Toward a Pharmacophore for Drugs Inducing the Long QT Syndrome: Insights from a CoMFA Study of HERG K^+ Channel Blockers

Andrea Cavalli,[†] Elisabetta Poluzzi,[‡] Fabrizio De Ponti,[‡] and Maurizio Recanatini*[,†]

Department of Pharmaceutical Sciences, University of Bologna, Via Belmeloro 6, I-40126 Bologna, Italy, and Department of Pharmacology, University of Bologna, Via Irnerio 48, I-40126 Bologna, Italy

Received March 25, 2002

In this paper, we present a pharmacophore for QT-prolonging drugs, along with a 3D QSAR (CoMFA) study for a series of very structurally variegate HERG K^+ channel blockers. The blockade of HERG K^+ channels is one of the most important molecular mechanisms through which QT-prolonging drugs increase cardiac action potential duration. Since QT prolongation is one of the most undesirable side effects of drugs, we first tried to identify the minimum set of molecular features responsible for this action and then we attempted to develop a quantitative model correlating the 3D stereoelectronic characteristics of the molecules with their HERG blocking potency. Having considered an initial set of 31 QT-prolonging drugs for which the HERG K^+ channel blocking activity was measured on mammalian transfected cells, we started the construction of a theoretical screening tool able to predict whether a new molecule can interact with the HERG channel and eventually induce the long QT syndrome. This in silico tool might be useful in the design of new drug candidates devoid of the physicochemical features likely to cause the above-mentioned side effect.

Introduction

The long QT syndrome (LQTS) is characterized by the prolongation of the QT interval of the surface electrocardiogram and is associated with an increased risk of torsades de pointes, a ventricular tachyarrhythmia that may degenerate into ventricular fibrillation and sudden death.[1] Several congenital and acquired disorders can lead to prolongation of the QT interval; of special interest is the fact that numerous agents, belonging to different drug classes, have been associated with QT prolongation and occurrence of torsades de pointes. Recently, several regulatory interventions have involved drugs for which this potentially fatal risk was recognized only after marketing authorization.[2] Screening methods for the early detection of an effect on the QT interval are therefore required during the drug development process, and several in vivo and in vitro methods are now available.[3,4]

The QT interval is defined as the time interval between the onset of the QRS complex and the end of the T wave and therefore includes both the ventricular depolarization and repolarization intervals. Although several pathophysiological mechanisms can lead to prolongation of the QT interval,[5−7] the key mechanism for drug-induced QT prolongation is the increased repolarization duration through blockade of outward K^+ currents (especially the delayed rectifier repolarizing current, I_K). In particular, most of the QT-prolonging drugs have been shown to inhibit the K^+ channels encoded by the human ether-à-go-go related gene (HERG), at the basis of the rapid component of I_K named I_{Kr}.[8−12] HERG K^+ channel blockade is therefore the most important mechanism through which QT-prolonging drugs increase cardiac action potential duration.[13] Notably, blockade of HERG K^+ channels forms the basis of the therapeutic effect of class III antiarrhythmic drugs, but for all other drugs, it is an unwanted side effect that must be detected as early as possible during drug development. IC_{50} for inhibition of HERG K^+ channels expressed in different cell lines is considered a primary test to study the QT-prolonging potential of a compound.[14]

Since none of the existing in vitro tests to assess the QT-prolonging potential of a compound has an absolute predictive value,[4] the availability of in silico methods in the early phase of drug development would dramatically increase the screening rate and would also lower the costs compared to experimental assay methods. At the present time, to our knowledge, no theoretical screening method for the evaluation of QT-prolonging properties of molecules is publicly available. In an attempt to explore the possibility of building an in silico screening system based on the known structure−activity relationships (SAR) of LQTS-inducing noncardiac drugs, we report here a pharmacophore for such drugs, based on an initial set of molecules taken from an organized list of QT-prolonging compounds.[15] Furthermore, we present a 3D QSAR model obtained by means of the CoMFA technique that attempts to correlate the physicochemical features of the drug molecules with their blocking activity toward the HERG K^+ channel. The obtained CoMFA model was validated by predicting the biological activities of a set of molecules not used in deriving the 3D QSAR equations. We propose this model as a first step toward the development of a tool able to predict whether a molecule of pharmacological interest bears structural features likely to induce LQTS.

* To whom correspondence should be addressed. Phone: +39 051 2099720. Fax: +39 051 2099734. E-mail: mreca@alma.unibo.it.
[†] Department of Pharmaceutical Sciences.
[‡] Department of Pharmacology.

10.1021/jm0208875 CCC: $22.00
Published on Web 07/25/2002

Chart 1. QT-Prolonging Drugs Considered in the Paper[a]

