ELSEVIER

3D-Quantitative structure–activity relationships of synthetic antileishmanial ring-substituted ether phospholipids

Agnes Kapou,[a,b] Nikolas P. Benetis,[a] Nikos Avlonitis,[a] Theodora Calogeropoulou,[a] Maria Koufaki,[a] Efi Scoulica,[c] Sotiris S. Nikolaropoulos[b] and Thomas Mavromoustakos[a,*]

[a]*Institute of Organic and Pharmaceutical Chemistry, National Hellenic Research Foundation, 48 Vas. Constantinou Av., 116 35 Athens, Greece*
[b]*Laboratory of Pharmaceutical Chemistry, Department of Pharmacy, School of Health Sciences, University of Patras, 26500 Rion-Patra, Greece*
[c]*Department of Clinical Bacteriology, Parasitology, Zoonoses and Geographical Medicine, School of Medicine, University of Crete, 76309 Heraklion, Greece*

Abstract—The application of 2D-NMR spectroscopy and Molecular Modeling in determining the active conformation of flexible molecules in 3D-QSAR was demonstrated in the present study. In particular, a series of 33 flexible synthetic phospholipids, either 2-(4-alkylidene-cyclohexyloxy)ethyl- or ω-cycloalkylidene-substituted ether phospholipids were systematically evaluated for their in vitro antileishmanial activity against the promastigote forms of *Leishmania infantum* and *Leishmania donovani* by CoMFA and CoMSIA 3D-QSAR studies. Steric and hydrophobic properties of the phospholipids under study appear to govern their antileishmanial activity against both strains, while the electrostatic properties have no significant contribution. The acknowledgment of these important properties of the pharmacophore will aid in the rational design of new analogues with higher activity.
© 2006 Elsevier Ltd. All rights reserved.

1. Introduction

A relatively new antileishmanial drug, *miltefosine*, a phosphocholine analogue, has been proposed and approved as an orally active drug for the improvement of the classical therapies for the treatment of leishmaniasis containing *antimony*, *meglumine antimoniate*, and *sodium stibogluconate*. There are however serious problems with both the efficacy and the toxicity of the above-mentioned drugs. The lack of knowledge for the mechanisms of their action has become an additional obstacle in the improvement of these cures of leishmaniasis, while parallel attempts for vaccine development against the parasite have not reached the aim until today.[1]

These drawbacks call for a systematic evaluation of the known antileishmanial drugs and the development of improved variations with reduced side effects. To this end, the genome of *Leishmania major* has been recently sequenced by Ivens et al.[2] with a possible application in the revelation of the proteins of the parasite that are potential targets for drug treatments. This kind of data can be used in combination with the recent developments of QSAR, in particular the 3D methodologies CoMFA[3] (comparative molecular field analysis) and CoMSIA[4] (comparative molecular similarity index analysis) which are valuable tools for deriving information about the possible pharmacophores.

In the present work, 3D-QSAR (3D-quantitative structure–activity relationship) methods are employed for the first time on previously reported[5] antileishmanial and anticancer ring-substituted ether phospholipids, which belong to the same chemical category as miltefosine. Other rigid antileishmanial compounds such as chalcones have been studied earlier by QSAR.[6] The real challenge for the compounds treated in the present study was to deal with their increased flexibility since the applications of CoMFA and CoMSIA require the optimized 3D-conformations of all molecules. The problem of flexibility has been addressed before[7–9] but to our knowledge no 3D-QSAR study has confronted a set of

* Corresponding author. Tel./fax: +302107273869; e-mail: tmavro@eie.gr

molecules with flexible alkyl residues of different length, so successfully. The present report designates the importance of CoMFA and CoMSIA to develop a consistent model based on using a template derived from experimental NMR data and theoretical calculations.

The 3D-QSAR models derived by this study, describing quantitatively the spatial and electrostatic requirements of the pharmacophore, will be useful in guiding the synthesis of molecules with improved activity, that is, by incorporating the results in the SYBYL module Leap-Frog. LeapFrog is designed to generate new ligand structures starting either from a receptor-cavity structure, usually a biopolymer, or as in our case, a CoMFA model.[10]

Lipophilic alkylphospholipids influence membrane physiology of Leishmania in several ways, such as ether lipid metabolism, GPI biosynthesis, and signal transduction.[5] The difference in the sensitivity of the two strains to different ether phospholipids was related to the dissimilarity of the membrane sterols and the different lipid content of the two strains of Leishmania.[11] The effects of ether phospholipids on the thermotropic properties of model membranes were studied using DSC (Differential Scanning Calorimetry).[5] In spite of numerous attempts aiming to elucidate the mechanism of antileishmanial activity, the molecular target of this class of compounds is not known yet. However, this is not an obstacle for 3D-QSAR with CoMFA, in which the receptor structure is not necessary for a structure–activity model to be created. In fact, structural information about the unknown molecular target can be derived from CoMFA[3] and CoMSIA.[4]

CoMFA and CoMSIA analyses require that the active or at least the equilibrium 3D-conformations of the compounds involved are available. Although the compounds under study are extremely flexible, there was adequate information supporting the proposed conformation of the active representative compound **20** (1-{2-[(5-adamantylidenepentyloxy)hydroxy-phosphinyl-oxy]ethyl}-*N,N,N*-trimethylammonium inner salt) (Fig. 1 and Table 1), derived by the combined application of 2D-NMR experiments and Molecular Modeling.[5] Compound **20** has served as a template to

build the 3D-conformation of the other ether phospholipids of the present study. CoMFA and CoMSIA, additionally to their predictive power, have constituted a means of confirmation of the already proposed conformation of the adamantylidene-substituted ether phospholipid **20**.
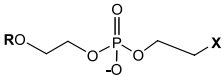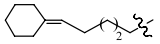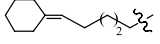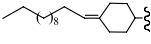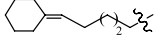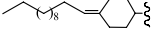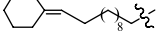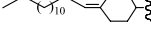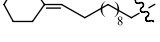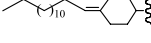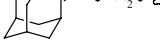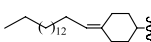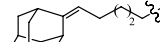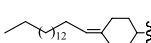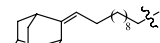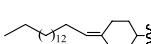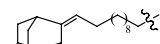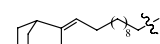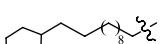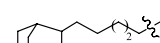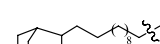
Avlonitis et al.[5] in their study focused on the correlation of the in vitro antileishmanial activity with the presence of cycloalkyl moieties in the lipid portion of alkyl-phosphocholines in combination with different polar head-groups. Several of the new synthetic compounds were found to be more active than the control compound, *miltefosine*, as seen in the following observations: (i) The 2-(4-alkylidene-cyclohexyloxy)ethyl analogues **4**–**12** and the 11-adamantylideneundecyl analogues **23**–**25** were active against *Leishmania donovani* and *Leishmania infantum*, and several of these were more potent than miltefosine. (ii) The 2-cyclohexyloxyethyl derivatives **1**–**3** and compounds **14**–**16**, bearing a short alkyl chain, were inactive, exhibiting IC_{50} values higher than 100 μM. However, compounds **20**, **22**, and **27**, though bearing a short pentyl alkyl chain, were very active against *L. donovani*. (iii) Generally, the trimethylammonium analogues were more potent than the respective *N*-methylpiperidino and *N*-methylmorpholino congeners. (iv) In general, the introduction of a 10–14-carbon unsaturated alkyl chain in position 4- of the 2-cyclohexyloxyethyl series and the presence of an unsaturated 5–11-carbon alkyl chain in ω-cycloalkylidene-substituted ether phospholipids lead to more active compounds against both Leishmania strains. (v) However, the replacement of the cyclohexylidene group in compounds **14** and **16** by an adamantylidene moiety (compounds **20**, **22**) rendered them active against *L. donovani*, as was the case for the saturated analogue of compounds **20**, **27**. (vi) The absence of a double bond in the cyclohexylidene-substituted compound **26** resulted in decreased activity. However, in the adamantylidene-substituted derivatives **27** and **28**, the absence or presence of the double bond did not seem to affect their potency.

