Pharmacological Activity and Membrane Interactions of Antiarrhythmics: 4D-QSAR/QSPR Analysis

Christian D. P. Klein[1,2] and A. J. Hopfinger[1,3]

Purpose. This study was done to explore the relationships of both macroscopic and molecular level physicochemical properties to in-vivo antiarrhythmic activity and interactions with phospholipid membranes for a set of cationic-amphiphilic analogs.

Methods. The 4D-QSAR method, recently developed by Hopfinger and co-workers (1), was employed to establish 3D-QSAR/QSPR models. Molecular dynamics simulations provided the set of conformational ensembles which were analyzed using partial least squares regression in combination with the Genetic Function Approximation algorithm to construct QSAR and QSPR models.

Results. Significant QSAR models for in-vivo antiarrhythmic activity were constructed in which logP (the partition coefficient), and specific grid cell occupancy (spatial) descriptors are the main activity correlates. LogP is the most significant QSAR descriptor. 4D-QSPR models were also developed for two analog-membrane interaction properties, the change in a membrane transition temperature and the ability of the analogs to displace adsorbed Ca^{2+}-ions from phosphatidylserine monolayers.

Conclusions. Spatial features, represented by grid cell occupancy descriptors, supplement partition coefficient, which is the most important determinant of in-vivo antiarrhythmic activity, to provide a comprehensive model for drug action. The QSPR models are less significant in statistical measures, and limited to interpretation of possible molecular mechanisms of action.

INTRODUCTION

The voltage gated sodium channel is present in tissues that perform the propagation of electrical impulses. In humans, these tissues are neurons and myocardial cells. The voltage gated sodium channel plays a major role in the depolarization phase of an impulse (2).

A variety of substances block or inhibit the voltage gated sodium channel and, therefore, produce both local anesthesia and antiarrhythmic effects. From the pharmacological point of view, antiarrhythmic drugs are usually subdivided into four classes with several subgroups within some of the classes. Class I antiarrhythmic agents include the sodium channel blockers. In this respect, their mode of action cannot be distinguished from

local anesthetics. Local anesthetics, like lidocaine, are, in fact, widely used as antiarrhythmics (3).

Class I antiarrhythmics in clinical use cover a wide structural range from quinidine to relatively simple local anesthetics like lidocaine. This paper focuses on the latter. The common features of this sub-class of antiarrhythmics are:

i. A relatively large lipophilic moiety which often consists of an aromatic system and an alkyl chain. In most cases, the alkyl chain is attached to the aromatic part by an ester or amide linkage.

ii. A hydrophilic group at the end of the alkyl chain, often a tertiary nitrogen, which is protonated under physiological conditions. However, sodium channel blockers need not be charged, and some neutral compounds such as benzocaine or polidocanol are also local anesthetics (3).

The underlying pharmacological event occurring at the channel and causing its inactivation is not exactly known for class I antiarrhythmics. Two major hypotheses concerning the mode of action are:

i. The voltage gated sodium channel possesses a specific binding site for antiarrhythmics. This model is, for example, supported by the stereospecific effects observed with some antiarrhythmics (4). Moreover, the pharmacological effect of Class I antiarrhythmics is not the same for all compounds. Quinidine, for example, blocks sodium channels in the open state, while lidocaine interacts with both open and closed channels (2,5,6).

ii. The activity is due to a relatively non-specific uptake of the substances in myocardial or neuronal cell membranes—a hypothesis similar to the classical Meyer-Overton explanation of general anesthesia (7). Changes in the thermodynamic and dynamic properties of the membrane are believed to cause the inhibition of sodium channels. Strong support for this theory is the ability of compounds of high structural diversity to produce the same effect (2,6,8).

In order to elucidate the structural parameters responsible for the sodium-channel blockade, and other membrane-related effects, a set of substituted N,N-Dimethyl-3-phenylpropyl-amines was synthesized and subsequently tested in different biological and biophysical systems (9,10). Antiarrhythmic potency was assayed in Langendorff-preparations of guinea pig hearts. The drug-effect on excitability was measured as the drug-concentration leading to a 50% increase in threshold intensity of 50 Hz alternating current to induce ventricular arrhythmia (AC_{50} value) (11).

Biophysical methods were used to to characterize nonspecific interactions of the compounds with phospholipid membranes. The ability to replace $^{45}Ca^{2+}$ from phosphatidylserine monolayers was measured as the drug-concentration causing a 50% decrease in adsorbed $^{45}Ca^{2+}$ (IC_{50}) (12). It has to be pointed out, that the IC_{50} value does **not** represent the binding affinity of a compound to a receptor protein. Rather, this measure reflects a relatively non-specific, but nevertheless, concentration-dependent uptake of a compound into phospholipid (PL) membranes, and the subsequent displacement of Ca^{2+}-ions from their "binding sites" at the membrane/water interface (12).

The logarithm of the partition coefficient (logP) of each of the compounds was determined by HPLC using an RP18 column

---

and methanol/phosphate buffer (pH 7.4) mobile phases having different water contents. This system was calibrated using a series of alkylbenzene homologs with known (octanol/water) logP values. The logP values were obtained by extrapolation of the k' (capacity factor) values to 100% water (9).

The influence of membrane-interacting compounds on the thermodynamic properties of a PL membrane was monitored via differential scanning calorimetry of liposome preparations (13). In particular, the main phase transition temperature, T_t, was measured. For this purpose the PL dipalmitoylphosphatidic acid proved to be most suitable, as its undisturbed transition occurs at a relatively high temperature (337 K), and the position of the second signal, formed upon addition of a membrane-active compound, is independent of the molar ratio [drug/PL] (14,15). A hypothetical explanation for this phenomenon can be inferred by the coexistence of membrane domains consisting of a [drug/PL] mixture at a fixed molar ratio and domains containing pure PL.