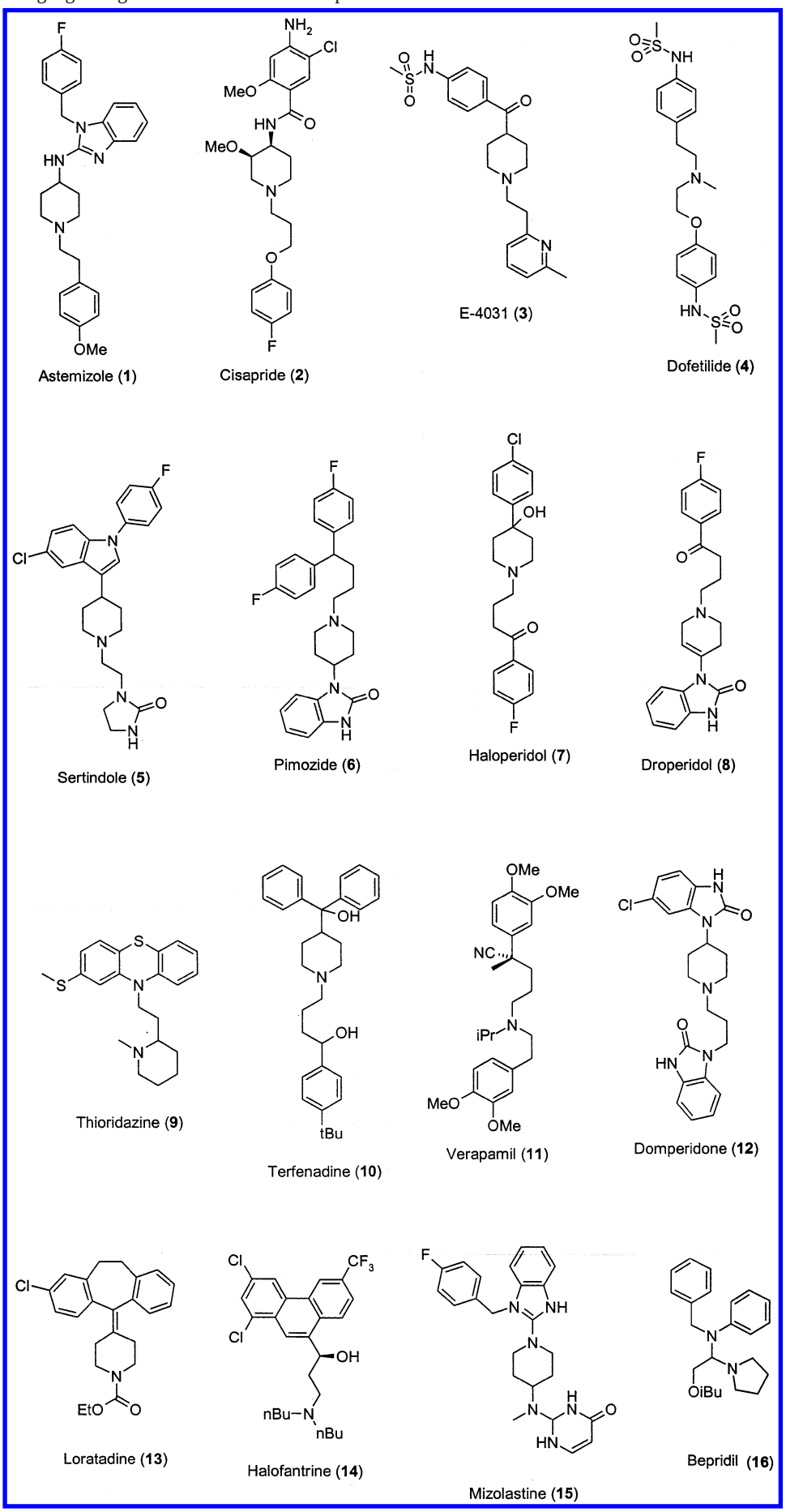


Chart 1 (Continued)

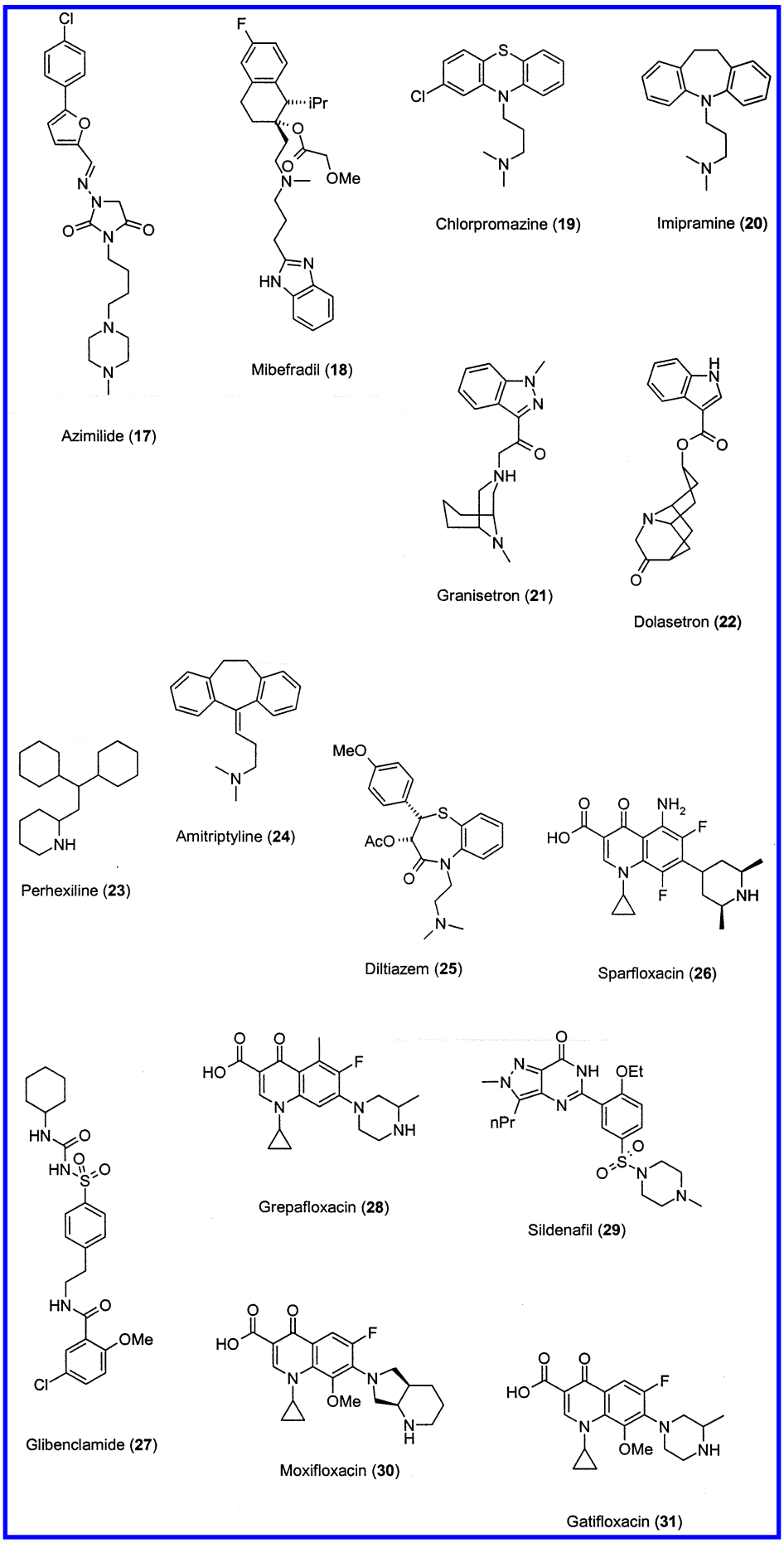
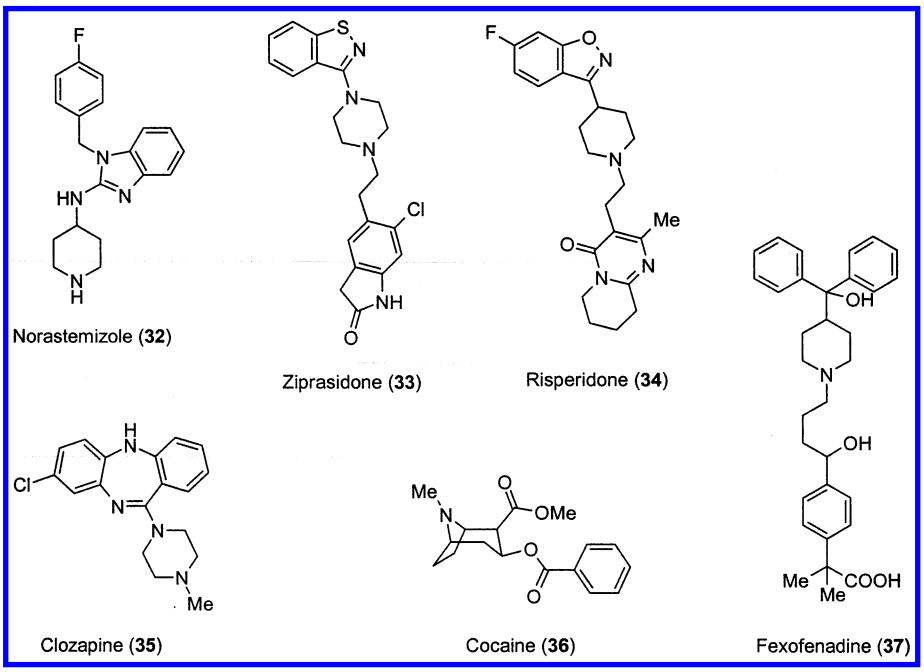


Chart 1 (Continued)



[a] In the CoMFA analysis, compounds **1**−**31** and **32**−**37** were the training and the test sets, respectively.