In conclusion, a long alkyl chain seemed to be a prerequisite for the activity against *L. infantum*, while this was not the case for *L. donovani* against which compounds **20**, **22**, and **27** are very active. In most cases,
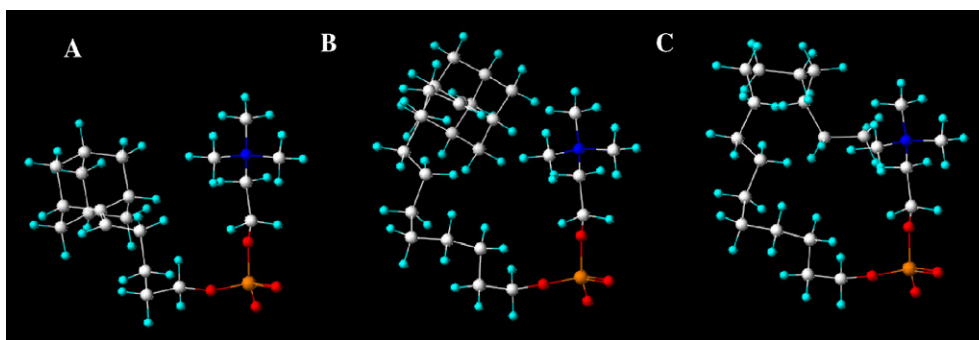
Figure 1. (A) The optimized 3D-conformation of compound **20** in solution as was concluded by a combination of NMR experiments and Molecular Modeling.[5] (B) The 3D-conformation of compound **23** resulted from that of **20**, by means of the field-fit procedure. Compound **23** served as the template for determining the conformation of the long-chain compounds. (C) The 3D-conformation of miltefosine (hexadecylphosphocholine), a phospholipid used originally as an antitumor drug and approved in Germany and India in 2003 as the first oral drug against leishmaniasis.[12]

Table 1. The chemical structures of the training set
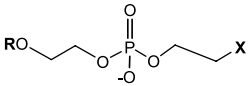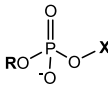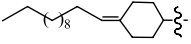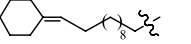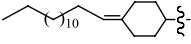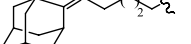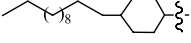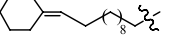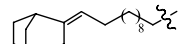
Compound	R	X	Compound	R	X
1		$-N^+(CH_3)_3$	13. Miltefosine		$-CH_2CH_2N^+(CH_3)_3$
2		H_3C^+N (piperidino)	14		$N^+(CH_3)_3$
3		H_3C^+N O (morpholino)	15		H_3C^+N (piperidino)
4		$-N^+(CH_3)_3$	16		H_3C^+N O (morpholino)
6		H_3C^+N O (morpholino)	17		$N^+(CH_3)_3$
7		$N^+(CH_3)_3$	19		H_3C^+N O (morpholino)
8		H_3C^+N (piperidino)	20		$N^+(CH_3)_3$
10		$-N^+(CH_3)$	21		H_3C^+N (piperidino)
11		H_3C^+N (piperidino)	23		$N^+(CH_3)_3$
12		H_3C^+N O (morpholino)	24		H_3C^+N (piperidino)
			25		H_3C^+N O (morpholino)
			26		$N^+(CH_3)_3$
			27		$N^+(CH_3)_3$
			28		$N^+(CH_3)_3$
			29		$N^+(CH_3)_3$
			32		$N^+(CH_3)_3$

The molecules in entries number **20** and **23** were used as templates in the alignment of the molecules. (The missing entries are shown in Table 2.)

the trimethylammonium substitution of the head-group was followed by increased potency compared to the *N*-methylpiperidino and *N*-methylmorpholino substitution.

The CoMFA and CoMSIA models, which are proposed here, verify the above conclusions and provide even more quantitative information which can be used by the SYBYL LeapFrog module in order to generate sug-

Table 2. The chemical structures of the test set

Compound	R	X	Compound	R	X
5		H_3C^+N (piperidine)	18		H_3C^+N (piperidine)
9		H_3C^+N–O (morpholine)	22		H_3C^+N–O (morpholine)
33		$-N^+(CH_3)_3$	30		$N^+(CH_3)_3$
			31		$N^+(CH_3)_3$

Table 3. Observed antileishmanial activity in combination with cytotoxicity and hemolysis data

Compound	Potency promastigotes IC$_{50}$ (μM)		Toxicity	
	L. infantum LEM235	*L. donovani LEM703*	Cytotoxicity IC$_{50}$ (μM)	Hemolysis HC$_{50}$ (μM)
1	>100	>100	—	—
2	>100	>100	—	—
3	>100	>100	—	—
4	3.25 ± 0.65	7.08 ± 1.2	25.954 ± 3.472	40.645 ± 6.651
5	23.07 ± 3.6	22 ± 3.25	—	—
6	16.46 ± 1.8	50.67 ± 3.6	—	>100
7	5.25 ± 0.45	3.91 ± 0.21	20.155 ± 3.193	>100
8	11.4 ± 2.4	29.7 ± 3.6	65.009	>100
9	15.5 ± 1.8	38.6 ± 3.2	20.389 ± 2.604	58.753 ± 2.23
10	21.19 ± 2.6	45.1 ± 7.2	—	>100
11	6.5 ± 1.7	21.96 ± 1.99	25.708 ± 1.261	>100
12	3.7 ± 0.71	16.22 ± 2.29	18.877 ± 3.512	>100
13	23.9 ± 4.21	8.72 ± 0.73	28.603 ± 2.478	38.257 ± 2.826
14	>100	>100	—	—
15	>100	>100	—	—
16	>100	>100	—	—
17	5.2 ± 1.5	2.4 ± 0.6	31.309 ± 2.774	68.31 ± 3.387
18	47.6 ± 7.33	8.7 ± 1	—	>100
19	22.8 ± 1.7	8.25 ± 0.25	—	—
20	>100	4.99 ± 1.5	>500	>100
21	>100	>100	—	—
22	>100	46.85 ± 8.7	—	—
23	6.75 ± 2.4	3.16 ± 0.63	59.503 ± 2.114	>100
24	22.58 ± 3.4	5.41 ± 1.14	168.7	>100
25	6.64 ± 1.2	5.09 ± 1.86	37.449 ± 0.689	>100
26	8.4 ± 0.8	10.3 ± 1.3	35.47 ± 2.283	>100
27	>100	4.02 ± 0.72	59.503 ± 2.114	>100
28	5.97 ± 1.06	2.88 ± 0.72	168.7	>100
29	>100	26.338	28.703 ± 3.217	59.816 ± 1.982
30	>100	8.562 ± 1.405	30.736 ± 3.165	76.28 ± 4.722
31	>100	16.613	39.713 ± 0.326	66.614 ± 5.293
32	38.045 ± 2.482	3.916 ± 0.757	74.953	>100
33	5.65 ± 1.93	9.49 ± 1.4	22.182 ± 2.475	75.733 ± 5.338
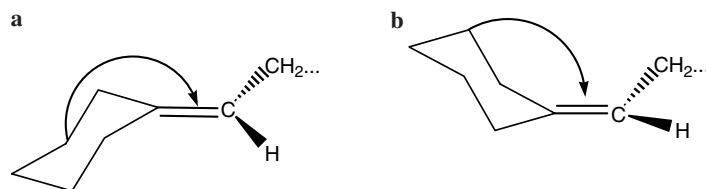
Figure 2. Axial dissymmetry. In (a) the cyclohexylidene ring lies below the plane, which is defined by the double bond and the atoms attached to it, while in (b) the cyclohexylidene ring lies above that plane.
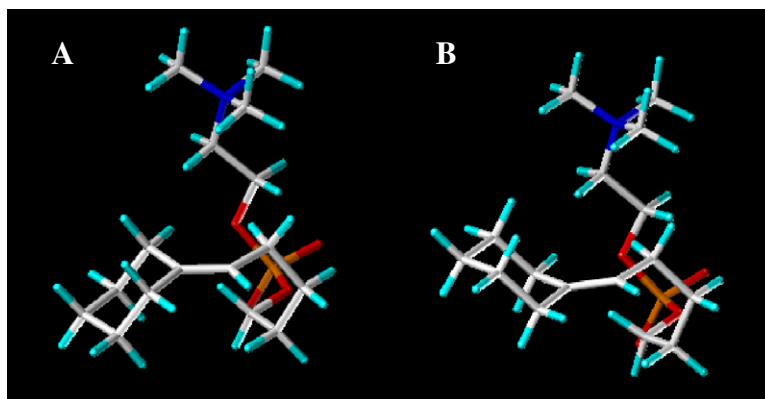
Figure 3. The two isomers of compound **14** are shown.

