial configuration. In detail the three modes can be categorized as follows: **a**: The *N*-atoms of the isoquinoline parts are aligned onto each other and so are the two *O*-functions in positions 4′ and 5′ of the naphthalene parts. **b**: The naphthalene parts of the test ligands are aligned onto the isoquinoline part of the reference ligand and *vice versa*. **c**: The test ligands are turned by 180°, as compared to **a**, which means that the *N*-functions are not aligned any more, while the *O*-functions still are.

Figure 3 shows these three possible alignments, exemplarily for dioncophylline C (**3a**) and ancistrocongoline A (**4c**), each with dioncopeltine A (**1a**) as the reference ligand. If the axial configurations of the test ligand and the reference molecule are identical,[59] as is the case for dioncophylline C

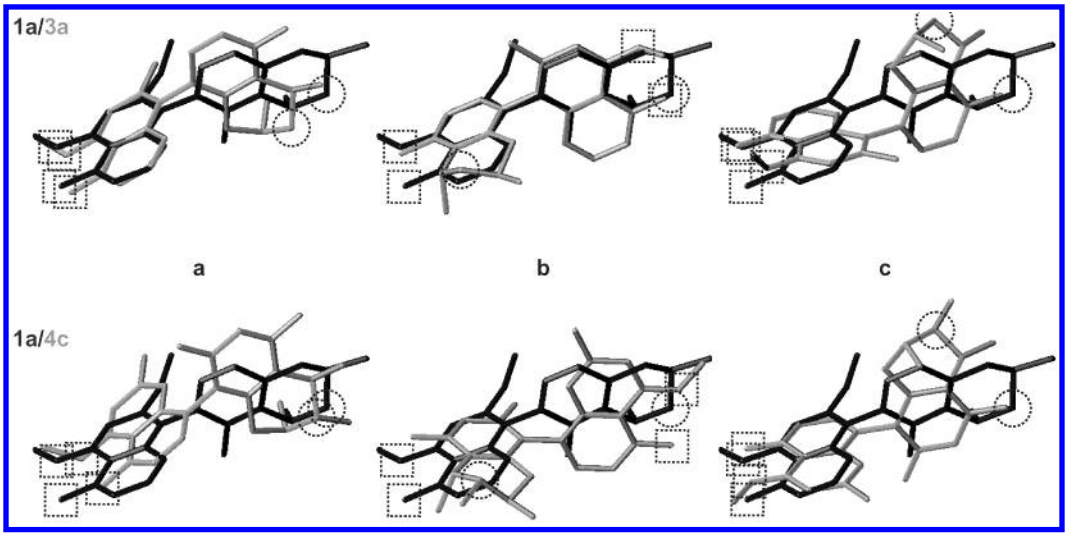

Figure 3. Three different FLEXS alignments of dioncophylline C (**3a**, top half, light) with dioncopeltine A (**1a**, dark) and of ancistrocongoline A (**4c**, lower half, light) with dioncopeltine A (dark); the circles accentuate the nitrogen atom in the isoquinoline part (hydrogens have been omitted for clarity), and the dotted squares the oxygen functions on the naphthalene portion.

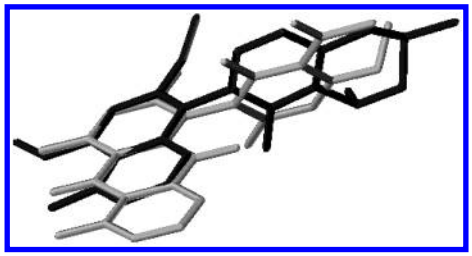

Figure 4. The phenylanthraquinone-naphthylisoquinoline alignment: knipholone (**6b**, light) vs dioncopeltine A (**1a**, dark).

(**3a**) and dioncopeltine A (**1a**, upper half of Figure 3), the alignment seems to fit somewhat better at first glance. But even if this is not the case, as for ancistrocongoline A (**4c**) and dioncopeltine A (**1a**, lower half of Figure 3), all three ways are still possible.

The described procedure resulted in approximately 10−50 correctly configured FLEXS-alignment possibilities per test ligand. For almost all naphthylisoquinoline alkaloids these proposals included the three possibilities shown in Figure 3, independent of whether the axial configuration was the same or not. As the ranking of these three possibilities varied between the different test ligands, none of them could be identified as generally advantageous over the others.

According to our experience, an *N*-alkylation of the isoquinoline part of naphthylisoquinoline alkaloids usually decreases the in vitro antimalarial activity dramatically, which led to the idea of manually altering the original FLEXS alignment by omitting all alignments different from **a**. This alignment is referred to as FLEXS(II). Compared to the initial FLEXS approach, the alignment of nine naphthylisoquinoline alkaloids (**3d**, **3e**, **3f**, **3h**, **4a**, **4b**, **4e**, **4i**, and **4j**) had to be changed. The alignments of the structures not belonging to the naphthylisoquinoline alkaloids (**6**−**14**) were left unchanged.

Despite the obvious structural differences of knipholone and its derivatives (group **6** in Figure 2) as compared to the other compounds in the data set, their molecular alignment was surprisingly straightforward. The anthraquinone part was placed upon the naphthalene portion of dioncopeltine A (Figure 4), and the phenyl ring was positioned onto the aromatic part of the isoquinoline moiety, letting the oxygen substitution patterns of the two molecules correspond perfectly.

Still, a remaining problem was the fact that, due to the various substitutions patterns of naphthylisoquinoline alkaloids, there was no compound that contained all possible structural elements (like OH or OMe groups in all respective positions) so that it could have been used as a reference for all the other compounds. For instance, if the hydroxy function at C5′ of the reference ligand **1a** (marked by a circle in Figure 5) was replaced by a methoxy group in the test ligand, there was no counterpart in the reference ligand during the alignment process. As a result of the use of **1a** as the reference ligand, a methoxy group in this position was not correctly matched in the final FLEXS(II) alignment (Figure 5). To avoid unnecessary noise the corresponding torsional angles were manually adjusted, as also in similar cases (e.g. for the 8-oxygen function, see Figure 5)

GASP Alignment.[38,60] GASP uses a genetic algorithm for superposition of a set of flexible molecules. The evolution of better models is evaluated by a fitness function which takes into account the protein−ligand interactions of different acceptors and donors. Molecules are represented by a chromosome encoding conformational information and intermolecular mappings between analogous important structural features that may be required for activity, in pairs of molecules (i.e., H-bond donor proton, acceptor lone pair, and ring center). The conformational data includes angles of rotation around flexible bonds.

The fitness of an alignment is characterized by a combination of three criteria: the number and similarity of overlaid features, the volume overlap of the alignment (i.e., the common volume of the aligned molecules), and the internal van der Waals energy of the molecular conformations (for a detailed description of the procedure, see references 38 and 60).

It is important to bear in mind that GASP aligns *pharmacophoric* features previously identified in the ligands and not necessarily *structural* similarities. The program has initially been designed to confirm or reject a particular pharmacophoric hypothesis. Molecular field analysis methods (CoMFA and CoMSIA), by contrast, are known to react very

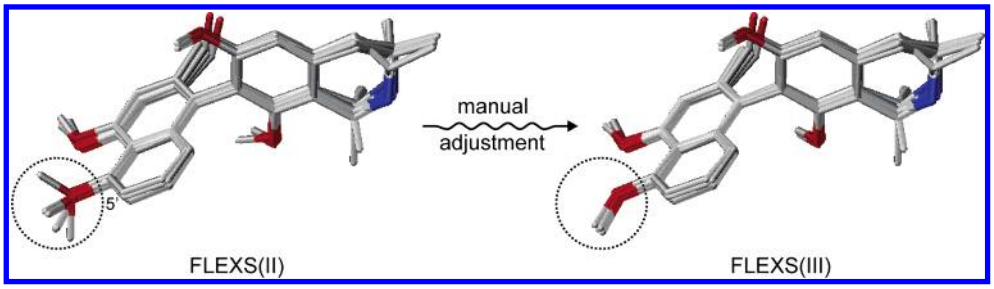

Figure 5. Incorrect alignment (see dotted circle) of dioncopeltine A (**1a**, OH at C5′) with several other, likewise 7,1′-coupled naphthylisoquinolines (with OMe at C5′), necessitating a manual adjustment.

sensitively even to minimal structural differences, which makes it necessary to align common structural elements as precisely as possible. The structural diversity of naphthylisoquinoline alkaloids, however, resulting mainly from different coupling positions and from the phenomenon of axial chirality, makes it impossible to pay attention to all such structural elements simultaneously, so that the GASP approach should possibly still be suited for the alignment problem.