Methods

1. QT Prolonging Drugs Selected for the Study. In principle, a pharmacophore accounting for a certain pharmacological action should be built by taking into consideration all (or most) of the compounds able to elicit such an action. In the case of the QT-prolonging effect, besides the class III antiarrhythmic drugs, several other agents are known to induce this response and new entries are reported almost monthly at an accelerated pace. Indeed, prolongation of the QT interval by nonantiarrhythmic drugs is not an unusual finding, but the yet-unanswered question is to what degree QT prolongation has to be considered clinically significant. In this regard, in recent papers, we proposed a list of noncardiac QT-prolonging drugs, based on specific clinical and nonclinical criteria, and aimed to make a starting point to maintain a "consensus list" to be periodically updated.[15] It seemed reasonable to consider this database (about 140 compounds at publication time) as the source of molecules to be used in the construction of the pharmacophore. However, since the pharmacophoric scheme must refer to a well-defined molecular target (i.e., the HERG K^+ channel), we selected from the whole list only those drugs for which HERG K^+ channel inhibition had been reported. The final set of molecules shown in Chart 1 was obtained by combining our list[15] with Fenichel's database,[16] which also includes drugs used as antiarrhythmics; we selected only those drugs for which IC_{50} values for inhibition of HERG K^+ channels expressed in mammalian cells (HEK, CHO, COS, neuroblastoma cells; see Table 1) were available. We decided not to include IC_{50} values obtained in nonmammalian cell lines, such as *Xenopus laevis* oocytes, since it is now recognized that the use of these systems leads to a significant underestimation of a drug's potency as a HERG K^+ channel blocker (the highly lipophilic environment in *Xenopus* oocytes limits access of the drug to its site of action; experiments are carried out at 22 °C instead of 37 °C).[14,17]

Table 1. Observed and Calculated HERG K^+ Channel Blocking Activity of Compounds **1**−**31**

compound	IC_{50} (nM)	pIC_{50obsd}	pIC_{50fit}[a]	Δ
astemizole (**1**)	0.9[b]	9.04	8.53	0.51
cisapride (**2**)	6.5[b]	8.19	7.96	0.23
E-4031 (**3**)	7.7[b]	8.11	7.85	0.26
dofetilide (**4**)	9.5−15[b]	7.91	7.67	0.24
sertindole (**5**)	14[b]	7.85	8.04	−0.19
pimozide (**6**)	18[b]	7.74	7.80	−0.06
haloperidol (**7**)	28.1[b]	7.55	7.58	−0.03
droperidol (**8**)	32.2[b]	7.49	7.82	−0.33
thioridazine (**9**)	35.7[b]	7.45	7.23	0.22
terfenadine (**10**)	56−204[b]	6.89	7.22	−0.33
verapamil (**11**)	143[b]	6.84	7.05	−0.21
domperidone (**12**)	162[b]	6.79	6.88	−0.09
loratadine (**13**)	173[b]	6.76	5.83	0.93
halofantrine (**14**)	196.9[c]	6.70	6.81	−0.11
mizolastine (**15**)	350[b]	6.45	6.65	−0.20
bepridil (**16**)	550[d]	6.26	6.30	−0.04
azimilide (**17**)	560[c]	6.25	6.15	0.10
mibefradil (**18**)	1430[d]	5.84	5.75	0.09
chlorpromazine (**19**)	1470[c]	5.83	5.68	0.15
imipramine (**20**)	3400[c]	5.47	5.98	−0.51
granisetron (**21**)	3730[b]	5.42	5.64	−0.22
dolasetron (**22**)	5950[b]	5.22	4.99	0.23
perhexiline (**23**)	7800[b]	5.11	5.18	−0.08
amitriptyline (**24**)	10000[b]	5.00	5.66	−0.66
diltiazem (**25**)	17300[b]	4.76	5.02	−0.26
sparfloxacin (**26**)	18000−34400[c]	4.58	4.39	0.19
glibenclamide (**27**)	74000[e]	4.13	4.07	0.06
grepafloxacin (**28**)	50000−104000[c]	4.11	4.35	−0.24
sildenafil (**29**)	100000[b]	4.00	3.50	0.50
moxifloxacin (**30**)	103000−129000[c]	3.93	3.82	0.11
gatifloxacin (**31**)	130000[c]	3.89	4.16	−0.27

[a] Calculated from the non-cross-validated CoMFA model. [b] In human embryonic kidney (HEK) cells. [c] In Chinese hamster ovary (CHO) cells. [d] In African green monkey kidney derived cell line COS-7. [e] In neuroblastoma cells.

Compounds **1**−**31** were used for the pharmacophore generation, and they also constituted the training set for the CoMFA procedure. In this connection, for each drug, the HERG blocking potency expressed as an IC_{50} value is reported in Table 1. It is remarkable that the drugs considered in this study span a potency interval as HERG K^+ channel blockers of more than 5 log units.

To assess the predictive ability of the CoMFA model, we considered a test set of molecules (**32**−**37**, Chart 1) taken from Fenichel's list, or whose activity as a HERG K^+ channel blocker was reported after the publication of our list (e.g., cocaine[18,19]).

2. Molecular Modeling. Here, we describe the construction of the three-dimensional (3D) models of the molecules, the conformational search followed by a cluster analysis, the pharmacophore generation, and the 3D QSAR analysis based on the comparative molecular field analysis (CoMFA) procedure.[20] All the molecular modeling studies were carried out by means of the SYBYL[21] and MacroModel[22] software running on a Silicon Graphics workstation.

2.1. Construction of the Models. The 3D models of the molecules were either retrieved[23] from the Cambridge Structural Database (CSD),[24] or modeled by adding functional groups on crystallographic skeletons. Thus, very few models were built by assembling fragments retrieved from the standard SYBYL library. The following molecules were directly retrieved from the CSD (the CSD code is in parentheses): astemizole, **1** (ZENREP); cisapride, **2** (KEYOB); pimozide, **6** (PIMOZD); haloperidol, **7** (HALOPB); droperidol, **8** (KAMCIK); thioridazine, **9** (WAVCEB); verapamil, **11** (CURHON); domperidone, **12** (BEQJUC); loratadine, **13** (YOVZEO); halofantrine, **14** (SATRAG); bepridil, **16** (CEZBAK); chlopromazine, **19** (DUKTOS); imipramine, **20** (IMIPRB); perhexiline, **23** (DEPGUA01); amitriptyline, **24** (YOVZEO); diltiazem, **25** (CEYHUJ01); sparfloxacin, **26** (COQWOU); glibenclamide, **27** (DUNXAL); sildenafil, **29** (CAXZEG); risperidone, **34** (WASTEP); clozapine, **35** (CMPDAZ10); and cocaine, **36** (COCAIN10). The following molecules were built by adding fragments to crystallographic skeletons: E-4031, **3**, was built starting from BERGUA; terfenadine, **10**, from YIHJOO; mizolastine, **15**, from CELNUG and LEJKUG; azimilide, **17**, from LIAXBUD; mibefradil, **18**, from BZDMAZ; granisetron, **21**, from KUSZED; dolasetron, **22**, from KAMCIK; grepafloxacin, **28**, and moxifloxacin, **30**, from NIVQAK; gatifloxacin, **31**, from COQWOU; ziprasidone, **33**, from DEDCIY, NUXWAE, and JIVAO. Norastemizole (**32**) and fexofenadine (**37**) are metabolites of astemizole (**1**) and terfenadine (**10**), respectively, and were built by modifying the structure of the parent compounds. Finally, only two molecules were built de novo, namely, dofetilide, **4**, and sertindole, **5**.