Table 4. Observed versus calculated activity values ($\log_{10}(1000/IC_{50})$) for Molecules in the training set

Compound	L. infantum				L. donovani			
	Observed activity		Predicted activity		Observed activity		Predicted activity	
	IC_{50} (μM)	$\log(1000/IC_{50})$	CoMFA	CoMSIA	IC_{50} (μM)	$\log(1000/IC_{50})$	CoMFA	CoMSIA
1	>100	1.00	0.99	1.03	>100	1.00	1.09	1.01
2	>100	1.00	0.95	0.94	>100	1.00	1.02	0.99
3	>100	1.00	0.99	0.95	>100	1.00	1.00	0.91
4	3.25 ± 0.65	2.49	2.24	2.20	7.08 ± 1.2	2.15	1.76	1.68
6	16.46 ± 1.8	1.78	2.04	2.08	50.67 ± 3.6	1.30	1.59	1.58
7	5.25 ± 0.45	2.28	2.17	2.16	3.91 ± 0.21	2.41	—	—
8	11.4 ± 2.4	1.94	1.91	1.93	29.7 ± 3.6	1.53	1.38	1.48
10	21.19 ± 2.6	1.67	—	—	45.1 ± 7.2	1.35	1.72	1.69
11	6.5 ± 1.7	2.19	2.20	2.28	21.96 ± 1.99	1.66	1.60	1.65
12	3.7 ± 0.71	2.43	2.28	2.33	16.22 ± 2.29	1.79	1.64	1.57
13	23.9 ± 4.21	1.62	1.82	1.74	8.72 ± 0.73	2.06	2.18	2.09
14	>100	1.00	1.06	1.06	>100	1.00	—	—
15	>100	1.00	0.88	0.95	>100	1.00	1.01	1.15
16	>100	1.00	0.92	0.95	>100	1.00	0.96	1.01
17	5.2 ± 1.5	2.28	—	—	2.4 ± 0.6	2.62	2.32	2.23
19	22.8 ± 1.7	1.64	1.68	1.57	8.25 ± 0.25	2.08	2.27	2.16
20	>100	1.00	1.10	1.09	4.99 ± 1.5	2.30	—	—
21	>100	1.00	0.92	0.94	>100	1.00	0.95	1.04
23	6.75 ± 2.4	2.17	2.19	2.09	3.16 ± 0.63	2.50	2.37	2.44
24	22.58 ± 3.4	1.65	1.91	2.06	5.41 ± 1.14	2.27	2.21	2.35
25	6.64 ± 1.2	2.18	2.10	1.99	5.09 ± 1.86	2.29	2.32	2.36
26	8.4 ± 0.8	2.08	1.99	1.93	10.3 ± 1.3	1.99	2.27	2.22
27	>100	1.00	1.11	1.10	4.02 ± 0.72	2.40	—	—
28	5.97 ± 1.06	2.22	2.24	2.23	2.88 ± 0.72	2.54	2.36	2.36
29	>100	1.00	0.95	0.96	26.338	1.62	—	—
32	38.045 ± 2.482	1.42	1.46	1.54	3.916 ± 0.757	2.42	2.51	2.57

gestions for the synthesis of new active compounds. The contour maps are used to create a 'negative' matrix in the place of the unknown active site and variations of the used ligands can be generated as long as they fit better into the 'imaginary' active site.

2. Results and discussion

A very good correlation was observed in CoMFA and CoMSIA for both *L. infantum* and *L. donovani* as it is demonstrated by the very high values of r^2_{conv}. Additionally, the credibility of the models is evidenced by the high values of r^2_{cv} and r^2_{bs} (Table 6). However, the real significance of the proposed models is verified by the good predictions of the activity of compounds belonging to the test set (Tables 5 and 6).

As it is shown in Figure 4, the CoMFA and the CoMSIA models for the *training set* fit rather well the experimental data of the antileishmanial activity. However, in the case of *L. infantum* the correlation is better, especially for the CoMFA model. The predictivity of all models is also demonstrated on the same figure in

which the compounds of the test set are shown in magenta squares.

Preliminary runs showed that compounds **10** and **17** for *L. infantum* and **7**, **14**, **20**, **27**, and **29** for *L. donovani* were outliers and were omitted during PLS runs. Compound **10** was an outlier because although it has a choline group it is not as active as expected against *L. infantum*, while compound **17** was not expected to be that active bearing a cyclohexylidene and not an adamantylidene group. In the case of *L. donovani*, compound **7** is much more active than any long-chained compound in this category. Compounds **14** and **20** differ only in the terminal ring-group of the lipophilic region. Compound **14** is inactive as a typical short-chained compound in its category, while **20** is one of the most active compounds against *L. donovani*. Compound **27** is also unexpectedly active because it is also a short-chained compound. This inconsistency 'confuses' the analysis, and subsequently compounds **20** and **27** were omitted from the PLS treatment. Finally, compound **29** is an outlier because it is much less active than any compound bearing a chain of 11 carbon atoms.

Table 5. Observed versus calculated activity values ($\log_{10}(1000/IC_{50})$) for molecules in the test set

Compound	*L. infantum*				*L. donovani*			
	Observed activity		Predicted activity		Observed activity		Predicted activity	
	IC_{50} (μM)	$\log(1000/IC_{50})$	CoMFA	CoMSIA	IC_{50} (μM)	$\log(1000/IC_{50})$	CoMFA	CoMSIA
5	23.07 ± 3.6	1.64	2.07	2.03	22 ± 3.25	1.66	1.59	1.61
9	15.5 ± 1.8	1.81	2.03	2.02	38.6 ± 3.2	1.41	1.42	1.46
33	5.65 ± 1.93	2.25	2.16	2.17	9.49 ± 1.4	2.02	2.05	1.86
18	47.6 ± 7.33	1.32	1.81	1.76	8.7 ± 1	2.06	2.23	2.22
22	>100	1.00	0.96	0.95	46.85 ± 8.7	1.33	0.94	0.97
30	>100	1.00	1.02	1.03	8.562 ± 1.405	2.07	2.26	2.17
31	>100	1.00	1.22	1.32	16.613	1.78	2.48	2.55