The 4D-QSAR molecular modeling method was employed to quantify the relationship between structure and antiarrhythmic activity, as measured by AC_{50}. 4D-QSAR analysis, developed by Hopfinger and co-workers (1), in summary, consists of first generating a sample ensemble of the conformations available to every compound in a training set using molecular dynamics simulation, MDS. Each conformation is then aligned in a grid cell lattice, and the grid cell occupancies of user-defined atom-types are computed for each conformation of the ensemble. Multiple alignments can be efficiently explored. The resulting grid cell occupancy profile of a compound can be regarded as a representation of its three-dimensional dynamic behaviour. The grid cell occupancy profiles of the whole training set form an array of descriptors, similar to the one obtained in CoMFA (16), which is subjected to data reduction using a combination of partial least squares, PLS, (17) and a genetic algorithm, GA, tool called the Genetic Function Approximation (18). Consequently, 3D-QSAR equations are derived as a function of conformations, alignment and distribution of atom types. It is this multiplicity in conformation and alignment which represents the fourth dimension of the 4D-QSAR formalism. The utility of a combined GA/PLS approach in molecular modeling has recently been illustrated in publications by Rogers and Dunn (19) and Hasegawa et al. (20).

The 4D-QSAR method was also applied to establish quantitative structure property relationships, QSPRs, for the biophysical properties, i.e. phase transition temperature of DPPA liposomes (T_t) and replacement of $^{45}Ca^{2+}$ from phosphatidylserine monolayers (IC_{50}).

MATERIALS AND METHODS

The SAR Dataset

The chemical structures of all analogs used in the 4D-QSAR analysis, and the corresponding biological activities and biophysical properties are given in Table I. Some of the antiarrhythmic activities (AC_{50}) are missing, or could not be determined with the desired accuracy, and the activities for two chiral compounds (19 and 20) were measured using the racemate. Hence, several compounds from the parent data set had to be excluded from the 4D-QSAR analysis of antiarrhythmic activity. Likewise, the phase transition temperatures of DPPA

liposomes (T_t) and the IC_{50} values have not been measured for all compounds. It was therefore necessary to construct three training subsets from the parent data set.

Building the Molecules

All molecules were built using the building tool and library fragments of Chemlab-II (21). The side chains were assigned all-trans conformations. The side chain tertiary amino groups are protonated under physiological conditions (see pKa values, Table I). Accordingly, the compounds were built as monocations.

Semi-empirical Calculations

Calculations at the semi-empirical level were carried out to assign partial atomic charges and to further optimize molecular geometry. The AM1 Hamiltonian, implemented in the MOPAC 6.0 (22) program, was employed.

Molecular Dynamics Simulation (MDS)

The molecular dynamics simulation package MOLSIM 3.0 (23) was used to carry out the MDS on the AM1-optimized molecules. The simulation conditions were as follows:

Step size: 0.001 ps
Steps: 10 000 for equilibration, 100 000 for production run
Temperature: 300 K, constant
Velocity rescaling/temperature bath: used, relaxation time 1.0 ps
Molecular dielectric: 3.5
Cutoff distance for nonbonded interactions: 8 Å
Trajectory file: updated every 100 MD steps, resulting in 1000 frames
Conformational restrictions: none
Initial velocities: randomly assigned

Thermodynamic equilibrium is reached after about 5 ps in the MDSs of each of the analogs. Therefore, two MDS runs were performed for every analog: The first MDS lasting 10 ps to allow equilibration, with the final geometry and velocities written to restart files, and the second (100 ps, all conditions as described above) to record the production trajectories used in the 4D-QSAR analysis.
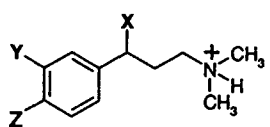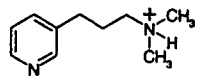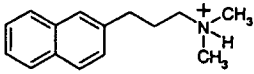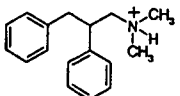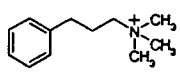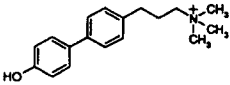
4D-QSAR ANALYSIS

Calculations of Grid Cell Occupancies

Grid cell occupancy profiles were calculated for the three alignments defined in Fig. 1. The chosen alignment atoms cover the entire common structural moiety of the test dataset. Since significant models were constructed for the endpoints AC_{50} and IC_{50} using these alignments, no other trial alignments were examined. For T_t, the transition temperature of DPPA liposomes, several additional alignments were explored without improvement in the resulting QSPR models.

A grid cell size of 1 Å was used. Cells having a variance in occupancy of less than 2% over the ensemble of all compounds were excluded from further analysis. Only absolute grid cell occupancy descriptors were considered in the analysis. Seven atom types of grid cell occupancies were computed and are listed as part of Fig. 1.