Another option to avoid the difficulties in aligning different coupling types might be to develop individual models for each structural subtype of the naphthylisoquinoline alkaloids, but this would also lead to small data sets with only few structures each. The major drawback would be only small ranges in the biological activity in some of the groups. This is already a problem with the entire data set, since it currently spans only 2.95 logarithmic units, which is just sufficient to generate valuable QSARs.[61] Moreover, similarities between the compounds in the data set make it highly likely that all have a similar mode of action and should therefore be treated as one single data set.

Different from FLEXS, GASP in principle allows the treatment of more than two molecules at a time, but state-of-the-art computer systems are not able to treat a data set with more than 60 — in part highly flexible — molecules in one single GASP run. Taking that into consideration, the initial idea was to generate a "basic alignment" of the two most active compounds, the 5,8′-coupled alkaloid dioncophylline C (**3a**) and dioncopeltine A (**1a**), which is 7,1′-coupled (see Figure 1), without any constraints except for the axial configuration. This should have led to an alignment very likely to be similar to the active conformation within a possible receptor. The resulting conformations of dioncophylline C and dioncopeltine A could then be used as templates for all of the remaining structures. Within GASP (since version 2.0) it is possible to define molecular constraints to input structures, which would have given us the opportunity to take into account the phenomenon of axial chirality without the need to manually sort out incorrectly configured compounds.

Already this "basic" alignment run, with only **3a** and **1a**, however, showed that there are unfortunately still problems with using constraints in GASP. To keep axial chirality intact, constraints were initially set for both molecules at their axis of $30° < \gamma < 150°$ for the torsion angles γ (see Figure 6). This resulted in an alignment with torsional angles very close to the constraints set, viz. about 30° to 35° for γ. Loosening the constraint to $20° < \gamma < 160°$ led to torsional angles in the aligned structures of about 23 for γ. However if we did not define any constraint at all, the GASP run resulted in

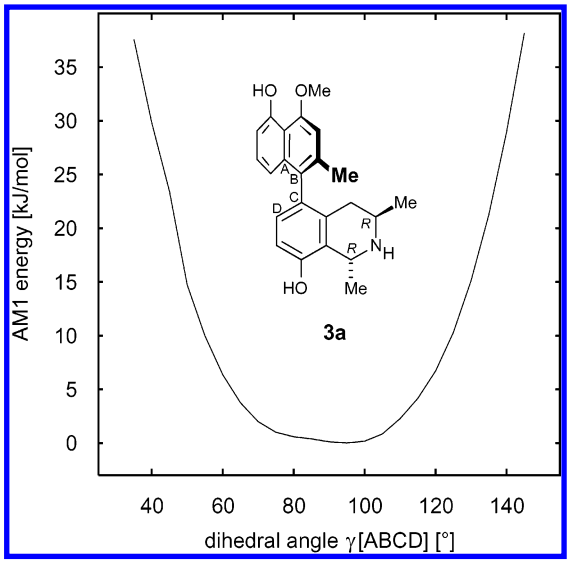

Figure 6. The change in potential energy during (restricted) rotation around the biaryl axis, exemplarily for dioncophylline C (**3a**); the AM1 energies were calculated, varying the dihedral angle [ABCD] in steps of 5°, and are given in kJ/mol relative to the minimum at 95°.

alignments with torsional angles in both molecules of $\gamma \sim 75°$, i.e., close to the thermodynamic minimum (Figure 6), which is therefore supposed to be the best possibility. Unfortunately, without constraints, alignments with wrong axial configurations were also produced.[62]

The only solution to the problem was to fix the torsional angles at the biaryl linkage of axially chiral compounds to 90°, which, in view of the rotational flexibility of naphthylisoquinoline alkaloids at the biaryl axis, is of course a drastic limitation. Using again dioncopeltine A (**1a**) as the template structure, three alignments per compound were generated with GASP and the one with the highest fitness score was used for CoMSIA calculations.

To detect and treat H-bond donors and acceptors correctly, the GASP algorithm needs to ionize specific atom types, i.e., it adds a proton (H^+) to the SYBYL atom type N.3 (like in a tetrahydroisoquinoline). In a similar way lone pairs are added to hydroxy and methoxy groups as well as dummy atoms for the identification of ring centroids during the alignment process. All these additional atoms were manually removed directly after the GASP run, to make sure they would not interfere with the subsequent molecular field calculations.

Calculation of the CoMSIA Fields. Steric, electrostatic, and hydrophobic as well as H-bond donor and acceptor similarity indices $A_{f,k}$ were calculated at each lattice intersec-

tion of a regular-spaced grid of 2.0 Å by using the formula[26]

$$A^{q}_{f,k}(j) = -\sum_{i} \omega_{\text{probe},k}\omega_{ik}e^{\alpha r^2_{iq}}, \quad (1)$$

where i = summation index over all atoms of the molecule j under investigation; ω_{ik} = actual value of the physicochemical property k of atom i; $\omega_{\text{probe},k}$ = probe atom with charge +1, radius 1 Å and hydrophobicity +1; and r_{iq} = mutual distance between probe atom at grid point q and atom i of the test molecule. The attenuation factor α was initially set to the recommended default value of 0.3; optimum values have been reported to be between 0.2 and 0.4.[63]

The CoMSIA region was automatically defined, extending the van der Waals volume of all molecules in x, y, and z directions by 4 Å. During the cross-validation analysis, column filtering was set to 2.0 kcal/mol.

PLS Analysis. For the correlation between molecular field descriptor values and biological activities the method of partial least-squares (PLS) regression[61,64] was used. To obtain the optimum number of principal components the leave-one-out (LOO) cross-validation was utilized. While this is the most commonly applied procedure,[61] LOO cross-validation is usually considered to be a measure of how good the model represents the data in the training set. In particular for most diverse data sets, this may not be sufficient to evaluate the prediction capability of an obtained model. In LOO every compound is omitted once and predicted to yield a q^2_{LOO} value, whereas in an n-fold-cross-validation the data set is first randomly divided into groups with approximately equal numbers of compounds in each group; these groups are then sequentially omitted and predicted using a model based on the remaining compounds to calculate q^2_n. Because the selection of the left-out structures is random in each run, it is necessary to do an n-fold-cross-validation multiple times, and the standard deviation (SD) of all runs can be taken to evaluate the overall stability of the model.[65] In this work, q^2_n was calculated for $n = 3$, 5, and 10. Every run was repeated 100 times.

RESULTS AND DISCUSSION

Training and Test Sets. Usually the complete data set is divided into a training and a test set; the latter is used to evaluate the final results after model derivation. Because our group continuously tests new compounds for antimalarial activity, we used all of the substances available at the beginning of this study to derive the model and have then used all new compounds tested during this work as the test set. This test set includes two compounds each from groups **1** and **2** and one compound each from groups **3**, **4**, and **6** (these substances have been labeled accordingly in Tables 1 and 2).