Table 6. Summary of results from CoMFA and CoMSIA

Summary of results from CoMFA and CoMSIA				
	L. infantum		*L. donovani*	
	CoMFA	CoMSIA	CoMFA	CoMSIA
r^2_{cv}	0.795	0.843	0.801	0.730
s_{press}	0.275	0.240	0.281	0.328
r^2_{conv}	0.953	0.942	0.899	0.879
SE	0.132	0.146	0.201	0.219
Components	3	3	2	2
F values	134.34	107.56	79.748	65.554
$Pr^2 = 0$	0.000	0.000	0.000	0.000
r^2_{pred}	0.902	0.909	0.815	0.746
r^2_{bs}	0.976	0.965	0.922	0.905
SD_{bs}	0.013	0.015	0.032	0.045
Fraction				
Steric	0.934	0.423	0.965	0.400
Electrostatic	0.066	0.060	0.035	0.090
Hydrophobic	—	0.517	—	0.509
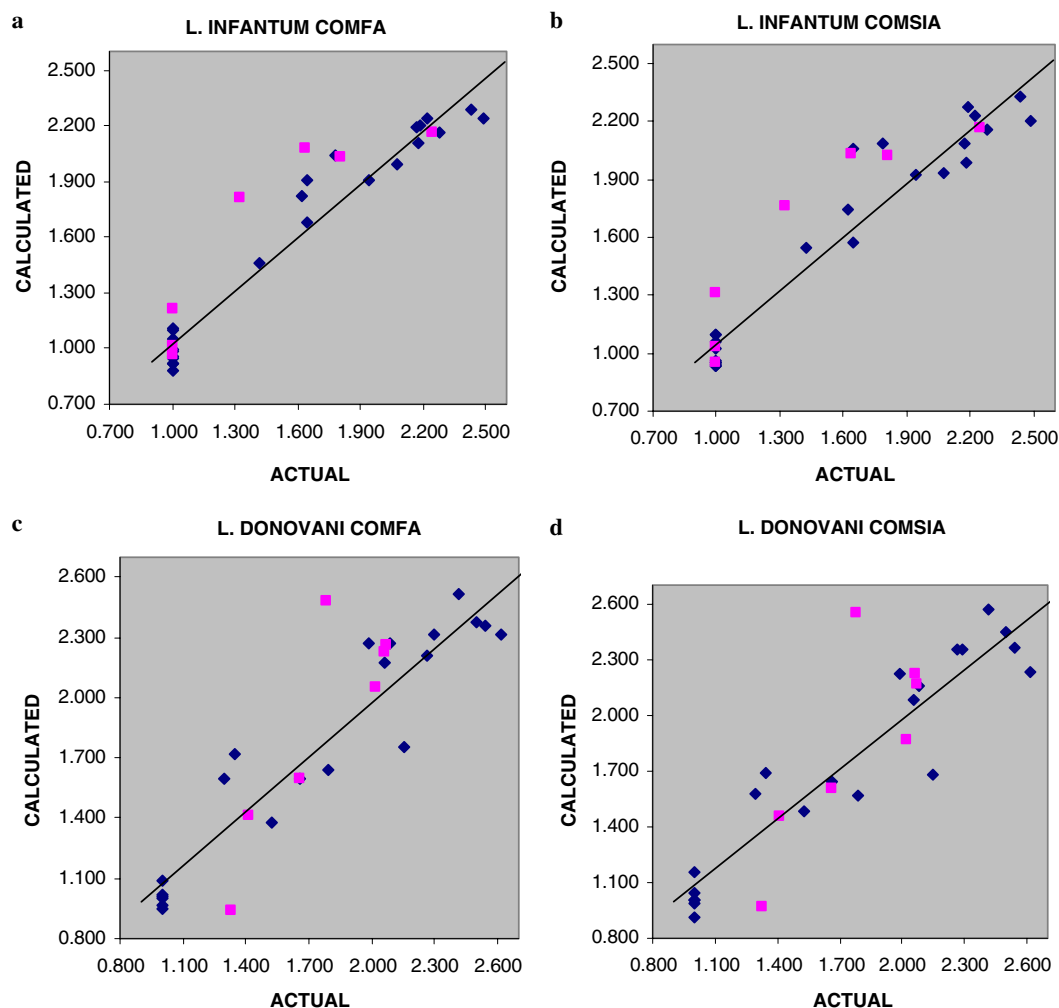
Figure 4. The plots of calculated versus observed activity based on the corresponding models: (a) *L. infantum*—CoMFA, (b) *L. infantum*—CoMSIA, (c) *L. donovani*—CoMFA, and (d) *L. donovani*—CoMSIA. Blue rhombs and magenta squares refer to the training and test set, respectively.

The following two general conclusions could be drawn from the characteristics of all 3D contour maps:

1. Sterics, as was expected (also shown from the relative fractions of the steric and electrostatic fields in Table 6), predominate in all models since the polar head-group remains basically unchanged in all compounds as far as the charged atoms are concerned. For this reason the polar part of the head-group of each compound was used for the superimposition on the template.
2. The overall preference for compounds comprising an extended alkyl chain, mainly in the case for *L. infantum*, was predicted by all models. It is evident that more compact and bulkier conformations, bringing the alkyl part of the molecule closer to the head-group, as is the case for such compounds, are generally preferred.

2.1. Graphics interpretation

For the interpretation of the CoMFA and CoMSIA sterics and electrostatic contour maps the following color conventions were made.

Sterics: Green contours indicate regions in which bulky groups are associated with increased antileishmanial

activity, while yellow contours indicate regions in which bulk is not favored.

Electrostatics: Red contours indicate regions in which the presence of negative charge increases activity, while blue contours indicate regions in which positive charge increases activity.

Additionally, for the interpretation of the CoMSIA hydrophobicity contour maps the following color conventions were made.

Cyan polyhedra indicate regions in which hydrophobicity is favored, while white polyhedra indicate regions in which the increase of hydrophobicity is correlated with reduced activity.

2.1.1. *Leishmania infantum*. From the CoMFA contour maps of *L. infantum* it is evident that the steric effect determines activity. Bulky groups, such as an adamanty-lidene, a cyclohexylidene group or an extended alkyl chain, in the lipophilic region of compounds comprising a long alkyl chain increase activity, while they are not favored in the corresponding part of compounds comprising a short alkyl chain. On the other hand, bulky

A. Kapou et al. / Bioorg. Med. Chem. 15 (2007) 1252–1265

1259

substituents are not favored in the region of the head-group, indicating a preference for the *N,N,N*-trimethy-lammonium moiety over piperidine and morpholine-substituted head-groups (Fig. 5).

Similar results were obtained with CoMSIA. All three *L. infantum* CoMSIA contour maps came out from the same PLS run, correlating the anti-*L. infantum* activity with the CoMSIA *steric*, *electrostatic*, and *hydrophobicity* parameters simultaneously (Fig. 6). However, we can visualize the requirements for each CoMSIA parameter separately (Figs. 7–9). According to the relative weights of the three CoMSIA parameters, as determined by the combined PLS run, steric effects and hydrophobicity appear to play the most important role in the determination of the QSAR, while electrostatics were not that important and concern mainly the presence of a double bond. A small red contour in the region corresponding to the double bond of long-chained compounds indi-

cates the preference for negative charge there. The position of the double bond in short-chained compounds, as is indicated by the small blue contour in that area, seems to render them inactive. The preference for bulky groups, the way it was described in CoMFA, seems to be correlated in CoMSIA, with increased hydrophobicity, as was expected (Fig. 5). Therefore, either parameter can be used to describe the activity of the compounds against *L. infantum*.