Table I. Chemical Structures, Biophysical Properties and Antiarrhythmic Activity of the Parent Dataset

| Compound #	X	Y	Z	log P[a]	pKa[b]	Mol. weight[c]	log AC$_{50}$[d] [μM]	T$_t$[e] [°K]	log IC$_{50}$[f] [μM]
1	H	H	Cl	2.21	10.59	198.72	4.31	302	3.85
2	H	H	Br	2.37	10.62	243.17	x	x	x
3	H	H	HO	0.35	x	180.27	3.29	314	2.72
4				0.57	10.68, 7.15	165.26	vl	x	3.60
5	H	H	Phenyl	2.99	9.88	240.37	4.80	285.3	5.40
6	H	H	H	1.50	10.96	164.27	3.55	308.7	3.00
7	H	H	Me	2.00	10.84	178.30	4.02	303.1	x
8	H	H	Et	2.45	10.63	192.32	3.94	300.4	x
9	H	H	n-Pr	2.89	10.54	206.35	4.85	295.5	x
10	H	H	i-Pr	2.75	10.49	206.35	4.22	299.4	x
11	H	H	t-Bu	2.96	10.03	220.38	4.52	297	5.40
12	H	H	MeO	1.48	10.91	194.30	3.62	x	x
13	H	H	4-Cl-phenyl	3.45	9.96	274.81	5.70	x	x
14	H	H	4-MeO-phenyl	2.95	9.85	270.39	4.96	296	5.40
15	H	H	4-EtO-phenyl	3.30	9.82	284.42	4.80	x	x
16				2.62	10.27	214.33	4.52	297	4.35
17	H	Cl	MeO	1.92	10.76	228.74	4.44	x	x
18	Phenyl	H	H	2.76	10.00	240.37	4.37	308	4.85
19	Phenyl	H	Phenyl	3.36	9.85	316.47	x	293.5	x
20				2.59	9.93	240.37	4.13	309	x
21	H	H	NO$_2$	1.50	10.74	209.27	vl	307	4.15
22				0.44	x	178.30	vl	307	3.40
23	H	H	4-HO-phenyl	2.10	9.32	256.37	4.43	348.8	4.40
24	H	Cl	HO	0.97	x	214.71	3.82	294	x
25				2.17	x	284.42	vl	x	x
26	H	H	NH$_2$	0.12	10.87, 7.34	179.29	3.00	x	x

[a] Measured by RP-HPLC (9)

[b] Measured by aqueous titration of hydrochlorides (9).

[c] Calculated, MOPAC 6.0 (22).

[d] Concentration to elevate the threshold of alternating current to induce arrhythmia in isolated guinea-pig hearts by 50% (9, 29).

[e] Main phase transition temperature of DPPA-liposomes/drug mixture (Tt of unperturbed DPPA: 337 K), measured by differential scanning calorimetry (DSC) (9, 10, 29).

[f] Drug concentration at which Ca^{2+}-binding to phosphatidylserine-monolayers is reduced to 50% of control value (9,29); x: not measured; vl: very low activity, could not be measured accurately.

The underlined values were included in the QSAR/QSPR analyses of the corresponding property endpoint.
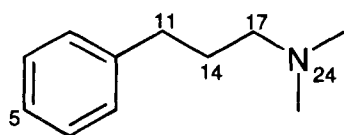
Alignment I: 24 - 17 - 14
Alignment II: 24 - 14 - 5
Alignment III: 14 - 11 - 5

Grid cell descriptor code number	Atom type
0	all
1	non-polar
2	polar / positive
3	polar / negative
4	H-bond acceptor
5	H-bond donator
6	aromatic

Fig. 1. Alignments, and grid cell descriptor codes.

The notation for reporting of grid cell data is as follows: Digits are separated by asterisks. The first three digits are the x, y, and z coordinates of the cell, the fourth digit defines the atom type, see Fig. 1. For example, the descriptor (0*0*0*6) specifies the occupancy of the grid cell at the origin (0*0*0) with atoms defined as "aromatic" (6).

Construction of QSAR Models/GFA-PLS

The conformation-independent descriptors logP and molecular weight (MW) and the dependent variables (-log AC_{50}, antiarrhythmic activity; T_t, phase transition temperature of DPPA liposomes; -log IC_{50}, replacement of $^{45}Ca^{2+}$ from phosphatidylserine monolayers) were separately added to the grid cell occupancy tables to complete each SAR training set. All columns were scaled to unit variance. The GFA program, Wolf 6.2, was used to derive QSAR models (18,24,25). The regression method was PLS with three components. A mutation probability of 100% was employed in all GFA runs, i.e. at every crossover a randomly generated, new basis function is added to the child. The starting population comprised 500 randomly generated models. Other GFA parameters (smoothing factor, number of crossover operations) were varied for all data subsets until optimal conditions were identified and are given in the **Results** section.

The ten correlation equations (QSAR models) with the lowest lack-of-fit score (*LOF*) for each trial alignment were crossvalidated (leave-one-out crossvalidation technique, implemented in Wolf 6.2; multiple linear regression is used to reconstruct models) in order to assess their predictive ability. The *LOF*, developed by Friedman (26), is a modified least square measure and the default ranking score of the Wolf program (18,19).

RESULTS

Background

The biophysical measurements were made in an attempt to develop indirect, but reliable, screening methods for antiarrhythmic activity that are independent of animal testing. Thus, the biophysical parameters, T_t and IC_{50}, could have been included as independent variables in the GFA analysis against the pharmacological endpoint. Unfortunately, the parent dataset is rather fragmented in regard to the biophysical measures IC_{50} and T_t. A GFA training subset, with all requisite descriptors, comprises no more than nine compounds of the total set. Since statistical significance is quite limited for models derived from such a small number of observations, we decided to construct three training subsets using the measured properties AC_{50}, T_t and IC_{50} as respective endpoints, and to develop independent GFA models for each of these training sets.