Outliers. There are several reasons for structures to appear as outliers, the most likely being an incorrect experimental value, but a different bioactive conformation or even an alternative binding mode is also possible. Our training set, containing 57 biarylic compounds (structures **12**−**14** were omitted at that time, because of their different sizes), was checked for outliers. After the first PLS runs it was found that two structures from group **6** have the largest residuals (≥1), independent of the alignment approach used (FLEXS or GASP):

Table 3. Results of the LOO Cross-Validation

alignment		biaryl compounds	all compounds
FLEXS	q^2	0.623	0.555
	N^a	8	6
FLEXS(II)	q^2	0.787	0.704
	N	9	8
FLEXS(III)	q^2	0.783	
	N	8	
GASP	q^2	0.686	0.600
	N	5	6

[a] Optimum number of PLS components.

- 8-*O*-methyl-1-*epi*-dioncophylline B (**5c**).
- 8-*O*-methyldioncophyllinol B (**5e**).

Both compounds are 7,6′-coupled naphthylisoquinoline alkaloids (see Figure 2 and Table 1) and differ only by a hydroxy group at C4 and the configuration at C1. The fact that the former is the most and the latter the least active compound in group **4** supports the assumption that the whole group may be wrongly aligned. However, none of the alternatives that we tried (similar to those in Figure 3) led to significant improvements in PLS analysis (data not shown).

Two further structures had to be omitted for the following reasons:

- 4-*O*-5-*O*-Didemethyldioncophylline A (**1m**) was synthesized by BBr_3-promoted *O*-demethylation of dioncophylline A (**1k**); it was found to be highly oxygen-sensitive and purification was thus very difficult; moreover it is questionable if the tested sample included only the correct atropisomer.[66] The model considers it to be more active than has been found experimentally, the reason for which might be that the substance had already partially decomposed before or during the testing process.
- Knipholone-anthrone (**6c**) is the only compound from group **6** with a reduced (hydrodeoxygenated) anthraquinone system and has, in contrast to the "true" phenylanthraquinones, proven to be quite cytotoxic in vitro.[67]

PLS Analysis. The results of the PLS analyses of the four alignment approaches including all five CoMSIA fields are summarized in Tables 3 (LOO cross-validation) and 4 (LMO cross-validation) for both, the data set containing only biaryl compounds (53 molecules), and for the data set including all 63 structures.

The results did not show any of the two alignment approaches to be generally superior to the other. However, the manual adjustment of the FLEXS alignment with respect to the isoquinoline nitrogen clearly improved the q^2 value from 0.623 to 0.787 (LOO), which underlined the importance of this atom for activity. Considering the rigorous constraints that had to be used with respect to the dihedral angle at the biaryl axis within the GASP alignment, a q^2 of 0.686 was surprisingly good, even better than the initial, unaltered FLEXS result. Hopefully, future versions of GASP may include improved constraint definitions and may therefore offer the possibility of superior automated alignments of biaryl compounds.

The automatic treatment of structures that correspond only to one biaryl half (groups **12**−**14**) did not lead to a general alignment rule because the substitution patterns were too inhomogeneous and the difference in shape compared to naphthylisoquinoline alkaloids was too large. Especially

Table 4. Results of the *n*-Fold-Cross-Validation

		biaryl compounds			all compounds		
alignment	*n*	10	5	3	10	5	3
FLEXS	mean q_n^{2a}	0.606	0.572	0.518	0.564	0.537	0.498
	highest q_n^2	0.685	0.685	0.673	0.646	0.647	0.637
	lowest q_n^2	0.481	0.410	0.253	0.463	0.369	0.154
	SD[b]	0.039	0.058	0.089	0.036	0.055	0.088
FLEXS(II)	mean q_n^2	0.757	0.714	0.651	0.675	0.640	0.579
	highest q_n^2	0.812	0.804	0.789	0.732	0.731	0.733
	lowest q_n^2	0.631	0.517	0.353	0.590	0.138	0.231
	SD	0.035	0.055	0.080	0.031	0.066	0.090
FLEXS(III)	mean q_n^2	0.749	0.720	0.651			
	highest q_n^2	0.799	0.806	0.794			
	lowest q_n^2	0.523	0.507	0.382			
	SD	0.037	0.051	0.083			
GASP	mean q_n^2	0.668	0.630	0.577	0.600	0.554	0.496
	highest q_n^2	0.735	0.757	0.742	0.663	0.672	0.635
	lowest q_n^2	0.538	0.373	0.211	0.404	0.298	0.277
	SD	0.040	0.069	0.086	0.040	0.063	0.086

[a] Mean of 100 runs. [b] Standard deviation.

benzyl substituted structures led to varying alignments. As an example, FLEXS and GASP did not prefer the alignment of an isoquinoline half (e.g. **12a** or **12b**) onto the corresponding portion of the template molecule (dioncopeltine A, **1a**), but placed it in the region of the biaryl axis. The additional use of groups **12**−**14** for model derivation only resulted in a slight degradation of q^2 from 0.787 to 0.704 with the FLEXS(II) alignment and from 0.686 down to 0.600 for GASP. However, even if groups **12** and **14** were included, a reliable prediction remains unlikely, due to the lack of a general alignment rule for these compounds. Therefore, groups **12** and **14** were excluded from further investigations.

The manually adjusted FLEXS(III) alignment did not show any better results than the FLEXS(II) approach (see Table 4): q^2 values did not improve significantly, which can have three reasons: Either the CoMSIA method is not sufficiently sensitive to distinguish between marginally varying torsional angles, or, second, it is indeed able to take into account even small structural changes, yet without the need for a totally exact alignment in this case. Third, it is also possible that the respective pharmacophoric elements are not essential for biological activity.

A comparison of GASP and FLEXS has to take into account several aspects:

- Just considering the mere calculation time, FLEXS is substantially faster than GASP. Depending on the ligands used, FLEXS provides between 50 and several hundred alignments, while the use of GASP needs about twice the time to calculate 4−5 alignments per test ligand.
- Since every single alignment has to be checked anyway, the larger number of possible alignments provided by FLEXS gives the user a much higher flexibility during the selection of the most promising ones.
- On the other hand, GASP can treat more than two compounds at a time; yet we have not been able to make use of that advantage because of the misleading constraint handling described earlier.

Taking all these findings into account, the FLEXS approach seems to be superior for our data set. The manual modification of the initial FLEXS alignment according to the isoquinoline-nitrogen atom clearly improved the model. The following tuning of the final QSAR model was based on the FLEXS(II) alignment.

Table 5: CoMSIA Models Based on Different Field Combinations (Final QSAR Model in Bold, for Field Definitions, See Figure 7)

	h d	s h d	e h d	h d a	s h d a	e h d a	all
q^2_{SAMPLS}	0.754	0.750	0.729	**0.813**	0.808	0.771	0.766
q^2	0.763	0.752	0.738	**0.818**	0.812	0.780	0.787
PRESS[a]	0.351	0.355	0.370	**0.308**	0.313	0.338	0.337
N[b]	8	7	8	**8**	8	8	9
r^2	0.978	0.970	0.982	**0.985**	0.979	0.982	0.987
s	0.107	0.124	0.097	**0.101**	0.104	0.096	0.084
F	243.6	208.2	298.9	**276.9**	257.5	304.1	357.8
fraction							
steric		0.122			0.066		0.048
electrostatic			0.352			0.238	0.223
hydrophobic	0.420	0.342	0.210	**0.234**	0.204	0.166	0.152
H-bond donor	0.580	0.537	0.437	**0.431**	0.411	0.350	0.337
H-bond acceptor				**0.334**	0.319	0.246	0.239

[a] Predictive error sum of squares. [b] Optimum number of PLS components.

Combination of Different CoMSIA Fields. It has been discussed whether the five different descriptor fields in CoMSIA are totally independent of each other;[63,68,69] such dependencies of the individual fields usually decrease the signal-to-noise ratio in the data[68] and lower the statistical significance of the results.[63] An evaluation which of the five CoMSIA fields are actually needed for a predictive model was performed by calculating all possible combinations of fields. The much faster SAMPLS[70] algorithm was used to reduce CPU-time.[71] Results are shown in Figure 7.