The molecular models obtained were first energy-minimized by using steepest descent and conjugate gradient until a convergence of 0.005 kJ mol^{-1} $Å^{-1}$ on the gradient was reached. Then, conformational searches were carried out in order to sample the potential energy surface. All of the classical molecular mechanics calculations were performed by using the MMFF force field,[25,26] which carries parameters for all the investigated molecules. Finally, the average conformers obtained from a cluster analysis (see below) were optimized by using the semiempirical Hamiltonian PM3[27] as implemented in the SYBYL package (keywords GNORM = 0.001, MMOK when needed).

2.2. Conformational Search. The conformational space of each molecule was sampled by means of Monte Carlo analysis[28] as implemented in the MacroModel software. In the Monte Carlo approach, the dynamical behavior of a molecule is simulated by randomly changing dihedral angle rotations or atom positions. Then, the trial conformation is accepted if its energy has decreased from the previous one. If the energy is higher, various criteria can be applied to determine whether the new conformation should be accepted or not. In our simulations, all of the dihedral angles of single linear bonds were allowed to move freely, and the trial conformation was accepted if the energy was lower than that of the previous conformation or if the energy was within a fixed energy window (100 kJ/mol). The number of Monte Carlo trials was set equal to 7000. A high number, usually in the order of hundreds, of conformers was generated. To classify the conformations obtained for each molecule, a geometrical cluster analysis was carried out on the output of Monte Carlo searches.

2.3. Cluster Analysis. Generally, a cluster analysis[29] provides for the most significant solutions, by using filtering screens based on one or more external criteria. Conformations were classified in terms of geometrical similarity, and in particular, two conformers were considered as belonging to the same family when the heavy atom's root-mean-square displacement (rmsd) was lower than 1 Å. In this way, between 10 and 20 families of conformers were generated for each molecule, and the average structure of each family was energy-minimized and used in the procedure of pharmacophore generation.

2.4. Pharmacophore Generation. An inspection of molecules **1**−**31** of Chart 1 shows that they are widely varied from a structural point of view, which implies an objective difficulty in the search for a common pharmacophore. To tackle this problem, we adopted a "constructionist" approach to the pharmacophore generation, consisting of the individuation of an initial reference structure (the template) onto which we overlapped molecules starting from those with similar geometric and spatial characteristics. This superimposition procedure led to the individuation of further pharmacophoric characteristics that were then used to add the most different molecules and to refine the initial alignment. The template chosen for this study was the crystal structure of astemizole (**1**), because this molecule is one of the most potent long QT-inducing drugs, and HERG channel blockers (IC_{50} = 0.9 nM, Table 1). The astemizole crystal structure was directly retrieved from the CSD and then geometrically optimized by means of PM3. Initially, three pharmacophoric points on the astemizole molecule were defined, namely, the basic nitrogen of the piperidine cycle (N) and the centers of mass (centroids C0 and C1) of the two close aromatic moieties. However, considering that several molecules of the training set bear an aromatic group connected to the basic nitrogen, a fourth pharmacophoric point was defined as the centroid (C2) of the phenyl ring belonging to the *N*-(*p*-methoxyphenylethyl) substituent of astemizole. Onto this template, the best-fitting conformer (within 20 kJ/mol from the global minimum) of the other molecules was then superimposed by using the corresponding pharmacophoric functions.

Not all the molecules displayed all four pharmacophoric points, and in such cases, the superimposition was based on the available points or it was further guided by other functions present on some molecules. Particu-

Table 2. Summary of the Pharmacophore Geometrical Parameters

distances (Å)		angles (deg)		angles between planes[a] (deg)		height above planes[a] (Å)	
C0−C1	4.3−6.7	C0−N−C2	134−173	P0−P1	50	N−P3	0.6
C0−C2	10.3−14.7	C0−N−C1	30−65	P0−P2	160	C0−P2	1.7
C1−C2	11.9−12.3	C1−N−C2	132−152	P1−P2	114	C1−P1	3.4
C0−N	5.2−9.1					C2−P0	1.2
C1−N	5.7−7.3						
C2−N	4.6−7.6						

[a] P0 is the plane containing C0, C1, and N. P1 is the plane containing C0, N, and C2. P2 is the plane containing C1, N, and C2. P3 is the plane containing C0, C1, and C2.

larly, in the case of compounds **2**, **5**−**8**, **12**−**15**, **17**, and **18**, the halogen atom located in the para position on one phenyl ring (C0) was used to reinforce the fit. As regards the quinolones, which are the least potent HERG channel blockers considered in this study, their structures were quite different from the structure of the template, which implied that they were superimposed to astemizole by first anchoring the molecular skeleton to the basic piperazine N atom. This oriented the centroid of the 4-piperidinone ring onto C0.

In Figure 1a, the overall superimposition of compounds **1**−**31** is shown and the positions of the pharmacophoric points are indicated. In particular, the simplified pharmacophoric scheme shown in Figure 1b was obtained by connecting the average spatial positions of the pharmacophoric points of all the molecules. Distances and angles defining the pharmacophore geometry are reported in Table 2.

2.5. 3D QSAR through CoMFA. The most critical point in the CoMFA procedure is the alignment of the molecules in Cartesian space, which is usually accomplished by following some hypothesis on the binding of compounds to the active site of their biological counterpart. In this study, the above-described pharmacophore based on astemizole was the starting point for the overall CoMFA alignment (training and test sets). Actually, the pharmacophoric superimposition was exploited to obtain the alignment for the 3D QSAR analysis, and in the cases where different conformations or orientations of the molecules were possible, the conformation/orientation giving the best statistical results was chosen.

A CoMFA table was built containing the biological activity data of the HERG channel blockers (the dependent variables, pIC_{50obsd}; Table 1) and the values of steric and electrostatic fields at discrete points of the Cartesian space surrounding the molecules (the independent variables). The fields were generated by using an sp^3 carbon atom with a formal charge of +1 as a probe. The region was generated automatically around the molecules by fixing a grid spacing of 1 Å. The statistical analysis was carried out by applying the PLS procedure to the appropriate independent variables and using the standard scaling method (COMFA_STD). Furthermore, to reduce the number of independent variables, an energy cutoff value of 30 kcal/mol was selected for both electrostatic and steric fields. The minimum σ for further filtering of the independent variables was set to 2.0 kcal/mol. Cross-validated PLS runs were carried out to establish the optimal number of components (the latent variables) to be used in the final fitting models. The number of cross-validated groups was always equal to the number of compounds (leave-one-out procedure), and the optimal number of latent variables was chosen by considering the lowest standard error of prediction (s_{cross}).