Correlation Between Conformation-independent Measures and Antiarrhythmic Activity

A crosscorrelation matrix for the three conformation-independent measures MW, logP and IC_{50} and the antiarrhythmic activity AC_{50} is given in Table II. Since the parent dataset consists of the common measures of these properties for only nine compounds, semiquantitative conclusions, at best, can be drawn from this correlation analysis. Nevertheless, the results indicate that logP, MW and IC_{50} are highly intercorrelated, while the phase transition temperature does not appear to have any significant linear correlation to any of the other biophysical properties.

Antiarrhythmic Activity: 4D-QSAR Analysis

All trial alignments yielded several GFA optimized models with $r^2 > 0.9$ and crossvalidated $r^2 > 0.8$ (see Table III a). The best models in terms of crossvalidated r^2, however, were obtained using alignment II. Therefore, models derived from alignment II were used in the further analysis.

Only linear basis functions were used in the GFA experiments for representation of the independent variables. Our experience using 4D-QSAR analysis suggests that a grid cell descriptor which is not significant in linear models shows only modest tendency to become significant when its non-linear dependence is considered. For the fitting of conformation-independent descriptors like logP, however, the use of non-linear basis functions has a physicochemical meaning. LogP often

Table II. The Crosscorrelation Matrix of the Conformation Independent Descriptors logP, MW and Tt and the Antiarrhythmic Activity -log AC_{50}, N = 9

	logP	MW	Tt	-log IC_{50}	-log AC_{50}
logP	1.00				
MW	**0.70**	1.00			
Tt	-0.43	0.07	1.00		
-log IC_{50}	**0.93**	**0.83**	-0.36	1.00	
-log AC_{50}	**0.93**	**0.85**	-0.32	**0.93**	1.00

Table IIIa. Summary of Statistical Measures of Fit for the Ten Best QSAR Models
for Antiarrhythmic Activity, Alignments I–III

model	I			II			III		
	LOF	r^2	xv-r^2	LOF	r^2	xv-r^2	LOF	r^2	xv-r^2
1	0.07	0.93	0.90	0.07	0.96	0.93	0.08	0.93	0.90
2	0.07	0.93	0.90	0.08	0.93	0.89	0.09	0.88	0.81
3	0.07	0.93	0.89	0.08	0.98	0.94	0.09	0.95	0.90
4	0.08	0.93	0.89	0.08	0.93	0.88	0.09	0.92	0.88
5	0.08	0.93	0.89	0.08	0.95	0.92	0.09	0.97	0.93
6	0.08	0.93	0.89	0.08	0.95	0.92	0.09	0.95	0.91
7	0.08	0.93	0.89	0.08	0.95	0.92	0.09	0.88	0.82
8	0.08	0.93	0.88	0.08	0.95	0.92	0.09	0.88	0.84
9	0.08	0.93	0.88	0.08	0.92	0.88	0.09	0.88	0.82
10	0.08	0.92	0.87	0.08	0.89	0.84	0.09	0.88	0.82

[N = 19]; LOF: Friedman's Lack-Of-Fit Score; xv-r^2: crossvalidated r^2; GFA options: 200000 Crossover operations, Smoothing factor = 1.5

shows a non-linear relationship to activity. In this study, however, significant QSAR models were generated using only a linear functional dependence for logP.

The ten best models for alignment II are given in Table III b. A linear crosscorrelation analysis of the grid-cell occupancy descriptors used by the ten best scored models was performed in order to detect pairs of cells (variables) which might be colinear.

Table IIIb. The Ten Best QSAR Models for Alignment II, Antiarrhythmic Activity

1 -log AC_{50} =
   2.61
   + 16.39 * (2*-2*11*0)
   + 0.65 * logP
   - 2.05 * (3*2*9*1)
   + 1.46 * (3*0*8*3)

2 -log AC_{50} =
   2.46
   + 0.70 * logP
   + 1.59 * (3*0*8*3)
   + 2.04 * (2*-2*9*6)

3 -log AC_{50} =
   2.67
   + 1.31 * (3*0*8*3)
   + 18.51 * (2*-2*11*0)
   - 2.84 * (4*1*9*0)
   + 0.65 * logP
   - 4.31 * (5*0*13*1)

4 -log AC_{50} =
   2.43
   + 7.51 *(1*-2*10*6)
   + 0.72 * logP
   + 1.62 * (3*0*8*3)

5 -log AC_{50} =
   2.43
   + 7.51 * (1*-2*10*6)
   + 0.72 * logP
   + 1.62 * (3*0*8*3)

6 -log AC_{50} =
   2.66
   + 0.64 * logP
   + 1.40 * (3*0*8*3)
   +15.30 * (2*-2*11*0)
   - 2.46 * (4*1*9*0)

7 -log AC_{50} =
   2.65
   + 0.97 * (2*0*9*3)
   + 15.59 * (2*-2*11*0)
   + 0.63 * logP
   - 1.82 * (3*2*9*1)

8 -log AC_{50} =
   2.73
   + 3.73 * (1*0*9*3)
   + 15.01 * (2*-2*11*0)
   + 0.61 * logP
   - 1.70 * (3*2*9*1)

9 -log AC_{50} =
   2.53
   + 0.68 * logP
   + 2.08 * (2*-2*9*6)
   + 1.04 * (2*0*9*3)

10 -log AC_{50} =
   3.09
   + 9.77 * (2*-2*11*0)
   + 0.46 * logP

The first three digits are x, y, z coordinates [Å] of the particular grid cell; fourth digit: atom type (see "methods" section).

The results are given in Table IV. As expected, the occupancies of neighboring cells tend to show a high degree of correlation. This can, for example, be seen for the two adjacent cells (2*-2*9*6) and (1*-2*9*6). No individual model of the ten best models, however, contains two highly correlated variables. The tendency to avoid redundant information can be partly attributed to the penalty for large models, which is introduced by the LOF score.