The first five models, using a single CoMSIA field, indicated that H-bonding interactions as well as the hydrophobic field are more important than the traditional steric and electrostatic CoMFA fields. The two best single fields used together gave rise to the best two-field model: hd ($q^2_{SAMPLS} = 0.754$, $N = 7$). The addition of the H-bond acceptor field led to the hda model (0.813, 9), thus providing the best overall model. The addition of the steric or the electrostatic field to hd yielded a reduced q^2_{SAMPLS} in both cases: shd (0.750, 10) and ehd (0.729, 7), which again underlined the negligible value of these two field types.

The results of the ordinary (i.e. not-SAMPLS) PLS analyses for the most promising field combinations, as shown in Table 5, confirmed the previous SAMPLS calculations. Most of the information from the molecular fields was already represented by the H-bond donor and the hydrophobicity field (hd model). However, the additional use of the H-bond acceptor field further enhanced the value of the model, leading to $q^2 = 0.818$. For these reasons, we consider hda to be the best possible combination.

Attenuation Factor. As mentioned above, the similarity indices are calculated using a Gaussian-type distance dependence.[26] The attenuation factor α (see eq 1) was initially set to the recommended[26] value of 0.3. Larger α values result in a stronger attenuation of the distance-dependence of similarity considerations and therefore local features become more important over global molecular similarity,[26] while smaller ones favor more global structure elements by a higher averaging of local features.[26] Hence, the optimum α-value is highly dependent on the actual problem, and we calculated q^2_{SAMPLS} for $0.1 < \alpha < 0.8$ in steps of 0.1 based on the

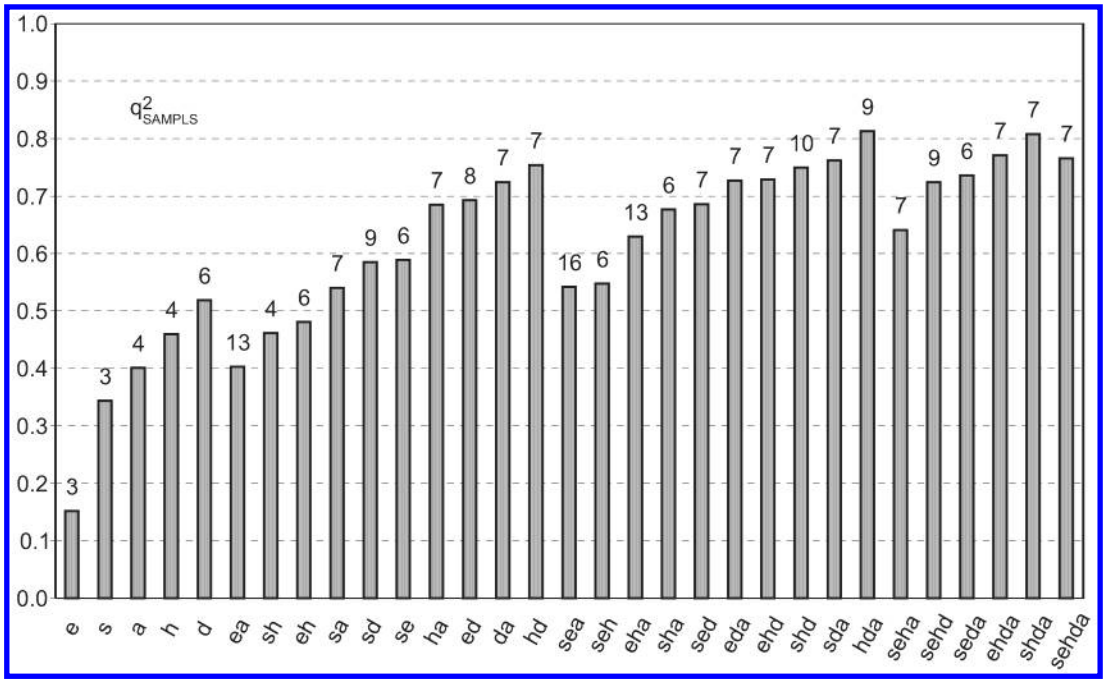

Figure 7. Results of the 31 possible CoMSIA field combinations (s = steric, e = electrostatic, h = hydrophobic, d/a = H-bond donor/acceptor), in the order of the number of fields per model, and q^2_{SAMPLS} (LOO cross-validation using the SAMPLS algorithm).[70] The numbers on top of the bars indicate the optimum number of pls components (*N*) for each analysis.

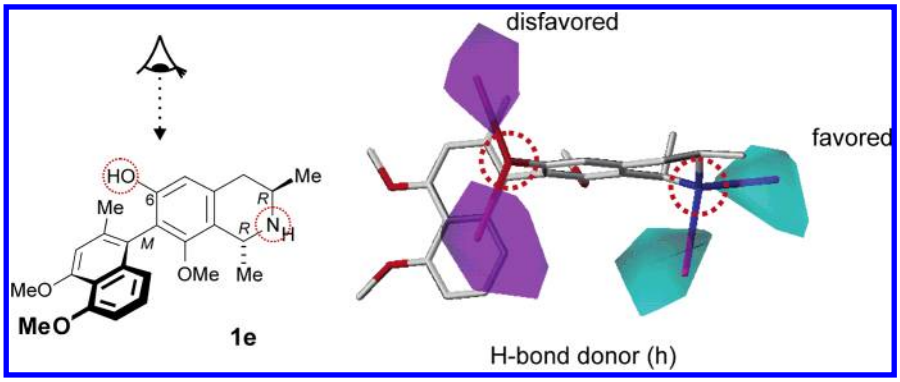

Figure 8. Ancistrocongoline D (**1e**) within the H-bond donor field (right) viewed from above as indicated on the formula (left). Areas where H-bond donor groups on the ligand are disfavored, are colored in blue, green bulks indicate H-bond donors to have a positive effect on biological activity. The red dotted circles marked the pharmacophoric elements corresponding to the contour plots (see text). Also shown are the dummy atoms used by CoMSIA to model a potential receptor site for specific groups, which are located directly inside the corresponding contour plots. The dummy atoms have been visualized using the "%comsia_info" expression generator within SYBYL.

FLEXS(II) alignment. In this case, however, only decreased q^2_{SAMPLS} values for both higher and lower values of α were obtained (data not shown).

Interpretation of CoMSIA Field Plots. The results of the best CoMSIA model (hda, $q^2 = 0.818$) were visualized by colored 3D contour plots. According to the contributions of the individual fields to the final, non-cross-validated hda model (Table 5), the H-bond donor field is obviously the most important one. Figure 8 shows the structure of ancistrocongoline D (**1e**) within the corresponding field plot. Additionally the dummy atoms created for the donor field calculation are shown. Compound **1e** was chosen because, in contrast to **1a** and **3a**, it includes both structural elements important to this field: One is obviously the *N*-atom in the isoquinoline moiety, surrounded by the H-bond donor favored contour plot (green-colored). These contours mark the areas of a potential target protein to which a possible ligand can donate hydrogens. The other element, illustrated by blue contours, visualizes areas that disfavor H-bond donors (Figure 8); these areas are attributable to a free hydroxy function at C6 of naphthylisoquinoline alkaloids, e.g. in **1e**, indicating this functional group to be disadvantageous for activity. To further demonstrate this dependence, the dummy atoms belonging to these two functional groups have been visualized (Figure 8) using the SYBYL "%comsia_info" expression generator;[72] they all point directly into the corresponding contour plots.