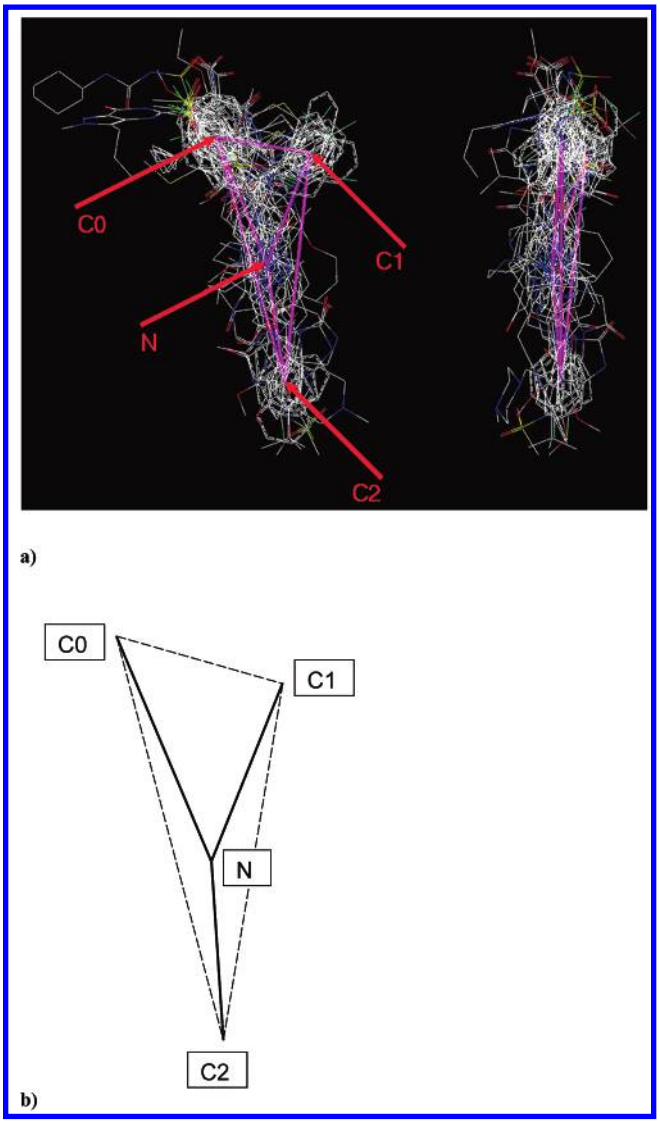


Figure 1. (a) Orthogonal views of the superimposition of compounds **1**−**31** from which the pharmacophore was obtained. In the left view, the pharmacophoric frame is shown in magenta and the centroids of the pharmacophoric functions are indicated. (b) "QT pharmacophore" (geometric details are reported in Table 2).

Results

Pharmacophore. The pharmacophore for the QT-prolonging potential obtained from the drugs shown in Chart 1 (**1**−**31**) is depicted in Figure 1b. How this pharmacophore was derived from the set of molecules considered in the present work is described above and illustrated in Figure 1a. All conformers aligned repre-

Table 3. Summary of the CoMFA Statistical Parameters

q^2	0.767
s_{cross}	0.744
F	172.87
r^2	0.952
s	0.336
optimal number of components	3
steric field contribution	0.716
electrostatic field contribution	0.284
r^2_{pred} [a]	0.744

[a] $r^2_{pred} = (SD - PRESS)/SD$, where SD is the sum of the squared deviations of each observed activity value for each molecule of the test set from the mean of the observed activity values of the training set (pIC_{50obsd} values of Table 1) and PRESS is the sum of the squared deviations between predicted and observed values (Δ of Table 4).

sent low-energy conformations of the molecules, and it can be seen that the final alignment shows a satisfactory superimposition of the pharmacophoric points. The pharmacophore is made by three aromatic moieties connected through a nitrogen function that is a tertiary amine throughout the whole set of molecules. This renders the compounds protonated at physiological pH, and the presence of the amine function in all the molecules might reasonably suggest the importance of a positively charged group in conferring biological activity. The three aromatic moieties of the pharmacophore schematically represented by C0, C1, and C2 are not carried by all the molecules. Actually, while C0 is always present in the QT-prolonging drugs here considered, only some molecules bear the functions identified as C1 and C2. Moreover, a few drugs of the set also show a polar or polarizable function (mostly carboxylic or sulfonamidic groups, or halogen atoms, respectively) on the C0 aromatic ring (Figure 1a), seemingly responsible for further modulation of the activity.

To better describe the pharmacophoric features, some geometric parameters are reported in Table 2. It is immediately apparent that the ranges of distances and angles between the points are rather wide, which might be a consequence of the relevant structural diversity of the set of molecules. However, the pharmacophoric scheme was drawn in the simplest way, in view of its application to a growing number of structurally diverse molecules.

How all the chemical functions in the pharmacophore can affect the biological activity toward the HERG K^+ channel of QT-prolonging drugs will be discussed in light of the 3D QSAR analysis results.

CoMFA Model. The results of the CoMFA analysis performed on the training set of the HERG K^+ channel blockers **1**−**31** are shown in Table 3, where the main statistics for both cross-validated and non-cross-validated PLS analyses are reported. The 3D QSAR model selected on the basis of the minimum s_{cross} value criterion has an optimal number of components equal to 3, and descriptive and predictive abilities are evaluated by the statistic parameters $r^2 = 0.952$ and $s = 0.336$, and $q^2 = 0.767$ and $s_{cross} = 0.744$, respectively. The predictive properties of the CoMFA model were more rigorously tested by calculating the HERG blocking potency of a set of molecules not used in the derivation of the model (**32**−**37**, Chart 1). The result is an $r^2_{pred} = 0.744$, which is satisfactory. Note that despite the small number of compounds considered in the test

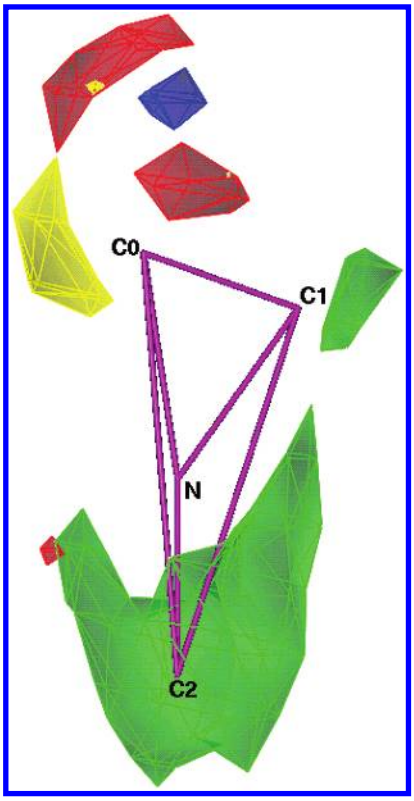


Figure 2. View of the steric and electrostatic CoMFA STDEV*COEFF contour maps. The regions where increasing the molecules' volume increases HERG K^+ channel blocking activity are green (0.028 level), and the region where increasing the volume decreases activity is yellow (−0.022 level). The electrostatic contours indicate an increase of activity with increasing positive (red, 0.010 level) and negative (blue, −0.012 level) charge, respectively. The pharmacophoric frame is shown for reference.

set, the variation of potency is rather high (about 3 log units), and it is recalculated by the model in an acceptable way, even if the extreme points (**32** and **37**) are not well predicted (but only **32** exceeds the s_{cross} value, Table 3).