The residuals of error (difference between calculated and predicted activity) for the ten best models range from +0.4 to -0.4. A linear crosscorrelation matrix of the residuals of error, given in Table V, shows that several of these models provide similar information, since they have high crosscorrelation coefficients. That is, each model is not distinct/unique from all of the others. It is possible to identify two apparent sub-populations, or families, of models. Models 2, 4 and 9 form one of these families, models 1, 5, 6, 7, 8, and 10 the other. Model 3 cannot be clearly classified.

As the GFA model population evolves from its random starting point, more significant independent variables propagate. At the same time, the frequency of less significant variables decreases. Therefore, from an inspection of the variable usage in the final population, conclusions regarding the validity of particular variables can be drawn. The most often used dependent variables in the final population and the ten best models for alignment 1 are given in Table VI.

GFA 4D-QSAR models were also constructed only using grid cell occupancy descriptors. The best model from such a GFA experiment has an r^2 of 0.90 and a crossvalidated r^2 of 0.85. It is noteworthy that in the resulting GFA model population grid cell (2*-2*11*0), which is the second most significant descriptor in the GFA population based upon grid cell and conformation-independent descriptors, becomes the most important descriptor. This observation underlines the significance of this cell (spatial location) and the robustness of the GFA/PLS analysis.

Replacement of $^{45}Ca^{2+}$-ions from Phosphatidylserine Monolayers: 4D-QSPR Analysis

The ability to replace $^{45}Ca^{2+}$-ions from phosphatidylserine (PS) monolayers has been measured for twelve compounds.

Table IV. Crosscorrelation Matrix for the Grid Cell Occupancy Descriptors Found over the the Ten Best 4D-QSAR Models, logP and the Measured Activity (AC_{50}); Bold: Correlation Coefficient ≥ 0.6

	2*–2*11*0	4*1*9*0	5*0*13*1	3*2*9*1	1*–2*10*6	2*–2*9*6	1*0*9*3	2*0*9*3	logP	Meas. Activity
3*0*8*3										
2*–2*11*0	1.00									
4*1*9*0	**0.68**	1.00								
5*0*13*1	0.56	0.35	1.00							
3*2*9*1	**0.73**	**0.95**	0.43	1.00						
1*–2*10*6	**0.80**	0.44	0.27	0.44	1.00					
2*–2*9*6	**0.78**	0.40	0.21	0.39	**0.97**	1.00				
1*0*9*3	–0.47	–0.45	–0.22	–0.40	–0.34	–0.37	1.00			
2*0*9*3	–0.50	–0.48	–0.23	–0.43	–0.37	–0.40	**0.92**	1.00		
logP	**0.62**	**0.60**	0.34	0.58	0.35	0.40	**–0.67**	**–0.75**	1.00	
Meas. Activity	**0.78**	0.53	0.33	0.54	0.57	**0.61**	–0.48	–0.57	**0.90**	1.00

Although the resulting training set is small, the 4D-QSPR analysis was successful in deriving several significant correlation models. The most important statistical scores for alignments I to III are given in Table VII a.

Alignment I gave the best results in the GFA, as measured by crossvalidated r^2. Variable usage over the ten best models for alignment I and the whole population is summarized in Table VII b (individual models are not shown).

Phase Transition Temperature of DPPA Liposomes: 4D-QSPR Analysis

Table VIII a shows the statistical measures of fit obtained for the three trial alignments. Although it was possible to obtain excellent correlations for all alignments, the predictive potential of each of these models, as indicated by the crossvalidated r^2, is relatively poor. Alignment I yielded better crossvalidated r^2 scores than the other alignments, although its other statistical measures are of a lower quality than for alignments 2 and 3. The variable usage over the ten best models for alignment I, and over the corresponding GFA model population, is given in Table VIII b. Noteworthy is the absence of the two whole-molecule descriptors, logP and MW, among the ten best models (individual models are not shown).

DISCUSSION

Antiarrhythmic Activity

LogP is identified as by far the most significant descriptor. LogP as a single independent variable in linear regression analysis yields the following correlation equation.

$$-\log AC_{50} = 2.93 + 0.61 * \log P$$
$$N = 19 \quad r^2 = 0.81 \quad xv - r^2 = 0.76 \tag{1}$$

In the best 4D-QSAR model (model 3 of Table III b), logP accounts for about 90% of the total variance in activity. The conformation-independent descriptor MW, on the other hand, is only of minor significance as a descriptor.

The importance of lipophilicity on antiarrhythmic activity is not surprising, and has been pointed out previously (e.g., by Rauls & Baker (27)). The "screening" model is an isolated organ. Pharmacokinetic effects play a major role in this model, and lipophilicity, as a determinant for the penetration "potency" of a compound, can be expected to have a large influence on the expression of total activity. Cationic-amphiphilic drugs have long been known to accumulate in intact heart tissue (28).