Figure 9 shows the alignment of dioncopeltine A (**1a**) and dioncophylline C (**3a**) within the hydrophobic (h) and the H-bond acceptor (a) plots. The h plot shows a large gray (hydrophobic-disfavored) region near the 2′-CH_2OH function of **1a**. This group is present in seven compounds of the data set, including three out of the four most active substances (**1a**, **1j**, and **10**) and the gray region clearly disfavors hydrophobic atoms and thus favors the *hydrophilic* properties of the hydroxy function. The same functional group is also considered as important in the H-bond acceptor plot, where it is surrounded by magenta contours. Similar to the H-bond

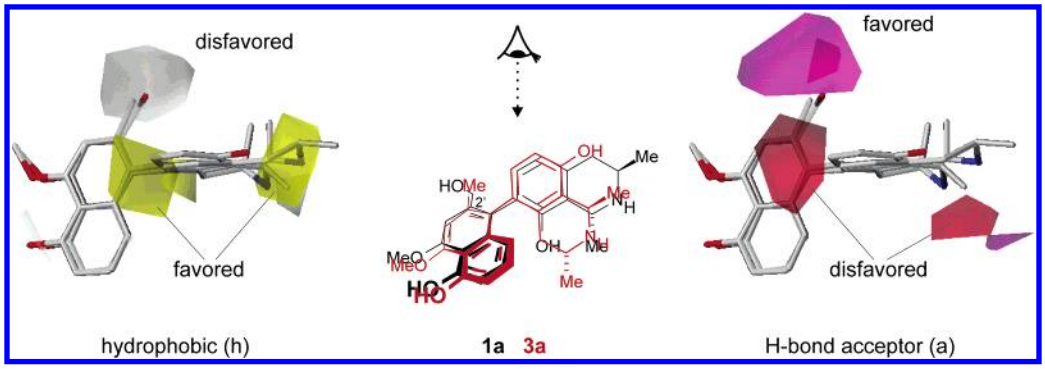

Figure 9. The alignment of dioncopeltine A (**1a**, black, center) and dioncophylline C (**3a**, red), placed within the hydrophobic (h, left) and the H-bond acceptor (a, right) plots.

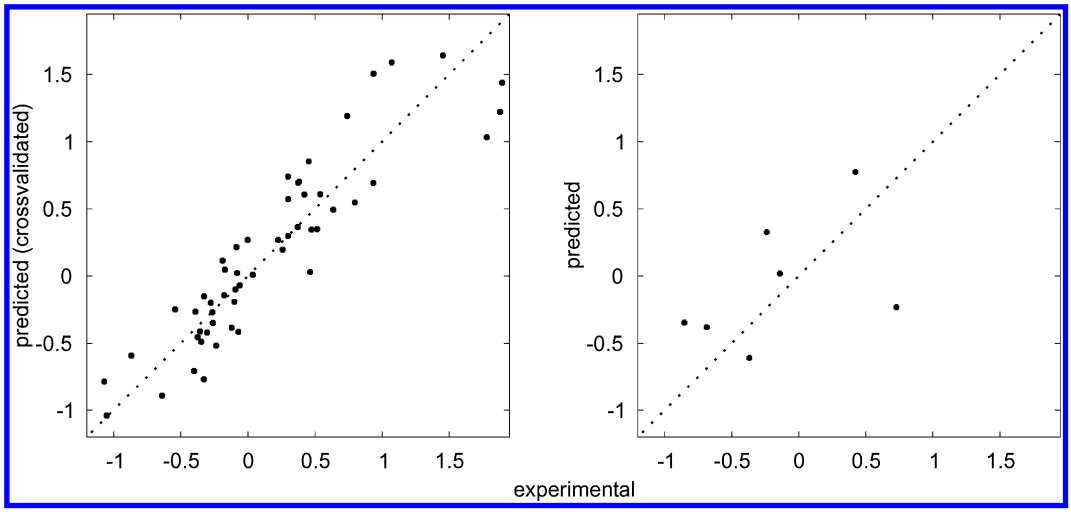

Figure 10. Graphical representation of the cross-validated predictions for the compounds in the training set (left) as well as the predicted values for the test set (right, see also Table 6).

donor field (Figure 8), position 6 in naphthylisoquinoline alkaloids also plays an important role in the hydrophobic and in the H-bond acceptor plots. The former favors a hydrophilic group, while the latter disfavors H-bond donors in this region.

In summary all three fields primarily support the presence of the *N*-atom in the isoquinoline moiety of the alkaloids. Making this group more exposed and thus easily accessible should be a rewarding synthetic goal presently approached by our group. Also important for an improvement of the activity are the findings concerning the 6-position: All of the three fields jointly indicate that a promising group should be (i) neither an H-bond donor (ii) nor an H-bond acceptor but should (iii) not contain hydrophobic atoms either. Demands (i) and (ii) are not contradictory, given the ampholytic character of the hydroxy function. According to all these findings such an OH group should have a strongly negative effect on antiplasmodial activity if located at C6.

Predicted Activities. While the CoMSIA contour plots qualitatively outline possible modifications of known compounds, another major use of a CoMSIA model is the prediction of the activity of new substances, to check new concepts for the design and synthesis of novel, synthetic compounds beforehand. Moreover, CoMSIA models can be used to select which compounds from an extract are the most promising ones and should then be isolated first, e.g. after their online detection and structural elucidation in plant extracts, using hyphenated methods such as LC-MS/MS-NMR-CD.[73−75]

In Table 6 and Figure 10 the (cross-validated) predictions of the compounds used for the final model derived, and the prediction of the test set compounds "unknown" to the QSAR, are shown.

Considering the problems during the alignment process of this very diverse data set, the final QSAR model using three CoMSIA fields represents the data in the training set very well, which is confirmed by a q^2 as high as 0.813 (hda model); looking at the predictions of compounds unknown to the model, two substances are not satisfactorily modeled: One of these is bulbine-knipholone (**6a**). Apparently the properties of the whole group of phenylanthraquinones are not sufficiently represented by the few compounds available (group **6**, Figure 2), although they seem to fit into the overall model. The second example is ancistrogriffine A (**1f**), which is not suitably predicted either, showing an unexpectedly high residual of 0.961 (possibly due to an incorrect experimental value).

All calculated activites result in a predictive r^2 value of 0.279 (for seven items), which is somewhat disappointing. Without taking **1f** and **6a** into consideration for the aforementioned reasons, the predictive r^2 value is raised to 0.578 (for five items). The experimental result for **1f** is presently being reevaluated.

Considering the alignments of different coupling types of naphthylisoquinoline alkaloids, however, one should expect the model to be able to predict structures out of groups **1**−**4** quite reasonably. Indeed, except for **1f**, all residuals ranged within ± 0.5 logarithmic units.

Table 6. Residues of Cross-Validated Predictions for All Compounds in the Training Set as Well as Of the Predicted -log(IC_{50}) Values for the Test Set (See Also Figure 10)