The CoMFA model accounting for the 3D QSAR of compounds **1**−**31** is illustrated by the contour maps shown in Figure 2. Sterically favorable regions (green) are located around the pharmacophoric points C1 and mostly C2, while the space around C0 seems sensitive to both steric and electrostatic properties of the molecules. Particularly, the yellow contour indicates that increasing bulk is detrimental for the activity, while the red and blue contours (which are not, as it might appear, on the same line but show the blue volume pointing outward with respect to the red ones) indicate a prevalent favorable effect of positively and negatively charged groups, respectively.

In Figure 3, the most (Figure 3a) and the least (Figure 3b) potent HERG K^+ channel blockers of the series (astemizole, **1**, and gatifloxacin, **31**, respectively) are shown along with the CoMFA contour maps in the orientation they assume in the CoMFA alignment. By inspection of the figures, it immediately appears that while relevant portions of astemizole indeed protrude within the favorable steric regions around C1 and C2 (green), gatifloxacin does not make contact with such regions with any of its parts. On the contrary, this molecule partially reaches the forbidden steric contour

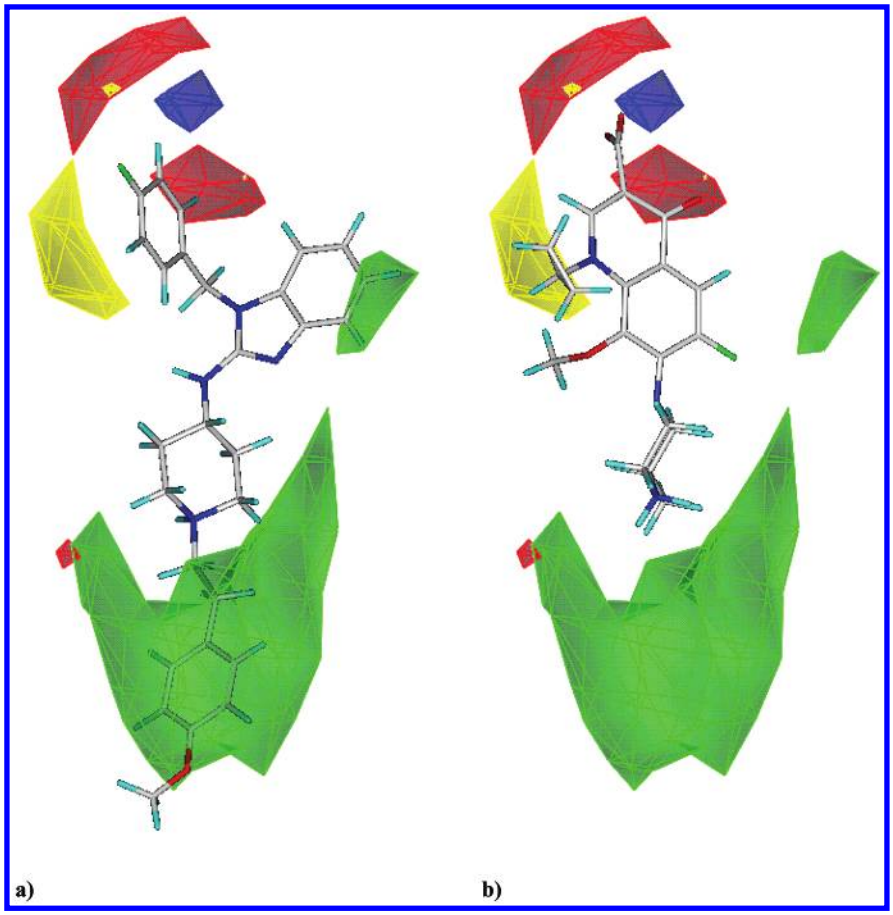


Figure 3. Different occupancy of the CoMFA regions by astemizole, **1** (a), and gatifloxacin, **31** (b).

around C0 (yellow), and moreover, its carboxylate function falls close to the electrostatic red region over C0, where positively charged groups would be required for an optimal HERG K^+ channel blocking activity.

Discussion

In this study, we present a pharmacophoric model for a series of structurally different molecules able to induce the LQTS. This pharmacophore (Figure 1b and Table 2) covers a number of prototypical structures recognized to be responsible for the LQTS. The presence of four relevant pharmacophoric points allowed us to allude to more than classical pharmacophoric geometrical elements (such as distances and angles between points). Actually, in our "QT pharmacophore", angles between the three different planes generated by the four pharmacophoric points, and heights of the pharmacophoric points above the planes were also identified. In principle, one might suggest that these additional geometrical elements should be taken into consideration when designing new drugs, to avoid the synthesis of molecules bearing such geometrical features and to minimize the risk of induced LQTS. On the other hand, we want to emphasize the "in progress" character of the present pharmacophore because of the growing number of new molecules endowed with this undesired biological property. Consequently, as the diversity of the molecular structures associated with LQTS increases, new chemical and geometrical features will have to be taken into account in the continuous process of pharmacophore building and refining. For instance, there is some not yet clearly defined crowding of polar groups in the region between N and C2 (Figure 2a), which might eventually result in the identification of a further pharmacophoric feature.

In an effort to refine the pharmacophore description by providing at the same time a tentatively predictive tool, we then performed a CoMFA analysis on the same set used for the construction of the pharmacophore, employing the HERG K^+ channel blocking data of the molecules as the dependent variable. This implied that we assumed the HERG K^+ channel blockade as the molecular mechanism through which drugs **1**−**31** cause the prolongation of the QT interval. We are aware of the fact that HERG K^+ channel blockade is the main, but not the exclusive, mechanism leading to QT interval prolongation;[5−7] however, for the compounds taken into consideration, experimental data exist confirming such hypothesis.[15] On the other side, relating the LQTS-inducing potential of the compounds of Chart 1 to the measured action against a defined biological target allowed us to attempt an interpretation of the 3D QSAR at the molecular level.