2.1.2. *Leishmania donovani*. The CoMFA contour map of *L. donovani* provides information similar to that of *L. infantum* (Fig. 10). Here too, the steric effect determines activity. An overall preference for bulky groups is demonstrated in the region of the alkyl chain of compounds comprising a long alkyl chain. However, the high activity of short-chained compounds **20** and **27** (omitted in the final analysis) is not compatible with the model and has to be further investigated. A prefer-
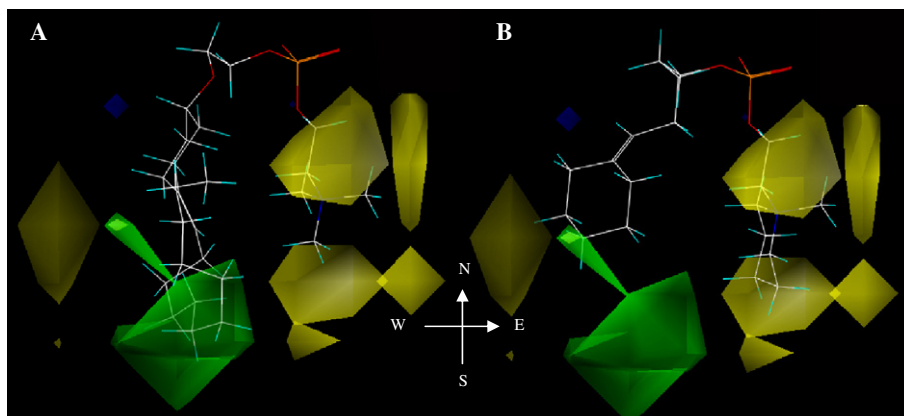
Figure 5. The *L. infantum* CoMFA STDEV*COEFF contour maps. (A) The very active compound **4**, which is displayed, occupies the green contour in the lipophilic region (S–W) with its 12C alkyl chain, while it leaves the S–E yellow contour in the area of the polar head-group unoccupied. These two structural features are according to the model associated with increased activity. (B) The inactive compound **15** combines three structural characteristics that are according to the model associated with reduced activity. The alkyl part is oriented toward the yellow contour in the lipophilic region (W), while the piperidine moiety occupies the yellow contours in the region of the head-group (N–E and S–E). Additionally, the double bond (electron excess) is oriented toward the small blue contour in the lipophilic region (N–W).
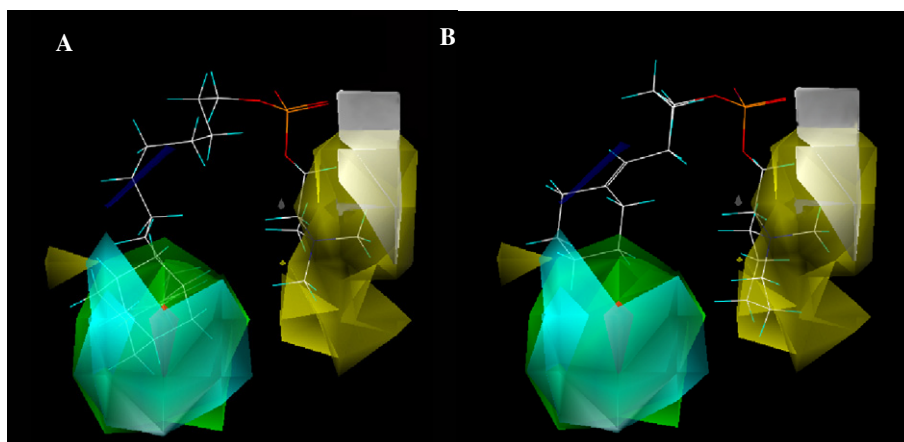
Figure 6. The *L. infantum* CoMSIA STDEV*COEFF contour maps (sterics, electrostatics, and hydrophobicity included). The significance of each color is explained in Figures 7–9. The preference for bulky groups in the region of the alkyl chain of long-chain compounds seems to be correlated with increased hydrophobicity. (A) The very active compound **23** is displayed. (B) The inactive compound **15** is displayed.
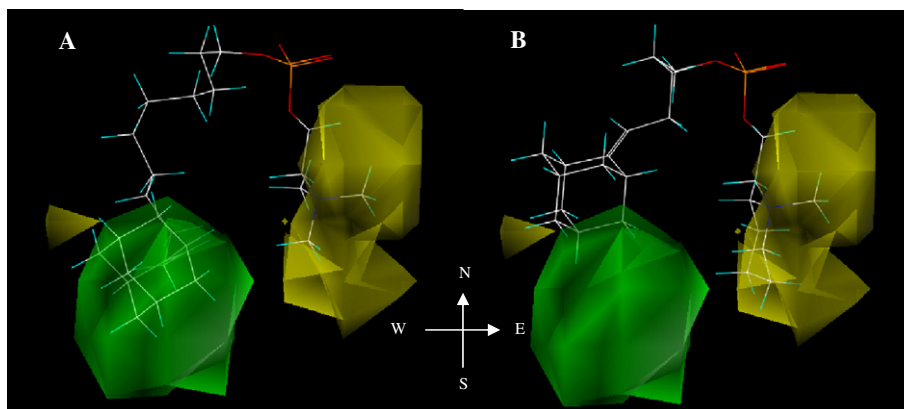
Figure 7. The *L. infantum* CoMSIA STDEV*COEFF steric contour map. (A) The very active compound **23**, which is displayed, occupies the green contour in the lipophilic region (S–W) with its adamantylidene-group, while it leaves the yellow contour in the area of the polar head-group unoccupied. These two structural features are according to the model associated with increased activity. (B) The inactive compound **21** combines three structural characteristics that are according to the model associated with reduced activity. The alkyl part is oriented toward the yellow contour in the lipophilic region (W), while the piperidine moiety occupies the yellow contour in the region of the head-group.
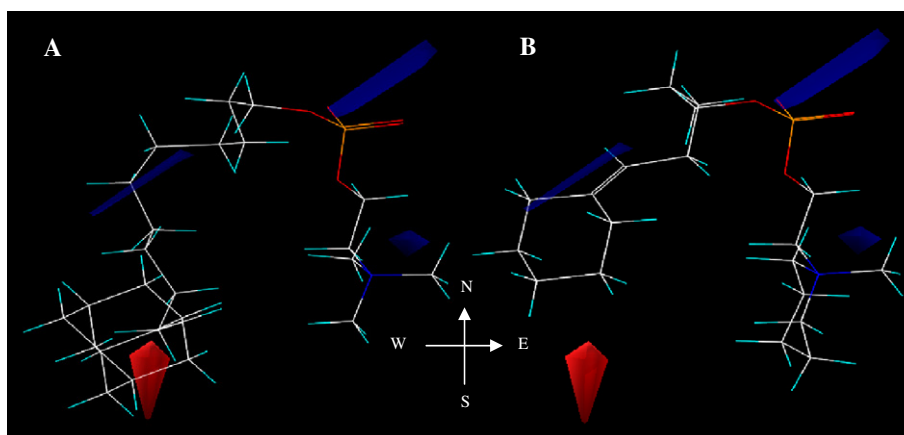
Figure 8. The *L. infantum* CoMSIA STDEV*COEFF electrostatics contour map. (A) The very active compound **23** is displayed. (B) The inactive compound **15** is displayed. Red polyhedra indicate regions in which negative charge is favored, while blue polyhedra indicate regions in which negative charge is disfavored. The double bond (electron excess) in compound **23** is oriented toward the red (S), while in compound **15** toward the blue contour (W). The blue polyhedra in the area of the head-group are related to noise due to a small location change of the O and N atoms of the phosphoric and ammonium groups in different molecules.

Figure 9. The *L. infantum* CoMSIA STDEV*COEFF hydrophobicity contour map. (A) The very active compound **7** is displayed. (B) The relatively inactive compound **21** is displayed. Cyan polyhedra indicate regions in which hydrophobicity is favored, while white polyhedra indicate regions in which the increase of hydrophobicity is correlated with reduced activity.

Figure 10. The *L. donovani* CoMFA STDEV*COEFF contour maps. Green contours indicate regions in which bulky groups increase activity against *L. donovani*, while yellow contours indicate regions in which bulk is not favored. Similarly, red contours indicate regions in which negative charge increases activity, while blue contours indicate regions in which negative charge is disfavored. (A) The very active compound **23** is displayed. (B) The inactive compound **2** is displayed. The red polyhedron (N) in the area of the head-group is related to noise due to small shifts of the phosphoric O, while the red polyhedron (S) accounts for a preference for the morpholine head-group over piperidine in some compounds.

Figure 11. The *L. donovani* CoMSIA STDEV*COEFF contour maps (sterics, electrostatics, and hydrophobicity included). Bulky groups in the region of the alkyl chain of long-chained compounds seem to be correlated with increased hydrophobicity. (A) The very active compound **23** is displayed. (B) The inactive compound **2** is displayed.
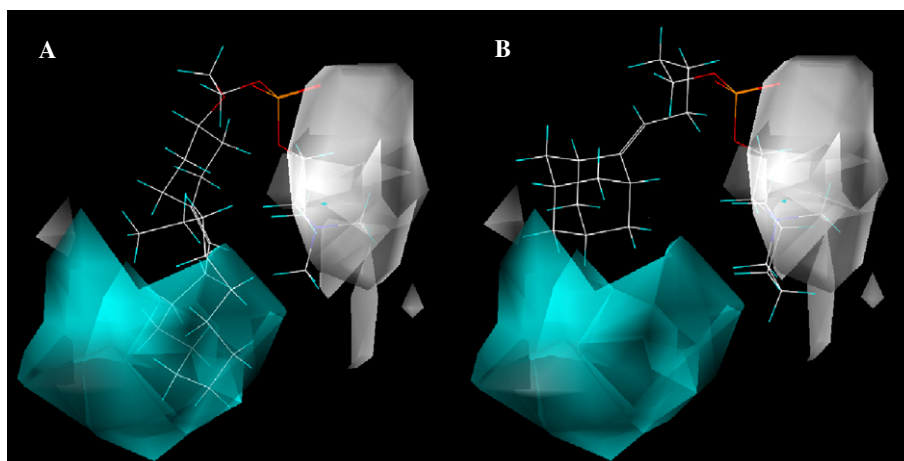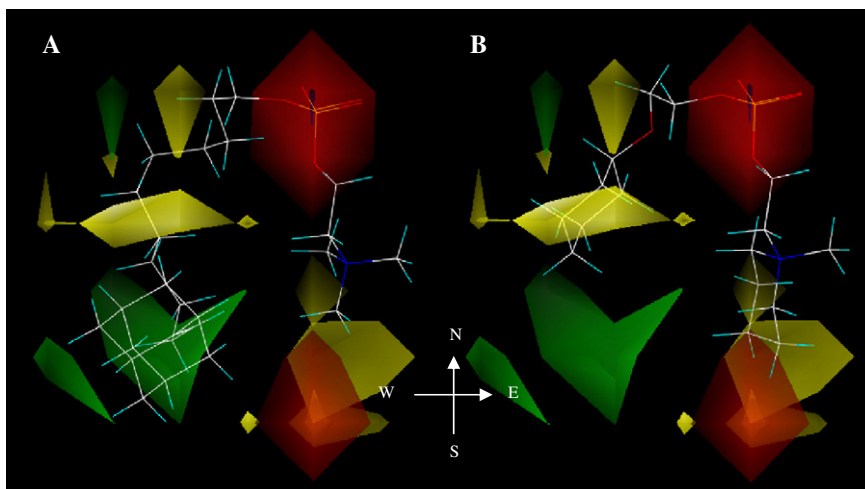