no.	trivial name	IC_{50}	residuals[a]	no.	trivial name	IC_{50}	residuals[a]
			Cross-Validated				
1a	dioncopeltine A	1.880	0.657	**4g**	ancistrolikokine A	0.464	0.433
1b	ancistrocladisine	−0.236	0.282	**4h**	ancistrolikokine D	−0.305	0.115
1c	*cis*−1,2-dihydroancistrocladisine	−0.327	−0.176	**4i**	ancistrobertsonine A	−1.051	−0.011
1d	*trans*−1,2-dihydroancistrocladisine	−0.265	0.004	**4j**	ancistrobertsonine C	−0.347	0.142
1g	ancistrobertsonine D	−0.639	0.253	**4k**	ancistrocladeine	−0.277	−0.079
1h	*N*-benzyldioncopeltine A	0.299	−0.441	**4l**	korupensamine A	0.738	−0.453
1i	*N*−5′-*O*-dibenzyldioncopeltine A	−0.001	−0.270	**5a**	dioncophyllinol B	1.074	−0.516
1j	habropetaline A	1.896	0.456	**5b**	1-*epi*-dioncophylline B	0.370	0.006
1k	dioncophylline A	0.419	−0.189	**5d**	dioncophylline B	0.636	0.142
1l	5′-*O*-demethyldioncophylline A	0.381	−0.321	**6b**	knipholone	−0.189	−0.303
2a	7-*epi*-dioncopeltine A	0.538	−0.072	**6d**	4-*O*-demethylknipholone	−0.071	0.343
2c	*N*-benzyl-7-*epi*-dioncopeltine A	0.257	0.061	**6e**	6-*O*-methylknipholone	−0.374	0.080
2d	7-*epi*-dioncophylline A	0.516	0.166	**6f**		−0.178	−0.034
2e	*N*-methyl-7-*epi*-dioncophylline A	−0.172	−0.221	**6g**		−0.121	0.263
2f	dioncophylline D	0.797	0.248	**7a**		0.935	0.242
3a	dioncophylline C	1.782	0.749	**7b**		0.375	−0.320
3b	hamateine	−0.261	0.086	**7c**		−0.080	−0.103
3d	ancistrolikokine B	0.226	−0.044	**7d**		−0.093	0.006
3e	ancistrolikokine C	−0.356	0.056	**8a**		0.473	0.128
3f	ancistrobertsonine B	−0.868	−0.276	**8b**		0.299	0.002
3g	*N*-methyldioncophylline C	0.453	−0.400	**8c**		0.036	0.027
3h	korupensamine B	0.936	−0.569	**8d**		−0.060	0.008
4a	ancistrocladine	−0.542	−0.293	**9a**	dioncolactone A	−0.085	−0.301
4b	*N*-methylancistrocladine	−0.329	0.438	**9b**	*N*-benzyldioncolactone A	−0.392	−0.127
4c	ancistrocongoline A	0.300	−0.273	**10**		1.454	−0.188
4e	ancistroealaine A	−0.401	0.305	**11**		−1.070	−0.285
4f	ancistroealaine B	−0.102	0.090				

no.	trivial name	IC_{50}	residuals[b]	no.	trivial name	IC_{50}	residuals[b]
			Predictions				
1e	ancistrocongoline D	−0.687	−0.306	**3c**	ancistrocongoline C	−0.853	−0.507
1f	ancistrogriffine A	0.729	0.961	**4d**	ancistrocongoline B	0.423	−0.351
2b	ancistrogriffine C	−0.138	−0.157	**6a**	bulbine-knipholone	−0.238	−0.565
2g		−0.369	0.240				

[a] Cross-validated. [b] Predicted by the final hda model.

CONCLUSIONS

In the present study we have established a 3D QSAR model for antiplasmodially active naphthylisoquinoline alkaloids, applying the CoMSIA method to a structurally diverse set of 53 biarylic compounds. These included mainly naphthylisoquinoline alkaloids of different coupling types, but also phenylanthraquinone structures and naphthylindenes, all of which could be fitted into one single model. Compounds of smaller size (which were available mainly from synthetic steps prior to biaryl coupling) could not be aligned unequivocally, which is the reason they were omitted in the final QSAR model. The alignment of the phenylanthraquinone structures (group **6**) was straightforward, but the prediction of their activities has to be treated with caution. It is possible that they just fit into the model because of the low degree of divergence of their biological activities.

Two automated alignment approaches were successfully applied; both FLEXS and GASP were able to generate highly useful alignment proposals, yet the need for consideration of axial chirality revealed several flaws. Both programs were unable to treat stereogenic biaryl axes correctly without tedious manual refinement.

The best derived CoMSIA model showed a cross-validated q^2 value of 0.813 (for eight components) and a predictive r^2 of 0.578. This provides the possibility to achieve reliable predictions on the antimalarial activity of naphthylisoquinoline alkaloids, at least within the most frequently occurring 7,1′-, 7,8′-, 5,1′-, and 5,8′-coupling types. Predicted activities for 7,6′-coupled naphthylisoquinoline alkaloids still need to be treated carefully. Maybe, the future availability of more (and even more diverse) representatives of that type (and of the related 7,3′-coupling type) will permit more accurate predictions.

The most important structural conclusions of this study are the importance of the nitrogen atom and the finding that an oxygen function at C6 of the isoquinoline moiety clearly decreases the antimalarial activity. These results will surely be helpful for the design, selection, and synthesis of new, more active synthetic naphthylisoquinoline-related molecules and support the goal of creating less sophisticated, but also hopefully even more active compounds.

ACKNOWLEDGMENT

We wish to thank Dr. Reto Brun and Dr. Ronald Kaminsky (Swiss Tropical Institute, Basel) for carrying out the biological testing as well as Dr. Knut Baumann and Professor Anton J. Hopfinger for helpful discussions. This investigation received financial support from the UNDP/World Bank/WHO Special Program for Research and Training in Tropical Deseases (TDR, ID No. 991049) and from the Fonds der Chemischen Industrie.

REFERENCES AND NOTES

(1) Bringmann, G.; Pokorny, F. The Naphthylisoquinoline Alkaloids. In *The Alkaloids*; Cordell, G. A., Ed.; Academic Press: New York, 1995; Vol. 46, pp 127−271.

(2) Bringmann, G.; Rübenacker, M.; Vogt, P.; Busse, H.; Aké Assi, L.; Peters, K.; von Schnering, H. Dioncopeltine A and Dioncolactone A: Alkaloids from *Triphyophyllum peltatum*. *Phytochemistry* **1991**, *30*, 1691−1696.

(3) Bringmann, G.; Rübenacker, M.; Weirich, R.; Aké Assi, L. Dioncophylline C from the Roots of *Triphyophyllum peltatum*, the First 5,1′-Coupled Dioncophyllaceae Alkaloid. *Phytochemistry* **1992**, *31*, 4019−4024.

(4) Bringmann, G.; Wohlfarth, M.; Rischer, H.; Schlauer, J. A new biosynthetic pathway to alkaloids in plants: acetogenic isoquinolines. *Angew. Chem., Int. Ed.* **2000**, *39*, 1464−1466.

(5) Govindachari, T. R.; Nagarajan, K.; Parthasarathy, P. C.; Rajagopalan, T. G.; Desai, H. K. Absolute Stereochemistry of Ancistrocladine and Ancistrocladinine. *J. Chem. Soc., Perkin Trans. 1* **1974**, *12*, 1413−1417.

(6) Bringmann, G.; Rübenacker, M.; Jansen, J.; Geuder, T.; Aké Assi, L. Isolation, Structure Elucidation, and Stereoselective Total Synthesis of Novel Alkaloids from *Triphyophyllum peltatum*. *Planta Med.* **1990**, *56*, 495−496.

(7) Bringmann, G.; Hamm, A.; Günther, C.; Michel, M.; Brun, R.; Mudogo, V. Ancistroealaines A and B, Two New Bioactive Naphthylisoquinolines, and Related Naphthoic Acids from *Ancistrocladus ealaensis*. *J. Nat. Prod.* **2000**, *63*, 1465−1470.

(8) Bringmann, G.; Feineis, D. Novel Antiparasitic Biaryl Alkaloids from Westafrican Dioncophyllaceae Plants. *Act. Chim. Thérapeut.* **2000**, *26*, 151−171.

(9) Bringmann, G.; Rübenacker, M.; Ammermann, E.; Lorenz, G.; Aké Assi, L. Dioncophyllines A and B as fungicides; European Patent EP 0515 856 A1 1992.

(10) Bringmann, G.; Zagst, R.; Schäffer, M.; Hallock, Y. F.; Cardellina II, J. H.; Boyd, M. The Absolute Configuration of Michellamine B, a 'Dimeric', Anti-HIV−Active Naphthylisoquinoline Alkaloid. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1190−1191.

(11) François, G.; Bringmann, G.; Phillipson, J. D.; Aké Assi, L.; Dochez, C.; Rübenacker, M.; Schneider, C.; Wéry, M.; Warhurst, D. C.; Kirby, G. Activity of Extracts and Naphthylisoquinoline Alkaloids from *Triphyophyllum peltatum*, *Ancistrocladus abbreviatus* and *A. barteri* against *Plasmodium falciparum in vitro*. *Phytochemistry* **1994**, *35*, 1461−1464.