Caution must be exerted when using biological data from different sources in QSAR studies: actually, the IC_{50} values listed in Table 1 were obtained from different cell lines expressing the HERG K^+ channel and this might, in principle, raise a question about the feasibility of the 3D QSAR investigation. This is the reason we carefully selected, for the training set, a relatively small number of compounds (31 out of a list of 140 drugs in the database), including only those drugs for which IC_{50} values for the inhibition of HERG K^+ channels expressed in mammalian cells were available. We believe

Table 4. Observed and Calculated HERG K^+ Channel Blocking Activity of Compounds **32**−**37**

compound	IC_{50} (nM)	PIC_{50obsd}	pIC_{50fit}[a]	Δ
norastemizole (**32**)	28[b]	7.55	6.72	0.83
ziprasidone (**33**)	152[b]	6.82	6.92	−0.10
risperidone (**34**)	163b	6.79	6.99	−0.20
clozapine (**35**)	191[b]	6.72	6.18	0.54
cocaine (**36**)	4.400−7.200[b]	5.24	5.44	−0.17
fexofenadine (**37**)	21570[b]	4.67	5.33	−0.66

[a] Calculated from the non-cross-validated CoMFA model. [b] HEK cells.

that this approach minimized the inhomogeneity of in vitro data and allowed us to use a larger number of compounds as a training set than would have been feasible using only our own experimental data for HERG K^+ channel blockade. Indeed, the statistics of the model can be used as a guide to evaluate the significance of the results. As shown from the data of Table 3 and also from both the pIC_{50fit} and Δ values of Table 1, the QSAR model is reasonably good. Although it is not a definitive model, it can be used as the basis for further refining and continuous probing through the use of a larger number of compounds as more experimental data become available.

Thus, from this study, a 3D QSAR model for the HERG K^+ channel blocking potency was generated (Table 3). We acknowledge that lower standard deviation values (especially s_{cross}) would have been desirable. A reason for the rather high dispersion of the (cross-validated) predicted activity data might be related to the aforementioned inhomogeneity of the biological system where the experimental IC_{50} values were measured. Different sensitivity of the cellular systems to the molecular perturbation caused by the drug can be responsible for some noise in the set of activity data not accounted for by the structural variation of the molecules. Moreover, the possibility exists that the compounds under investigation bind HERG K^+ channels in different states (i.e., open and close) and/or in different domains of the protein. As a consequence of that, the physicochemical requirements for the binding may change from one compound to another in a way not simply correlated to the structure of the molecule.

Whether this might be the case also for loratadine (**13**, Table 1), which is severely underestimated by the model (Δ = 0.93, i.e., more than twice the standard deviation of the non-cross-validated regression) is hard to say. Indeed, loratadine is one the most rigid molecules of the whole set **1**−**37**, and this suggests an entropic factor (unaccounted for by the CoMFA field equations) possibly favoring its interaction with the channel protein. However, the similar molecule amitriptyline (**24**, Table 1) is overestimated by the CoMFA model (Δ = −0.66), which might reflect an intrinsic difficulty of the model to treat rigid molecules (perhaps in the alignment step).

Finally, the reliability of a QSAR model is usually assessed by testing its ability to predict the activity of an external set of molecules interacting with the same biological counterpart, and to this aim, we tested the CoMFA model by predicting the biological activity of six molecules not included in the training set (**32**−**37**, Table 4). It turned out that the statistical model was able to predict reasonably well the HERG K^+ channel blocking potency of these compounds, giving a predictive correlation coefficient (r^2_{pred}) equal to 0.744. As regards the underestimation of norastemizole (Table 4, **32**, Δ = 0.83), it might be noted that the parent compound astemizole was also underestimated (Table 1, **1**, Δ = 0.51), whereas the difference of potency between the two compounds (1.49) was acceptably well predicted (1.81).

Considering the characteristics of the pharmacophore and of the 3D QSAR model together (as in Figure 2) can give a broad idea of the physicochemical/spatial factors associated with the ability to induce the LQTS by blocking the HERG K^+ channel. These features describe a rather flexible molecule ca. 12 (±2) Å long (see distances C0−C1 and C1−C2 in Table 3) bearing a central tertiary amine function and at least two aromatic moieties centered around C0 and C2. A further aromatic fragment in a position corresponding to C1, and polar functions on one side of the molecule (C0 region) can favor the activity. Of course, these characteristics should be avoided in the design of candidate drugs.

All the considerations above-reported on both the pharmacophore and the 3D QSAR model are ligand-based; i.e., they do not take into consideration the structure of the macromolecular target but were derived only from the physicochemical and structural properties of the drugs. Target-based studies of SAR can be carried out provided that a model of the 3D structure of the target protein (either theoretical or X-ray-crystallography-derived) is available. A comparison of ligand- and target-based models independently developed for the same ligand−target system can be attempted when looking for a deeper interpretation of the binding interactions occurring between the two molecules, and we have shown the feasibility of the approach in the case of enzyme−inhibitor systems.[30,31] As regards the HERG K^+ channel, no resolved crystallographic structure is known, but a theoretical 3D model of the docking complex between the protein and the drug MK-499 was recently reported.[32] In the present case, we might try to relate the above-illustrated pharmacophoric and CoMFA features to the characteristics of the HERG K^+ channel binding site as described by Mitcheson et al.[32] These authors, on the basis of their and others'[33] work, point out the importance of two residues among those forming the binding pocket of the channel (Tyr652 and Phe656), suggesting the possibility of π-stacking interactions between the aromatic rings of the amino acid side chains and aromatic fragments of the drug molecule. Unfortunately, we could not include MK-499 in our model because its HERG K^+ channel blocking activity was determined in *Xenopus* oocytes (which leads to significant underestimation of the IC_{50}[14,17]), but the superimposition of its structure on our pharmacophoric frame reveals a perfect match between the methylsulfonamido-substituted phenyl and C0, the piperidine nitrogen and N, and the cyano-substituted phenyl and C2. Although in Mitcheson's work other residues of the binding cavity of HERG K^+ channel were discussed and their role in drug binding was probed by selective mutations, no mention was made of the possible role of the charged nitrogen function present in the drugs investigated there (and present in all the compounds investigated here). Remarkably, the inspection of the

stereoview of the MK-499/HERG complex reported in the paper reveals the presence of some aromatic residues in proximity to the nitrogen atom of MK-499 that, if protonated, might be in a favorable position to establish π−cation interactions with those near phenyl rings. However, only an in-depth study of the drug/protein docking and the direct comparison of the CoMFA and channel models might help to assess the role of the amine and perhaps of other functions in the binding of drug molecules to the HERG K^+ channel.

Conclusions

In conclusion, we present a first pharmacophore for LQTS-inducing drugs, together with a 3D QSAR model for the HERG K^+ channel blocking activity. The former was derived on the basis of some compounds taken from an organized list of QT-prolonging drugs selected following the criterion of the homogeneity of both the mechanism causing the QT prolongation (HERG K^+ channel blockade) and the in vitro assay system (transfected mammalian cells). In an attempt to extend the understanding of the physicochemical characteristics underlying the biological activity of the drugs investigated and to provide a predictive tool for the HERG K^+ channel blocking activity, we developed a CoMFA model based on the same set of compounds used for the generation of the pharmacophore and tested through an external set of molecules. The results we obtained allow one to draw a broad description of the molecular features associated with the QT-prolonging ability and eventually to assess quantitatively the HERG K^+ channel blocking potential of new compounds.

The CoMFA model presented here offers an acceptable level of predictivity but cannot yet be proposed as a rapid and efficient in silico tool for the early identification of the LQTS-inducing activity. To develop a predictive virtual screening tool, the collection of a sufficiently large database of reliable and homogeneous biological data is needed. Since the efforts that are being carried out in this direction are providing a growing number of IC_{50} values for blockade of HERG K^+ channels,[15] it is conceivable that our approach may offer a continuously improvable in silico screening method for the QT-prolonging potential. In any case, at present, our model may be used in conjunction with existing in vitro methods (none of which has an absolute predictive value) to increase overall predictivity at the early stages of drugs development.