Figure 12. The *L. donovani* CoMSIA STDEV*COEFF steric contour map. (A) The very active compound **23** is displayed. (B) The inactive compound **2** is displayed. As with COMFA, green polyhedra indicate regions in which the accommodation of bulky groups is favored, while yellow contours indicate regions in which bulk is not favored.
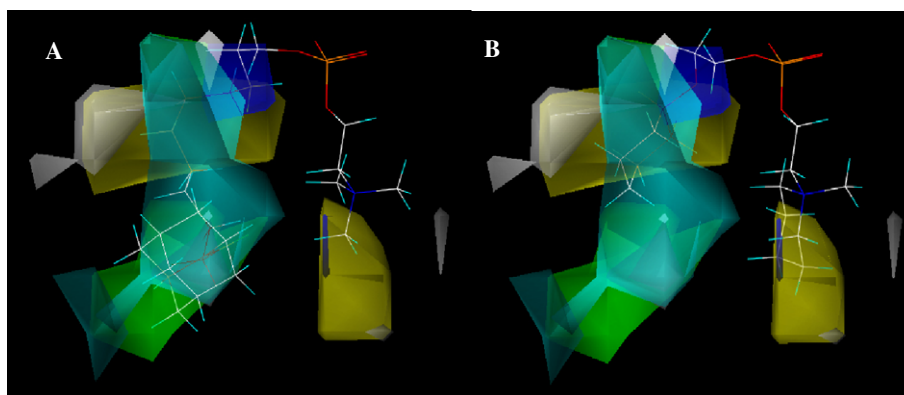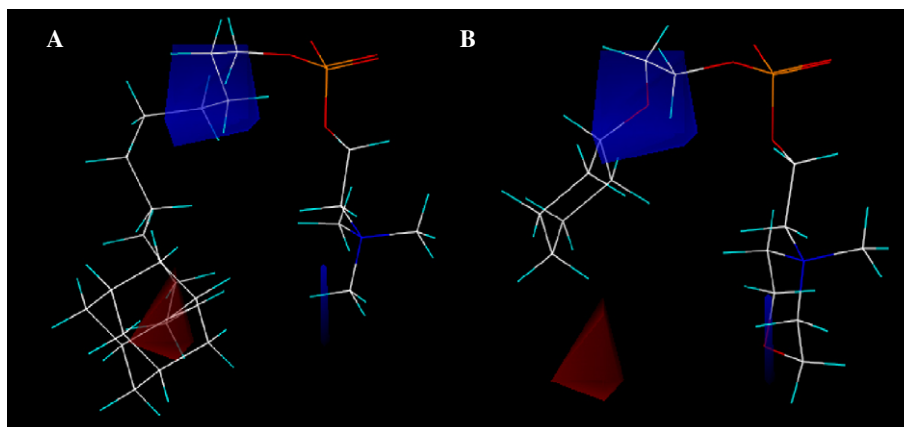
Figure 13. The *L. donovani* CoMSIA STDEV*COEFF electrostatics contour map. (A) The very active compound **23** is displayed. (B) The inactive compound **3** is displayed. Red polyhedra indicate regions in which negative charge is favored, while blue polyhedra indicate regions in which negative charge is disfavored.
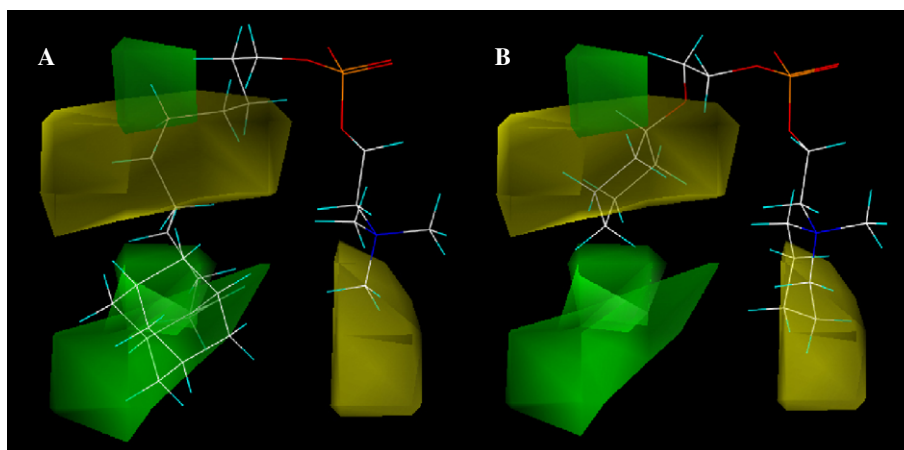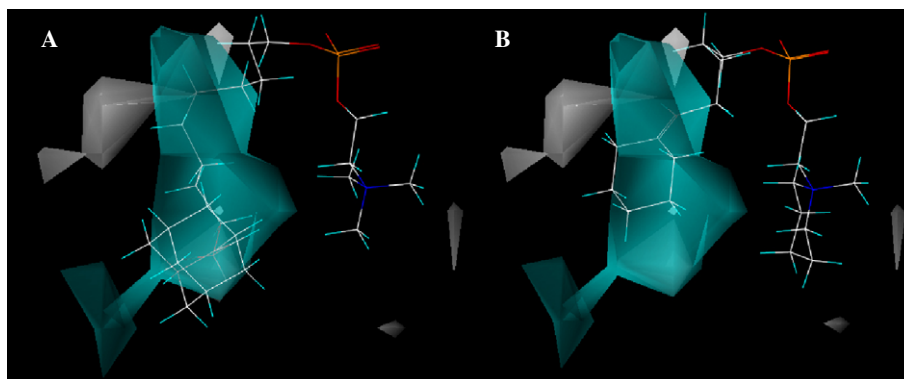

Figure 14. The *L. donovani* CoMSIA STDEV*COEFF hydrophobicity contour map. (A) The very active compound **23** is displayed. (B) The inactive compound **15** is displayed. Cyan polyhedra indicate regions in which hydrophobicity is favored, while white polyhedra indicate regions in which the increase of hydrophobicity is correlated with reduced activity.

ence for the *N,N,N*-trimethylammonium moiety over piperidine and morpholine-substituted head-groups is observed as with *L. infantum*.

As with *L. infantum*, all three *L. donovani* CoMSIA contour maps came out of the same PLS run, correlating the anti-*L. donovani* activity with the CoMSIA steric, electrostatic, and hydrophobicity parameters simultaneously (Fig. 11). Figures 12–14 provide a visualization of the requirements for each CoMSIA parameter separately. Steric effects and hydrophobicity appear to play the most important role in the determination of the QSAR, while electrostatics, although more highlighted than with *L. infantum*, concern mainly the presence and the position of the double bond in the alkyl chain. Also, as with *L. infantum*, the preference for bulky groups seems to be correlated with increased hydrophobicity, as was expected (Fig. 11).