Table V. Crosscorrelation Matrix for the Residuals of Error in Predicting $-\log AC_{50}$ of the Ten Best Models. Alignment II; Bold: Correlation Coefficient ≥ 0.6

Model	1	2	3	4	5	6	7	8	9	10
1	1.00									
2	0.45	1.00								
3	0.52	0.23	1.00							
4	0.44	**0.95**	0.26	1.00						
5	**0.89**	0.39	**0.61**	0.38	1.00					
6	**0.89**	0.38	**0.62**	0.37	**1.00**	1.00				
7	**0.93**	0.47	0.51	0.56	**0.85**	**0.84**	1.00			
8	**0.87**	0.48	0.46	0.51	**0.79**	**0.79**	**0.88**	1.00		
9	0.44	**0.95**	0.26	**1.00**	0.38	0.37	0.56	0.51	1.00	
10	0.57	0.59	0.38	**0.60**	**0.60**	**0.60**	**0.63**	**0.64**	**0.60**	1.00

Table VI. Descriptor Usage Over the 10 Best Models, and Over the Entire GFA Population, Alignment II, AC_{50}

Variable	% use in 10 best models	% use in whole population
logP	100.0	99.0
(2*–2*11*0)	70.0	26.6
(3*0*8*3)	60.0	12.0
(3*2*9*1)	30.0	2.2
(2*–2*9*6)	20.0	3.4
(4*1*9*0)	20.0	1.6
all others	10.0	

The first three digits are x, y, z coordinates [Å] of the particular grid cell; fourth digit; atom type (see "methods" section)

However, with the inclusion of descriptors dependent on molecular shape, such as the grid cell occupancy descriptors, more reliable QSAR models are obtained. The best model, given below, consists of logP and four grid cell occupancy descriptors. The statistical measures are better for this model than for the single logP model. One of the grid cell descriptors, (2*–2*11*0), is found in about 26% of all GFA QSAR models and is, therefore, to be regarded as very significant.

$$-\log AC_{50} = 2.67 + 1.31 * (3 * 0 * 8 * 3) +$$
$$18.51 * (2 *- 2 * 11 * 0) - 2.84 * (4 * 1 * 9* 0) +$$
$$0.65\ 8\ logP - 4.31 * (5* 0 * 13 * 1)$$
$$N = 19 \quad r^2 = 0.98 \quad xv - r^2 = 0.94 \tag{2}$$

Predictions were made for compounds 4, 21 and 22 on the basis of eqs. (2) and (3) and eq. (1). Equation (2) was selected because it is the the best model in terms of crossvalidated r^2. Equation (3) is particularly suited to complement eq. (2), since the residuals of error for both models are nearly uncorrelated (correlation coefficient: 0.23, see also Table V).

$$-\log AC_{50} = 2.46 + 0.70 * logP + 1.59 * (3 * 0 * 8 * 3)$$
$$+ 2.04 * (2 * -2 * 9 * 6)$$
$$N = 19 \quad r^2 = 0.93 \quad xv - r^2 = 0.89 \tag{3}$$

Due to their low activities, no explicit AC_{50} values were available for these compounds. The predicted activity values, ranging from 2.8 to 4.0, correspond to low-activity consistent with experiment.

Predictions of AC_{50} for the two stereoisomers of compound 20 were also made. The two 4D-QSAR models [eqs. (2) and (3)] each gave a predicted activity of 4.3 for both stereoisomers, while the logP-model predicted 4.5. The observed value (for the racemate) is 4.1.

Table VIIa. Summary of Statistical Measures of Fit for the Ten Best QSAR Models for $^{45}Ca^{2+}$-Replacement from PS Monolayers (IC_{50}), Alignments I–III

Model	I			II			III		
	LOF	r^2	xv-r^2	LOF	r^2	xv-r^2	LOF	r^2	xv-r^2
1	0.18	0.95	0.92	0.12	0.97	0.93	0.10	0.99	0.62
2	0.18	0.95	0.91	0.13	0.97	0.90	0.13	0.99	0.58
3	0.18	0.95	0.91	0.14	0.96	0.85	0.13	0.99	0.55
4	0.20	0.98	0.95	0.15	0.96	0.91	0.14	0.99	0.52
5	0.20	0.95	0.91	0.15	0.99	0.87	0.15	0.99	0.51
6	0.22	0.95	0.82	0.15	0.96	0.85	0.15	0.99	0.51
7	0.22	0.95	0.79	0.15	0.96	0.83	0.15	0.99	0.48
8	0.22	0.95	0.88	0.15	0.96	0.82	0.15	0.99	0.50
9	0.22	0.95	0.88	0.15	0.96	0.81	0.16	0.99	0.48
10	0.22	0.94	0.90	0.16	0.96	0.82	0.16	0.99	0.50

[N = 12]; LOF: Friedman's Lack-Of-Fit Score; xv-r^2: crossvalidated r^2; GFA options: 100000 Crossover operations, Smoothing factor = 1.2

Table VIIb. Descriptor Usage Over the 10 Best Models, and Over the Entire GFA Population, Alignment I, IC_{50}

Variable	% use in 10 best models	% use in whole population
2*0*4*1	80	22
logP	60	34
MW	20	47
4*–2*6*0	20	7
8*1*6*0	20	2
all others	10	

The first three digits are x, y, z coordinates [Å] of the particular grid cell; fourth digit: atom type (see "methods" section)

Table VIIIa. Summary of Statistical Measures of Fit for the Ten Best QSAR Models
for the Phase Transition Temperature (T_t) of DPPA Liposomes, Alignments I–III

	I			II			III		
Model	LOF	r^2	xv-r^2	LOF	r^2	xv-r^2	LOF	r^2	xv-r^2
1	37.1	0.92	0.63	8.23	0.99	0.07	7.62	0.99	0.05
2	37.2	0.92	0.62	8.98	0.99	0.15	7.76	0.99	0.13
3	37.9	0.92	0.62	12.6	0.99	0.11	8.38	0.99	0.17
4	37.9	0.92	0.62	13.1	0.99	0.06	8.53	0.99	0.18
5	38.1	0.92	0.61	13.4	0.98	0.13	9.94	0.99	0.14
6	38.1	0.92	0.62	17.8	0.99	0.08	10.0	0.99	0.11
7	38.3	0.92	0.62	18.0	0.99	0.03	10.3	0.99	0.13
8	38.9	0.92	0.61	20.1	0.97	0.11	10.5	0.99	0.02
9	38.9	0.92	0.61	20.7	0.98	−0.04	10.9	0.99	0.10
10	39.3	0.92	0.61	21.1	0.98	−0.01	12.1	0.99	0.15