(12) François, G.; Bringmann, G.; Dochez, C.; Schneider, C.; Timperman, G.; Aké Assi, L. Activities of extracts and naphthylisoquinoline alkaloids from *Triphyophyllum peltatum*, *Ancistrocladus abbreviatus* and *Ancistrocladus barteri* against *Plasmodium berghei* (Anka strain) in vitro. *J. Ethnopharmacol.* **1995**, *46*, 115−120.

(13) François, G.; Timperman, G.; Eling, W.; Aké Assi, L.; Holenz, J.; Bringmann, G. Naphthylisoquinoline alkaloids against malaria: evaluation of the curative potential of dioncophylline C and dioncopeltine A against *Plasmodium berghei* in vivo. *Antimicrob. Agents Chemother.* **1997**, *41*, 2533−2539.

(14) The World Health Report 2001; World Health Organization: Geneva, 2001.

(15) Wellems, T.; Plowe, C. V. Chloroquine-resistant malaria. *J. Infect. Dis.* **2001**, *184*, 770−776.

(16) Bringmann, G.; Breuning, M.; Tasler, S. The Lactone Concept: An Efficient Pathway to Axially Chiral Natural Products and Useful Reagents. *Synthesis* **1999**, 525−558.

(17) Bringmann, G.; Menche, D. Stereoselective Total Synthesis of Axially Chiral Natural Products: via Biaryl Lactones. *Acc. Chem. Res.* **2001**, *34*, 615−624.

(18) Bringmann, G.; Saeb, W.; Koppler, D.; François, G. Jozimine A ('Dimeric' Dioncophylline A), A Non-Natural Michellamine Analogue with High Antimalarial Activity. *Tetrahedron* **1996**, *52*, 13409−13418.

(19) Bringmann, G.; Tasler, S.; Pfeifer, R.; Breuning, M. The directed synthesis of axially chiral ligands, reagents, catalysts, and natural products: through the lactone methodology. *J. Organomet. Chem.* **2002**, *661*, 49−65.

(20) Bringmann, G.; Messer, K.; Wolf, K.; Mühlbacher, J.; Grüne, M.; Brun, R.; Louis, A. M. Dioncophylline E from *Dioncophyllum tholonnii*, the first 7,3′-coupled Dioncophyllaceous naphthylisoquinoline alkaloid. *Phytochemistry* **2002**, *60*, 389−397.

(21) François, G.; Chimanuka, B.; Timperman, G.; Holenz, J.; Plaizier-Vercammen, J.; Aké Assi, L.; Bringmann, G. Differential sensitivity of erythrocytic stages of the rodent malaria parasite *Plasmodium chabaudi chabaudi* to dioncophylline B, a highly active naphthylisoquinoline alkaloid. *Parasitol. Res.* **1999**, *85*, 935−941.

(22) Chimanuka, B.; François, G.; Timperman, G.; Heyden, Y. V.; Holenz, J.; Plaizier-Vercammen, J.; Bringmann, G. A comparison of the stage-specific efficacy of chloroquine, artemether, and dioncophylline B, against the rodent malaria parasite *Plasmodium chabaudi chabaudi in vivo*. *Parasitol. Res.* **2001**, *87*, 795−803.

(23) François, G.; Steenackers, T.; Timperman, G.; Aké Assi, L.; Haller, R. D.; Bär, S.; Isahakia, M. A.; Robertson, S. A.; Zhao, C.; Souza, N. J. D.; Holenz, J.; Bringmann, G. Retarded Development of Exoerythrocytic Stages of the Rodent Malaria Parasite *Plasmodium berghei* in Human Hepatoma Cells by Extracts from Dioncophyllaceae and Ancistrocladaceae Species. *Int. J. Parasitol.* **1997**, *27*, 29−32.

(24) François, G.; Timperman, G.; Steenackers, T.; Aké Assi, L.; Holenz, J.; Bringmann, G. In vitro inhibition of liver forms of the rodent malaria parasite *Plasmodium berghei* by naphthylisoquinoline alkaloids − structure−activity relationships of dioncophyllines A and C, and ancistrocladine. *Parasitol. Res.* **1997**, *83*, 673−679.

(25) Schwöbel, B.; Bringmann, G.; Harwaldt, P.; Croft, S.; Becker, K. Unpublished results.

(26) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130−4146.

(27) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(28) Yu, Y.-W. Probing into the Mechanism of Action, Metabolism and Toxicity of Gossypol by Studying its (+)- and (−)-Stereoisomers. *J. Ethnopharmacol.* **1987**, *20*, 65−78.

(29) Müller, G. E. Thalidomide: from tragedy to new drug discovery. *Chemtech* **1997**, *27*, 21−5.

(30) Klebe, G.; Mietzner, T.; Weber, F. Different Approaches Toward an Automatic Structural Alignment of Drug Molecules: Applications to Sterol Mimics, Thrombin and Thermolysin Inhibitors. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 751−778.

(31) Klebe, G.; Mietzner, T.; Weber, F. Methodological Development and Strategies for a Fast Flexible Superposition of Drug-size Molecules. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 35−49.

(32) Klebe, G.; Mietzner, T. A Fast and Efficient Method to Generate Biologically Relevant Conformations. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 583−606.

(33) Sybyl 6.7.1 available from Tripos, Inc., 1699 St. Hanley Road, Suite 303, St. Louis, MO, 63144.

(34) Lemmen, C.; Lengauer, T. Time-efficient flexible superposition of medium-sized molecules. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 357−368.

(35) Lemmen, C.; Lengauer, T.; Klebe, G. FlexS: A method for fast flexible ligand superposition. *J. Med. Chem.* **1998**, *41*, 4502−4520.

(36) Rarey, M.; Wefing, M.; Lengauer, T. Placement of medium sized molecular fragments into active sites of proteins. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 41−54.

(37) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.

(38) Jones, G.; Willet, P.; Glen, R. C. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 532−549.

(39) Although FLEXS has recently been fully integrated, this study was performed with the earlier command line version.

(40) Bringmann, G.; Menche, D. First, Atropo-enantioselective Total Synthesis of the Axially Chiral Phenylanthraquinone Natural Products Knipholone and 6′-*O*-Methylknipholone. *Angew. Chem., Int. Ed.* **2001**, *40*, 1687−1690.

(41) Bringmann, G.; Vitt, D.; Schmitt, S. unpublished results.

(42) Configurational instability is assumed but has not been rigorously proven for compounds **7a**−**d**, while it has been clearly evidenced (e.g. by NMR) for compounds **5a**−**e**, **9a**, **9b**, and **10** and is obvious for **8a**−**d**.

(43) Bringmann, G.; Messer, K.; Brun, R.; Mudogo, V. Ancistrocongolines A−D, New Naphthylisoquinoline Alkaloids from *Ancistrocladus congolensis*. *J. Nat. Prod.* **2002**, in press.

(44) Bringmann, G.; Teltschik, F.; Michel, M.; Busemann, S.; Rückert, M.; Haller, R.; Bär, S.; Robertson, A.; Kaminsky, R. Ancistrobertsonines B, C, and D as well as 1,2-Didehydro-ancistrobertsonine D from *Ancistrocladus robertsoniorum*. *Phytochemistry* **1999**, *52*, 321−332.

(45) Bringmann, G.; Messer, K.; Schwöbel, B.; Brun, R.; Aké Assi, L. Habropetaline A, an antimalarial naphthylisoquinoline alkaloid from *Triphyophyllum peltatum*. *Phytochemistry* **1999**, in press.