Seeking for the practical developments which can lead to a new improved antileishmanial drug from the present ether phospholipid series an outline of the toxicity data in combination to the potency is following. The non-toxic compounds **8**, **24**, and **32** displayed medium to high potency against both promastigote forms. The com-

pounds **4**, **7**, **11**, **17**, **25**, and **26** constitute another group of rather active but relatively toxic compounds. Finally, the most interesting group comprises compounds **20**, **23**, and **28**, which have, in common, adamantylidene in the lipophilic part and choline in the head-group. They are all non-toxic and very active, except for compound **20** against *L. infantum*. Therefore, **20** and **23** were used as templates in the 3D-QSAR analysis. However, compounds **20** and **27**, though inactive against *L. infantum*, were very potent against *L. donovani* and proved to be non-toxic, as well. On the other hand, compounds **29**–**31**, though inactive against *L. infantum*, were potent against *L. donovani* but displayed medium toxicity. Compounds **18** and **19** were not tested for toxicity since no material is left. However, their observed potencies do not justify for repeating their synthesis. All inactive compounds with potency $IC_{50} > 100$ μM were not evaluated with respect to toxicity.

3. Conclusions

The present systematic procedure correlating the antileishmanial activity of the newly synthesized, ring-substituted ether phospholipids against the *L. infantum* and

L. donovani strains, with the steric, electrostatic, and hydrophobicity parameters resulted in the corresponding CoMFA and CoMSIA activity models. The good correlation observed in all cases supports the proposed conformation of template **20** and justifies our selection of alignment rule.

CoMFA and CoMSIA provided similar results; however the latter revealed an inner correlation of the steric and hydrophobic effect. It is evident from the contour maps of both types of analyses of the anti-*L. infantum* and anti-*L. donovani* data that the *steric effect determines activity*. Bulky groups in the terminal of the lipophilic alkyl chain of compounds comprising a long alkyl chain increase activity, while such groups are not favored in the lipophilic part of short-chain compounds. Thus, bulkier conformations, bringing the alkyl part of the molecule closer to the head-group, as is the case for compounds comprising a long alkyl chain, are generally preferred. In general, the presence of the adamantylidene group leads to greater activity against both strains compared to the cyclohexylidene group.

On the other hand, bulky substituents are not favored in the region of the head-group, indicating a preference for the *N*,*N*,*N*-trimethylammonium moiety over piperidine and morpholine-substituted head-groups.

However, in the case of *L. donovani*, the high activity of the short-chain compounds **20** and **27** (omitted in the final analysis) is not compatible with the model and has to be further investigated.

Information concerning the electrostatics is quite limited because the charged part of the polar head-group remains unchanged in all molecules. However, the electrostatic CoMFA and CoMSIA fields reveal the preference for an unsaturated (electron excess) over a saturated alkyl chain in the most active long-chain compounds.

The preference for bulky groups, as was described in CoMFA, seems to be correlated in CoMSIA, with increased hydrophobicity, as was expected. Therefore, either parameter can be used to describe the activity of the compounds against both Leishmania strains.

The activity data against the promastigote forms of Leishmania along with the toxicity data of the ether phospholipids lead to a conclusive result: high activity and minimum toxicity can be achieved by a combination of an adamantylidene group in the lipophilic part and choline in the head-group as it is depicted by compounds **20**, **23**, and **28**.

In conclusion, an extended alkyl chain seems to be a prerequisite for the activity against *L. infantum*, while this is not always the case for *L. donovani* against which certain compounds bearing a 5-membered alkyl chain are very active, probably due to the presence of the adamantylidene group. In general, the presence of the adamantylidene group leads to greater activity against both strains compared to the cyclohexylidene group. Additionally, and in most cases, the trimethylammonium substitution

of the head-group is followed by increased potency compared to the *N*-methylpiperidino and *N*-methylmorpholino substitution.

All this detailed information obtained by the CoMFA and CoMSIA models, combined with forthcoming results from the intracellular activity analyses, can be used as input of a special SYBYL module that generates new possible active compounds with enhanced activity and other tailored properties.

4. Materials and methods

4.1. Chemical structure-antileishmanial activity and toxicity

The chemical structures of the studied analogues, comprising *N*,*N*,*N*-trimethylammonium, *N*-methylpiperidino, *N*-methylmorpholino, *trans*-(2-hydroxy-cyclopentyl)-*N*,*N*,*N*-trimethyl-ammonium or *trans*(2-hydroxy-cyclohexyl)-*N*,*N*,*N*-trimethylammonium headgroups, are shown in Tables 1 and 2, which correspond to the *training* and *test sets*, respectively. Both sets comprise compounds that belong to two chemical categories. In the first category (left column of the tables) the hexadecyl chain in the control compound, miltefosine (hexadecylphosphocholine), is replaced by 2-cyclohexyloxyethyl or 2-(4-alkylidene-cyclohexyloxy)ethyl groups, while in the second one (right column of the tables) by cyclohexylidenealkyl or adamantylidenealkyl groups.

The antileishmanial activity (Table 3) of the compounds as described by Avlonitis et al.[5] was evaluated in vitro against the promastigote forms of *L. donovani* and *L. infantum* using an MTT (3-(4,5-dimethylthiazol-2yl)-2,5-diphenyltetrazolium bromide)-based enzymatic method as a marker of cell viability. The activity was expressed as the IC_{50} concentration. The experimental data were converted for QSAR purposes to the $\log_{10}(1000/IC_{50})$ form because logarithms tend to give more evenly distributed values of the target properties.[13]

The cytotoxicity of the ether phospholipids was assessed in the human monocytic cell line THP1 (Table 3). THP1 monocytes infected with the appropriate *Leishmania* species are used for the evaluation of the leishmanicidal activity of compounds against the intracellular amastigote stages of the parasite.

As a quantitative measurement of the cell damage after incubation with different concentrations of ether phospholipids dual staining with SYBR-14 and PI (Molecular Probes, The Netherlands) was used. THP1 cell cultures were incubated at 1×10^6 cells/ml with different concentrations of the compounds ranging from 50 to 156 μM. After an incubation period of 72 h, approximately 4×10^6 cells were suspended in labeling buffer (10 mM HEPES, 150 mM NaCl, and 10% BSA, pH 7.4), and 10 μg/ml PI and 0.1 mg/ml SYBR-14 were added. The cultures were incubated at 37 °C for 30 min before analysis by flow cytometry.

Cell samples were analyzed on an Epics Elite model flow cytometer (Coulter, Miami, FL). The green fluorescence of SYBR-14 and the red fluorescence of PI were excited at 488 nm. At least 10,000 cells were analyzed per sample and each staining experiment was repeated twice. Data analysis was performed on fluorescence intensities that excluded cell autofluorescence and cell debris.

Since hemolysis is a problem associated with alkyl-phosphocholines which prevents their use in injectable form, it is important to study the hemolytic properties of the new compounds. Thus, several of the new analogues were evaluated for their hemolytic activity against red blood cells. The results are expressed as the concentration that exhibits 50% hemolytic activity (HC_{50}) and are shown in Table 3.

EDTA-preserved peripheral blood from healthy volunteers was centrifuged in order to remove serum and red blood cells were washed thrice in PBS. After the final wash, cells were distributed in 96-well microplates (100 microliters/well) and an equal volume of each compound concentration was added. The test compounds were diluted in PBS in concentrations ranging from 100 to 6.25 μM. Incubation proceeded at 37 °C for 1 h, red cells were then centrifuged at $800g$ for 10 min. Absorbance of the supernatants was measured at 550 nm with reference filter at 625 nm. The percentage of hemolytic activity of each drug at different concentrations was estimated as $(A - A_0)/(A_{max} - A_0) \times 100$, where A_0 is the background hemolysis obtained by incubation with PBS and A_{max} is the 100% hemolysis achieved after incubation in dH_2O.

4.2. Molecular Modeling

The conformation of **20** in solution (Fig. 1A) was determined by Avlonitis et al.[5] by applying a combination of 2D-NOESY NMR experiments and Conformational Search studies.

The conformations of all compounds rise from that of the template by the application of *field-fit* and geometry optimization using SYBYL's Powell and Steepest Descents energy minimization algorithms. Gasteiger-Hückel charges were assigned to each compound. All compounds were treated with field-fit using compound **20** as the template for compounds comprising a short alkyl chain and compound **23** (Fig. 1B) for compounds comprising a long alkyl chain.