[N = 16]; LOF; Friedman's Lack-Of-Fit Score; xv-r^2: crossvalidated r^2; GFA options: 50000 Crossover operations, Smoothing factor = 1

Graphical representations for ten conformations of compounds 13 (highly active) and 8 (inactive) are given in Fig. 2. All conformations shown are snapshots from the 100 ps MDS production run and lie within ± 2 kcal/mol from the average energy of the MDS. The QSAR model grid cells from model 3 of Table III b are also shown.

The prediction based on the partition coefficient alone (eqn. 1) for compound 13, the most active compound, yields a value substantially larger than is measured. The same logP-only model predicts too low an activity for compound 8. On the other hand, the 4D QSAR model 3 yields good predicted activities for both compounds. This implies that the grid cell occupancy descriptors for these two compounds provide important information complementary to lipophilicity. Thus, both compounds are especially suited to serve jointly for a representation of the relationship between the conformational profiles and significant 4D-QSAR grid cells.

Compound 13 occupies the key grid cells (4*1*9) and (2*-2*11). The occupancy of the latter grid cell with atoms of any type, shows a strong, positive correlation with activity. Cell (4*1*9) gives a weak, negative correlation when occupied by any type of atom. The "second" aromatic moiety of 13 can, as well as all other biphenyl congeners in the training set, occupy these two grid cells. Grid cell (5*0*13) is negatively correlated to activity and is only accessible to the alkyl chains of the p-alkoxy substituted compounds 14 and 15. Accordingly, both

of these compounds exhibit a lower antiarrhythmic activity than the p-chloro substituted compound 13.

Replacement of $^{45}Ca^{2+}$-ions from Phosphatidylserine Monolayers: 4D-QSPR Analysis

The enhancing influence of lipophilicity on the ability to replace $^{45}Ca^{2+}$ is not unexpected, since lipophilicity, in part, should govern the binding behaviour of cationic-amphiphilic

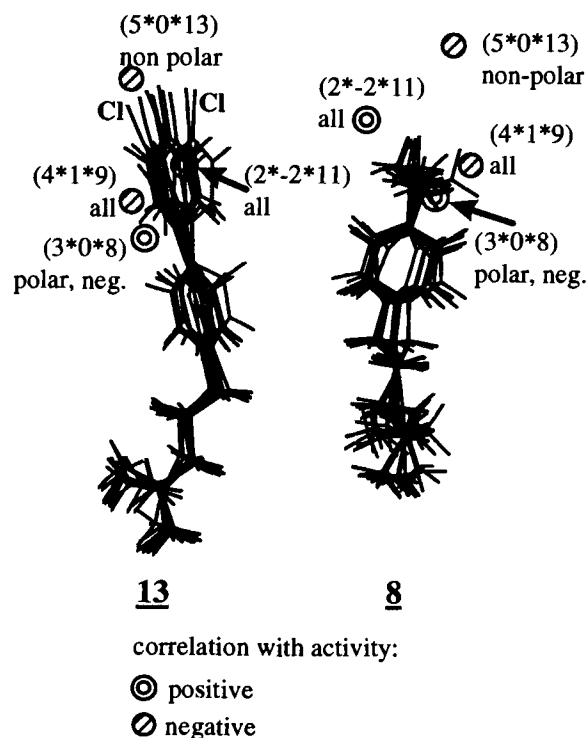correlation with activity:

◎ positive

⊘ negative

Fig. 2. Graphical representation of the grid cell descriptors of model 3, Table 4B, along with 10 aligned conformers of compounds 212 and 217, respectively. The perspective is different for the two compounds. Key grid cells "distant" from the structures of 212 and 217 arise because of substituents on other analogs in the training set and conformational freedom available to the analogs.

Table VIIIb. Descriptor Usage Over the 10 Best Models, and Over the Entire GFA Population, Alignment I, Tt

Variable	% use in 10 best models	% use in whole population
7*–1*6*0	60	5
7*0*6*0	30	7
9*0*8*6	30	27
10*0*7*6	30	11
9*–1*8*6	30	12
all others	≤ 20	

The first three digits are x, y, z coordinates [Å] of the particular grid cell; fourth digit: atom type (see "methods" section)

compounds to a PS monolayer. LogP and molecular weight are correlated (r = 0.79), and therefore the presence of MW in 20% of the ten best models, and about 50% of all models can be interpreted as a substitution for logP. This notion is supported by the fact that no single QSAR model contains both descriptors.

Phase Transition Temperature of DPPA Liposomes: 4D-QSPR Analysis

Lipophilicity, the most important descriptor for the two other endpoints considered in this study, is virtually uncorrelated to phase transition temperature, T_t, and does not appear in any of the ten best models. The independence of lipophilicity on phase transition behaviour of DPPA liposomes has been demonstrated in a qualitative manner previously by Borchardt *et al.* (29). In this study, compounds 5 and 18 were compared. Although these two compounds have about the same lipophilicity, their potency to influence T_t is different.