(46) Bringmann, G.; Saeb, W.; God, R.; Schäffer, M.; François, G.; Peters, K.; Peters, E.-M.; Proksch, P.; Hostettmann, K.; Aké Assi, L. 5′-*O*-Demethyldioncophylline A, a New Antimalarial Alkaloid from *Triphyophyllum peltatum*. *Phytochemistry* **1998**, *49*, 1667−1673.

(47) François, G.; Timperman, G.; Holenz, J.; Aké Assi, L. L.; Geuder, T.; Maes, L.; Dubois, J.; Hanocq, M.; Bringmann, G. Naphthylisoquinoline alkaloids exhibit strong growth-inhibiting activities against *Plasmodium falciparum* and *P. berghei* in vitro − Structure−activity

relationships of dioncophylline C. *Ann. Trop. Med. Parasitol* **1996**, *90*, 115−123.
(48) Bringmann, G.; Günther, C.; Saeb, W.; Mies, J.; Brun, R.; Aké Assi, L. 8-*O*-Methyldioncophyllinol B and revised structures of other 7,6'-coupled naphthylisoquinoline alkaloids from *Triphyophyllum peltatum* (Dioncophyllaceae). *Phytochemistry* **2000**, *54*, 337−346.
(49) Bringmann, G.; Günther, C.; Saeb, W.; Mies, J.; Wickramasinghe, A.; Mudogo, V.; Brun, R. Ancistrolikokines A−C: New 5,8'-Coupled Naphthylisoquinoline Alkaloids from *Ancistrocladus likoko*. *J. Nat. Prod.* **2000**, *63*, 1333−1337.
(50) Hallock, Y. F.; Manfredi, K. P.; Blunt, J. W.; Cardellina II, J. H.; Schäffer, M.; Gulden, K.-P.; Bringmann, G.; Lee, A. Y.; Clardy, J.; François, G.; Boyd, M. R. Korupensamines A−D, Novel Antimalarial Alkaloids from *Ancistrocladus korupensis*. *J. Org. Chem.* **1994**, *59*, 6349−6355.
(51) Bringmann, G.; Menche, D.; Brun, R.; Msuta, T.; Abegaz, B. M. Bulbine-Knipholone, a New Axially Chiral Phenylanthraquinone from *Bulbine abyssinica* (Asphodelaceae): Isolation, Structural Elucidation, Synthesis, and Antiplasmodial Activity. *Eur. J. Org. Chem.* **2002**, *7*, 1107−1111.
(52) Bringmann, G.; François, G.; Aké Assi, L.; Schlauer, J. The Alkaloids of *Triphyophyllum peltatum* (Dioncophyllaceae). *Chimia* **1998**, *52*, 18−28.
(53) Ridley, R. G.; Hofheinz, W.; Matile, H.; Jacquet, C.; Dorn, A.; Masciadri, R.; Jolidon, S.; Richter, W. F.; Girometta, M. A.; Urwyler, H.; Huber, W.; Thaitong, S.; Peters, W. 4-Aminoquinoline analogues of chloroquine with shortened side chains retain activity against chloroquine-resistant *Plasmodium falciparum*. *Antimicrob. Agents Chemother.* **1996**, *40*, 1846−1854.
(54) Desjardins, R. E.; Canfield, C. J.; Haynes, D.; Chulay, J. Quantitative assessment of antimalarial activity in vitro by a semiautomated microdilution technique. *Antimicrob. Agents Chemother.* **1979**, *16*, 710−718.
(55) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219−3222.
(56) The use of AM1 charges did not improve the results, but led to lower q^2-values; we attribute this to the fact that initially force-field geometries were used, for which AM1 charges might not be appropriate.
(57) Sadowski, J.; Gasteiger, J. From atoms and bonds to three-dimensional atomic coordinates: automatic model builders. *Chem. Rev.* **1993**, *93*, 2567−2581.
(58) Sadowski, J.; Gasteiger, J.; Klebe, G. Comparison of Automatic Three-Dimensional Model Builders Using 639 X-ray Structures. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1000−1008.
(59) We refer to the spatial distribution of the two biaryl moieties, which, due to the CIP formalism, does not necessarily imply that they have the same axial descriptors (i.e. *M* or *P*).
(60) Jones, G.; Willet, P.; Glen, R. C. Molecular Recognition of Receptor Sites Using a Genetic Algorithm with Description of Desolvation. *J. Mol. Biol.* **1995**, *245*, 43−53.
(61) Cramer, R. D.; Bunce, J. D.; Patterson, D. E. Crossvalidation, bootstrapping and partial least squares compared with multiple regression in conventional QSAR studies. *QSAR* **1988**, *7*, 18−25.
(62) We have discussed these problems with Tripos, Inc., where an improved version of GASP is currently under development.
(63) Böhm, M.; Stürzebecher, J.; Klebe, G. Three-Dimensional Quantitative Structure−Activity Relationships Analysis Using Comparative Molecular Field Analysis and Comparative Molecular Similarity Indices Analysis To Elucidate Selectivity Differences of Inhibitors Binding to Trypsin, Thrombin, and Factor Xa. *J. Med. Chem.* **1999**, *42*, 458−477.
(64) Lindberg, W.; Persson, J.-A.; Wold, S. Partial Least-Squares Method for Spectrofluorometric Analysis of Mixtures of Humic Acid and Ligninsulfonate. *Anal. Chem.* **1983**, *55*, 643−648.
(65) Talele, T. T.; Kulkarni, S. S.; Kulkarni, V. M. Development of Pharmacophore Aligment Models as Input for Comparative Molecular Field Analysis of a Diverse Set of Azole Antifungal Agents. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 958−966.
(66) Bringmann, G.; Saeb, W. unpublished results.
(67) Bringmann, G.; Menche, D.; Bezabih, M.; Abegaz, B. M.; Kaminsky, R. Antiplasmodial Activity of Knipholone and Related Natural Phenylanthraquinones. *Planta. Med.* **1999**, *65*, 757−758.
(68) Norinder, U. Recent Progress in CoMFA Methodology and Related Techniques. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 25−39.
(69) Kim, K. H.; Greco, G.; Novellino, E. A Critical Review of Recent CoMFA Applications. *Perspect. Drug Discovery Des.* **1998**, *12*, 257−315.
(70) Bush, B. L.; Nachbar, R. B. Sample-distance Partial Least Squares: PLS optimized for many variables, with application to CoMFA. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 587−619.
(71) Our experience has shown that q^2_{SAMPLS} in most cases is very similar to the usual q^2_{LOO}.
(72) To simulate hydrogen bonds in CoMSIA, dummy atoms are added to potential H-bond donor groups (in the donor field) and to acceptors groups (in the acceptor field).
(73) Bringmann, G.; Messer, K.; Wohlfarth, M.; Kraus, J.; Dumbuya, K.; Rückert, M. HPLC-CD On-Line Coupling in Combination with HPLC NMR and HPLC-MS/MS for the Determination of the Full Absolute Stereostructure of New Metabolites in Plant Extracts. *Anal. Chem.* **1999**, *71*, 2678−2686.
(74) Bringmann, G.; Wohlfarth, M.; Rischer, H.; Heubes, M.; Saeb, W.; Diem, S.; Herderich, M.; Schlauer, J. A Photometric, 'Biomimetic' Screening Method for Dimeric Naphthylisoquinoline Alkaloids and Complete On-Line Stereochemical Assignment of a Dimer, by the Analytical Triad, LC-MS, LC NMR, LC-CD in Crude Plant Extracts. *Anal. Chem.* **2001**, *73*, 2571−2577.
(75) Bringmann, G.; Lang, G. Full Absolute Stereostructures of Natural Products Right from Crude Extracts: The HPLC-MS/MS-NMR-CD 'Triad'. In *Progr. Molec. Subcell. Biol.*; Müller, W. E. G., Ed.; Springer Verlag: Berlin, in press.
(76) Bringmann, G.; Gulden, K.-P.; Busemann, S. *Tetrahedron* **2003**, in press.

CI025570S