4.3. Selection of the template molecule and alignment rule

Before the application of CoMFA or CoMSIA on a set of compounds, a template molecule has to be selected and its active conformation has to be determined. Then, the whole set has to be superimposed on the template conformation with the application of an alignment rule. In contrast to planar aromatic systems with conjugated double bonds or rigid ring systems, such as steroids, the active configuration of more flexible molecules is not easy to predict. In the current study, in addition to the flexibility issue, the length of the alkyl chain varies sig-

nificantly among the compounds, making the selection of the appropriate template molecule even more difficult. The conformation of all compounds was dictated by the conformation of compound **20**, which was selected to be the template because its 3D-conformation had already been determined by NMR and systematic conformational search.[5]

Although compound **20** has much lower activity against *L. infantum* than other molecules in the series, its selection as a template is justified for two reasons. First, it is the least toxic in the series and second, compounds with longer alkyl chains, $(CH_2)_8$ and $(CH_2)_{10}$, possessing higher bioactivity, present NMR spectra with the alkyl region consisting of peaks severely overlapping. Thus, observed NOEs cannot lead decisively to a specific low-energy conformer, as it has already been shown previously[5] for compound **23**. Compound **20** was used as the template for molecules with a short alkyl chain. In order to achieve uniformity in the treatment of all compounds comprising a long chain, we created a second template, compound **23**, based on the conformation of **20**. Compound **23** was superimposed on **20**, the field-fit procedure was applied, and minimization followed. Finally, the resulting conformation of compound **23** was used as the template for compounds with long alkyl chain.

The alignment of the molecules was based on atom-by-atom superimposition of the six skeletal atoms –O–P–O–C–C–N in the head-group fragment –CH_2–CH_2–O–PO_2–O–CH_2–CH_2–N– of the molecules, which is common in all compounds. This treatment implies that only the relation of the non-overlapping lipophilic part of the test molecules to the antileishmanial activity can be investigated.

Dihedral constraints of the template were applied on the –CH_2–CH_2–O–PO_2–O–CH_2–CH_2–N– group of all compounds. Then, all compounds were superimposed on the corresponding template molecule and the ***field-fit minimization*** procedure was applied. Field-fit seems to be the method of choice for flexible molecules, the atom number of which varies significantly within the series. The rigid-body fit could not be applied since there was no atom-to-atom correspondence for every atom, between each compound and the template. In the field-fit procedure, the non-superimposed part, that is the hydrophobic chain of the fitted molecule, was forced to adopt the molecular field of the template, and consequently, a conformation analogous to the template's. The outcoming conformations were then subjected to minimization with the constraints activated.

As manifested in the ^1H NMR spectra, compounds **4**–**12** and **14**–**19** display a type of isomerism called *axial dissymmetry* that increases conformational complexity. The reason for the observed isomerism is the debility of the cyclohexylidene-group to rotate around the double bond[14,15] and it is demonstrated in Figures 2 and 3.

During synthesis a mixture of the two isomers was formed. Due to the small differences in the computed

energy ($\leqslant$1 kcal/mol) of the corresponding conformations, the **a** conformers were selected for compounds **4**–**12** and the **b** conformers for compounds **14**–**19** because the corresponding conformers were more resembling to the templates.

4.4. CoMFA and CoMSIA parameters

The CoMFA field was generated and calculated at the intersections of a regularly spaced grid of 2.0 Å, extending 5.0 Å in the *X*, *Y*, and *Z* directions, beyond the largest molecule in the series. For the calculation of the steric and electrostatic CoMFA fields (Lennard-Jones 6–12 and Coulomb potential functions, respectively), the sp^3, +1.0 charged carbon atom was used as the probe atom and the standard TRIPOS force field was applied. The dielectric function was selected to be distance dependent. For both fields, cut-off values were set to 30 kcal/mol. This indicates that, any steric or electrostatic field value that exceeds this value will be replaced by the value 30 kcal/mol, thus making a plateau of the fields close to the center of any atom. Respectively in CoMSIA,[4] the steric, electrostatic, and hydrophobic similarity indices were evaluated (Gaussian function) at the intersections of a similar grid using the same probe atom according to the standard implementation of CoMFA in SYBYL. The attenuation factor α was set to 0.3.

4.5. Partial least squares (PLS) analyses

The series was divided into a *training set* of 26 compounds (miltefosine included) and a *test set* of 7 compounds (Tables 1 and 2). The test set includes compounds representing all categories of activity of the training set, that is, inactive, active, and very active compounds comprising all structural features that are important for activity.

PLS analyses were carried out using the standard implementation of the PLS fitting procedure in SYBYL. Both the experimental data in $\log_{10}(1000/IC_{50})$ as well as the model predictions are shown in Table 4. In both CoMFA and CoMSIA a *crossvalidated* (leave-one-out) run preceded the final run as a measure of the statistical significance of the model and in order to obtain the *optimal number of components* to use in the final run. In all cases, column filtering was set to 2.00, thus, those columns (lattice points) the energy variance of which was less than 2.0 kcal/mol were omitted from the analysis, reducing noise. A shortcoming of applying a filter on the final 3D-model maps is that some surfaces of closed polyhedra disappear because they include points with field variance smaller than 2.0 kcal/mol. For the creation of the

CoMFA field, '***CoMFA standard***' scaling was selected, while in the case of CoMSIA '***autoscale***' was selected.

In order to obtain confidence limits for every conventional COMFA and CoMSIA PLS run, bootstrapping was also performed (100 runs, column filtering 2.00).

Prediction of the activity of the 7 compounds of the test set (Table 5) proved the predictive ability of each CoMFA and CoMSIA model.

Predictive r^2 (r^2_{pred}) was computed using the corresponding SYBYL routine. Besides the conventional parameters (r^2_{conv}, SE), the values of the statistical parameters concerning cross-validation (r^2_{cv}, s_{press}), bootstrap (r^2_{bs}, SD$_{\mathrm{bs}}$), and test-set predictions (r^2_{pred}) are also listed on Table 6. The definitions of the above statistical parameters are given in reference.[16]

References and notes

1. Laboratory for Organic Chemistry at the Swiss Federal Institute of Technology (ETH) in Zurich http://news.bbc.co.uk/1/hi/health/4930528.stm.
2. Ivens, A. C. et al. *Nature* **2005**, *309*, 436.
3. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
4. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
5. Avlonitis, N.; Lekka, E.; Detsi, A.; Koufaki, M.; Calogeropoulou, T.; Scoulika, E.; Siapi, E.; Kyrikou, I.; Mavromoustakos, T.; Tsotiris, A.; Golic-Grdadolnic, S.; Makriyannis, A. *J. Med. Chem.* **2003**, *46*, 755.
6. Liu, M.; Wilairat, P.; Croft, S. L.; Tan, A. L.-C.; Go, M.-L. *Bioorg. Med. Chem.* **2003**, *11*, 2729.
7. Nicklaus, M. C.; Milne, G. W. A.; Burke, T. R., Jr. *J. Comput. Aided Mol. Des.* **1992**, *6*, 487.
8. Demeter, D. A.; Weintraub, H. J. R.; Knittel, J. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1125.
9. Radwan, A. A.; Gouda, H.; Yamaotsou, N.; Torigoe, H.; Hirono, S. *Drug Des. Discov.* **2001**, *17*, 265.
10. Receptor-Based Design: LeapFrog, Introduction to LeapFrog, *The Tripos Bookshelf, version 6.8,* **2001**, 3.
11. Escobar, P.; Matu, S.; Marques, C.; Croft, S. I. B. *Acta Trop.* **2002**, *81*, 151.
12. More, B.; Bhatt, H.; Kukreja, V.; Ainapure, S. S. *J. Postgrad. Med.* **2003**, *49*, 101.
13. QSAR/CoMFA: QSAR Techniques, Expressing Target Property Data, *The Tripos Bookshelf, version 6.8,* **2001**, 142.
14. Gawronski, J. K.; Madhav- Reddy, S.; Walborsky, H. M. *J. Am. Chem. Soc.* **1987**, *109*, 6726.
15. Duhamel, L.; Ravard, A.; Plaquevent, J. C. *Tetrahedron: Asymmetry* **1990**, *1*(6), 347.
16. Kharkar, P. S.; Desai, B.; Gaveria, H.; Varu, B.; Liorija, R.; Naliapara, Y.; Shah, A.; Kulkarni, V. M. *J. Med. Chem.* **2002**, *45*, 4858.