Using alignment 1 and the best scored model obtained for this alignment, predictions of T_t for the chiral compunds 19 and 20 were made. In both cases T_t had been measured for the racemates. The predicted T_t for compound 20 is in agreement with the measured value. On the other hand, the QSPR model fails to correctly predict the influence of compound 19 on the phase transition behaviour of DPPA liposomes.

The rather low predictiveness of the 4D-QSPR models for T_t might be explained from a physicochemical point of view. Unlike a proteinacious ion channel, the phospholipid membrane may not offer spatially defined, specific binding sites, mirrored by distinct and significant grid cell descriptors in a 4D-QSAR model.

4D-QSAR Analysis

This is only the second application of 4D-QSAR analysis to an *in vivo* activity endpoint, AC_{50}. In the previous *in vivo* application (1) logP, and lipophilicity substructure measures, were included in the GFA descriptor basis set. However, none of these lipophilicity descriptors survived to be among the descriptors of the best 4D-QSAR models. Thus, it is encouraging from this study to find logP the major descriptor for the best AC_{50} models, and used in combination with significant grid cell occupancy descriptors.

4D-QSAR analysis has not been used in QSPR model development previously. The IC_{50} and T_t QSPR models developed using 4D-QSAR analysis seem plausible. The lack of predictiveness in the T_t models may be a reminder of the limitations of using a QSAR approach highly dependent on spatial properties. If the endpoint property (T_t) is not dependent on specific spatial interactions, then reliable QSAR models cannot be expected. This reasoning supports the use of "2D" QSAR descriptors as part of the basis set of trial descriptors in the GFA phase of 4D-QSAR model development.

REFERENCES

1. A. J. Hopfinger, S. Wang, J. S. Tokarski, B. Jin, M. Albuquerque, P. Madhav, and C. Duraiswami. *J. Am. Chem. Soc.*, **119**: 10509–10524 (1997).
2. R. H. de Jong. *Local Anesthetics.* Mosby-Year Books Ltd., St. Louis, 1994.
3. H. Lüllmann, K. Mohr, A. Ziegler, and D. Bieger. *Pocket Atlas of Pharmacology.* Thieme Medical Publishers, Inc., New York, 1993.
4. R. S. Sheldon, N. J. Cannon, A. S. Nies, and H. J. Duff. *J. Pharmacol. Exp. Ther.* **33**:327–331 (1988).
5. L. M. Hondeghem and J. W. Mason. In: B. G. Katzung [ed.], *Basic and Clinical Pharmacology*, pp 151–168. Appleton & Lange, Norwalk (1987)
6. G. R. Strichartz. *Handbook of Experimental Pharmacology*, Vol. 81: Local Anesthetics. Springer Verlag, Berlin, Germany, 1987.
7. E. Overton. *Vierteljahresschrift Naturforsch. Ges. Zürich*, **44**:88–135 (1899).
8. A. G. Lee. *Nature* **262**:545–548 (1976).
9. M. Klingmüller. Ph.D. Thesis, Kiel, Germany, 1990.
10. S. I. Landmann. Med. Thesis., Kiel, Germany (1996).
11. S. Girke, K. Mohr, and S. Schrape. *Biochem. Pharmacol.* **38**:2487–2496 (1989).
12. H. Lüllmann, H. Plösch, and A. Ziegler. *Biochem. Pharmacol.* **29**:2969–2974 (1980).
13. B. Kursch, H. Löllmann, and K. Mohr. *Biochem. Pharmacol.* **32**:2589–2594 (1983).
14. R. Hanpft and K. Mohr. *Biochim. Biophys. Acta* **814**:156–162 (1985).
15. R. Hanpft. Ph.D. Thesis, Kiel, Germany, 1987.
16. R. D. Cramer, D. E. Patterson, and J. D. Bunce. *J. Am. Chem. Soc.* **110**:5959–5967 (1988).
17. W. G. Glen, W. J. Dunn III, and D. R. Scott. *Tetrahedron Computer Methodology* **2**:349–376 (1989).
18. D. Rogers and A. J. Hopfinger. *J. Chem. Inf. Comput. Sci.* **34**:854 (1994).
19. W. J. Dunn III, and D. Rogers. Genetic partial least-squares in QSAR. In J. Devillers (ed.), *Genetic Algorithms in Molecular Modeling*, Academic Press, London, 1996, pp. 109–130.
20. K. Hasegawa, Y. Miyashita, and K. Funatsu. *J. Chem. Inf. Comp. Sci.* **37**:306–310 (1997).
21. *CHEMLAB-II, Revision 11.0.* Molecular Simulations Incorporated, San Diego, CA (1993).
22. J. J. P. Stewart. *Mopac 6.0 Manual.* Frank J. Seiler Research Laboratory, United States Air Force Academy, CO 80840 (1990).
23. *Molsim Version 3.0.* The Chem21 Group, Inc., Lake Forest, IL 60045 (1994).
24. D. Rogers. *WOLF 5.5 Release Notes* (1993).
25. *WOLF 6.2 GFA Program.* D. Rogers and Molecular Simulations Inc., San Diego, CA (1994).
26. J. Friedman. *Multivariate Adaptive Regression Splines.* Technical report No. 102, Laboratory for Computational Statistics, Dept. of Statistics, Stanford University (Nov. 1988).
27. D. O. Rauls and J. K. Baker. *J. Med. Chem.* **22**:81–86 (1979).
28. H. Lüllmann, P. Timmermans, and A. Ziegler. *Eur. J. Pharmacol.* **60**:277–285 (1979).
29. K. Borchardt, D. Heber, M. Klingmüller, K. Mohr, and B. Müller. *Biochem. Pharmacol.* **42**(Suppl.):S61–S65 (1991).