References

(1) Viskin, S. Long QT syndromes and torsade de pointes. *Lancet* **1999**, *354*, 1625−1633.

(2) De Ponti, F.; Poluzzi, E.; Montanaro, N. QT-interval prolongation by non-cardiac drugs: lessons to be learned from recent experience. *Eur. J. Clin. Pharmacol.* **2000**, *56*, 1−18.

(3) Points to consider: the assessment of the potential for QT interval prolongation by noncardiovascular medicinal products; European Agency for the Evaluation of Medicinal Products (EMEA): Committee for Proprietary Medicinal Products (CPMP), 1997.

(4) De Ponti, F.; Poluzzi, E.; Cavalli, A.; Recanatini, M.; Montanaro, N. Safety of non-antiarrhythmic drugs that prolong the QT interval or induce torsades de pointes: an overview. *Drug Saf.* **2002**, *25*, 263−286.

(5) Nattel, S. The molecular and ionic specificity of antiarrhythmic drug actions. *J Cardiovasc. Electrophysiol.* **1999**, *10*, 272−282.

(6) Towbin, J. A.; Vatta, M. Molecular biology and the prolonged QT syndromes. *Am. J. Med.* **2001**, *110*, 385−398.

(7) Sheridan, D. J. Drug-induced proarrhythmic effects: assessment of changes in QT interval. *Br. J. Clin. Pharmacol.* **2000**, *50*, 297−302.

(8) Suessbrich, H.; Waldegger, S.; Lang, F.; Busch, A. E. Blockade of HERG channels expressed in Xenopus oocytes by the histamine receptor antagonists terfenadine and astemizole. *FEBS Lett* **1996**, *385*, 77−80.

(9) Suessbrich, H.; Schonherr, R.; Heinemann, S. H.; Attali, B.; Lang, F.; et al. The inhibitory effect of the antipsychotic drug haloperidol on HERG potassium channels expressed in Xenopus oocytes. *Br. J. Pharmacol.* **1997**, *120*, 968−974.

(10) Taglialatela, M.; Pannaccione, A.; Castaldo, P.; Giorgio, G.; Zhou, Z.; et al. Molecular basis for the lack of HERG K^+ channel block-related cardiotoxicity by the H1 receptor blocker cetirizine compared with other second-generation antihistamines. *Mol. Pharmacol.* **1998**, *54*, 113−121.

(11) Bischoff, U.; Schmidt, C.; Netzer, R.; Pongs, O. Effects of fluoroquinolones on HERG currents. *Eur. J. Pharmacol.* **2000**, *406*, 341−343.

(12) Kang, J.; Wang, L.; Chen, X. L.; Triggle, D. J.; Rampe, D. Interactions of a series of fluoroquinolone antibacterial drugs with the human cardiac K^+ channel HERG. *Mol. Pharmacol.* **2001**, *59*, 122−126.

(13) Vandenberg, J. I.; Walker, B. D.; Campbell, T. J. HERG K^+ channels: friend and foe. *Trends Pharmacol. Sci.* **2001**, *22*, 240−246.

(14) Cavero, I.; Mestre, M.; Guillon, J. M.; Crumb, W. Drugs that prolong QT interval as an unwanted effect: assessing their likelihood of inducing hazardous cardiac dysrhythmias. *Expert Opin. Pharmacother.* **2000**, *1*, 947−973.

(15) De Ponti, F.; Poluzzi, E.; Montanaro, N. Organising evidence on QT prolongation and occurrence of Torsades de Pointes with non-antiarrhythmic drugs: a call for consensus. *Eur. J. Clin. Pharmacol.* **2001**, *57*, 185−209.

(16) Fenichel, R. R. http://www.fenichel.net/QT%20stuff/pbydrug.htm.

(17) Netzer, R.; Ebneth, A.; Bischoff, U.; Pongs, O. Screening lead compounds for QT interval prolongation. *Drug Discovery Today* **2001**, *6*, 78−84.

(18) Zhang, S.; Rajamani, S.; Chen, Y.; Gong, Q.; Rong, Y.; et al. Cocaine blocks HERG, but not KvLQT1+minK, potassium channels. *Mol. Pharmacol.* **2001**, *59*, 1069−1076.

(19) Ferreira, S.; Crumb, W. J., Jr.; Carlton, C. G.; Clarkson, C. W. Effects of cocaine and its major metabolites on the HERG-encoded potassium channel. *J. Pharmacol. Exp. Ther.* **2001**, *299*, 220−226.

(20) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(21) *SYBYL*, version 6.8; Tripos, Inc.: St. Louis, MO, 2001.

(22) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R. M. J.; Lipton, M. A.; et al. MacroModel−an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **1990**, *1*, 440−467.

(23) Allen, F. H.; Kennard, O. 3D search and research using the Cambridge Structural Database. *Chem. Des. Autom. News* **1993**, *8*, 1 and 31−37.

(24) Fletcher, D. A.; McMeeking, R. F.; Parkin, D. The United Kingdom Chemical Database Service. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 746−749.

(25) Halgren, T. A. Merck Molecular Force Field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490−519.

(26) Halgren, T. A. Merck Molecular Force Field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* **1996**, *17*, 520−552.

(27) Stewart, J. P. P. Optimization of parameters for semiempirical methods I. Method. *J. Comput. Chem.* **1989**, *10*, 209−220.

(28) Chang, G.; Guida, W. C.; Still, W. C. An internal coordinate Monte Carlo method for searching conformational space. *J. Am. Chem. Soc.* **1989**, *111*, 4379−4386.

(29) Shenkin, P. S.; McDonald, D. Q. Cluster analysis of molecular conformations. *J. Comput. Chem.* **1994**, *15*, 899−916.

(30) Cavalli, A.; Greco, G.; Novellino, E.; Recanatini, M. Linking CoMFA and protein homology models of enzyme−inhibitor interactions: an application to non-steroidal aromatase inhibitors. *Bioorg. Med. Chem.* **2000**, *8*, 2771−2780.

(31) Recanatini, M.; Cavalli, A.; Belluti, F.; Piazzi, L.; Rampa, A.; et al. Structure−activity relationships of 9-amino-1,2,3,4-tetrahydroacridine-based acetylcholinesterase inhibitors: synthesis, enzyme inhibitory activity, QSAR, and structure-based CoMFA of tacrine analogues. *J. Med. Chem.* **2000**, *43*, 2007−2018.

(32) Mitcheson, J. S.; Chen, J.; Lin, M.; Culberson, C.; Sanguinetti, M. C. A structural basis for drug-induced long QT syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12329−12333.

(33) Lees-Miller, J. P.; Duan, Y.; Teng, G. Q.; Duff, H. J. Molecular determinant of high-affinity dofetilide binding to HERG1 expressed in Xenopus oocytes: involvement of S6 sites. *Mol. Pharmacol.* **2000**, *57*, 367−374.

JM0208875