



CoMFA and CoMSIA 3D-quantitative structure-activity relationship model on benzodiazepine derivatives, inhibitors of phosphodiesterase IV

Pierre Ducrot*, Charles R. Andrianjara & Roger Wrigglesworth

Pfizer Global Research and Development, Fresnes Laboratories, 3-9 Rue de la Loge, 94265 Fresnes, France

Received 18 January 2001; accepted 6 July 2001

Key words: 3D-QSAR, CoMFA, CoMSIA, H-QSAR, PDE4 inhibitors, benzodiazepines

Summary

Recently, we reported structurally novel PDE4 inhibitors based on 1,4-benzodiazepine derivatives. The main interest in developing benzodiazepine-based PDE4 inhibitors is in their lack of adverse effects of emesis with respect to rolipram-like compounds. A large effort has thus been made toward the structural optimization of this series. In the absence of structural information on the inhibitor binding mode into the PDE4 active site, 2D-QSAR (H-QSAR) and two 3D-QSAR (CoMFA and CoMSIA) methods were applied to improve our understanding of the molecular mechanism controlling the PDE4 affinity of the benzodiazepine derivatives. As expected, the CoMSIA 3D contour maps have provided more information on the benzodiazepine interaction mode with the PDE4 active site whereas CoMFA has built the best tool for activity prediction. The 2D pharmacophoric model derived from CoMSIA fields is consistent with the crystal structure of the PDE4 active site reported recently. The combination of the 2D and 3D-QSAR models was used not only to predict new compounds from the structural optimization process, but also to screen a large library of benzodiazepine derivatives.

Introduction

During the last decade, considerable attention has been focused on the cAMP-specific enzyme (PDE4) as an attractive target for novel anti-asthmatic and anti-inflammatory therapy. The relaxant potency of PDE4 inhibitors on human isolated bronchi in addition to their ability to suppress inflammatory cell activity support the concept that PDE4 inhibition represents an interesting approach toward the therapy of asthma, COPD and rheumatoid arthritis.

The first generations of selective PDE4 inhibitors were derived from the structure of rolipram but the clinical development of rolipram analogues has been hampered by the adverse effect of emesis. The characterisation of the mechanism that regulates this side effect through PDE4 inhibition has then triggered intensive investigations. The first striking finding from these studies is the presence of the high affinity

rolipram binding site (HARBS) in the structure of the PDE4 enzyme [1]. The relationships between high HARBS affinity and emesis potential have been first developed on the basis of the effects of the rolipram analogues on the acid secretion in rabbit gastric gland which is considered to be one of the biological components that induces the emesis effect.

On the other hand, some recent results tend to suggest that the emesis potency is related to the rolipram structure but not to the HARBS affinity. This hypothesis has been built on the pharmacological profile of some PDE4 inhibitors which induce a weak emesis effect despite their high affinities to HARBS [2]. The series of PDE4 inhibitors that display these characteristics are mainly composed of compounds structurally different from rolipram. These observations indicate a lack of data and knowledge on the relationship between PDE4 inhibitor structure, emesis potential and the HARBS affinity. No rational basis has been developed to reduce the adverse effect of emesis in the PDE4 inhibitors so far. Therefore, the approach that has been adopted to enhance the chance of success in

*To whom correspondence should be addressed.
E-mail: Pierre.ducrot@pfizer.com

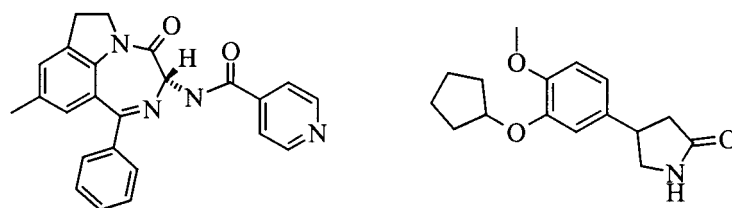


Figure 1. CI-1018 (or PD0168787, left) and Rolipram (right).

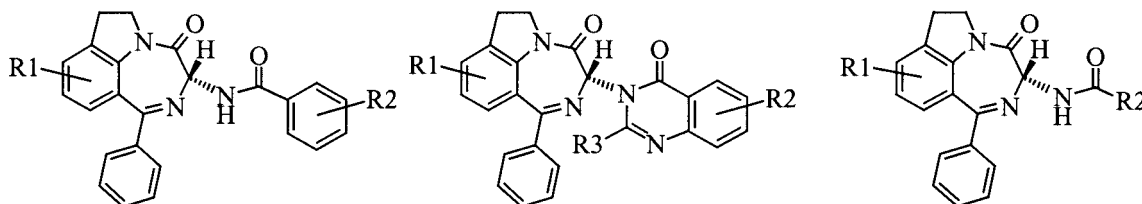


Figure 2. Three benzodiazepine series, A, B and C from left to right respectively.

the clinical development of PDE4 inhibitors consists of optimizing the structural diversity of each series as routinely used in the field of drug design.

A large variety of molecular structures have been reported to competitively inhibit the PDE4 enzyme with respect to cAMP. The most studied PDE4 inhibitor structures can be classified into three families: rolipram-related compounds, xanthine derivatives and nitraquazone analogues. This structural diversity of the PDE4 inhibitors suggests that the cAMP binding site presents a number of pharmacophores which are capable of interacting with different structural types of inhibitors.

Although the three dimensional structure of the PDE4 enzyme in the solid state has been recently published [3], no information on the inhibitor binding mode is available yet. Therefore, the mainstream approach to understand the structural basis for the affinity to the catalytic center is to focus on the search for the common pharmacophores of PDE4 inhibitors. Different pharmacophore models of the PDE4 inhibitors have been proposed in the literature [4]. These pharmacophores have been mainly obtained from the conformational analysis and the structural superimposition of the three families of PDE4 inhibitors previously mentioned, and led to a pharmacophoric model displaying up to two lipophilic centers and three ligand hydrogen bond acceptors.

Recently, we reported structurally novel PDE4 inhibitors based on the benzodiazepine derivatives (Figure 1) which both exhibit a strong inhibition of eosinophilic infiltration in sensitized Brown-Norway

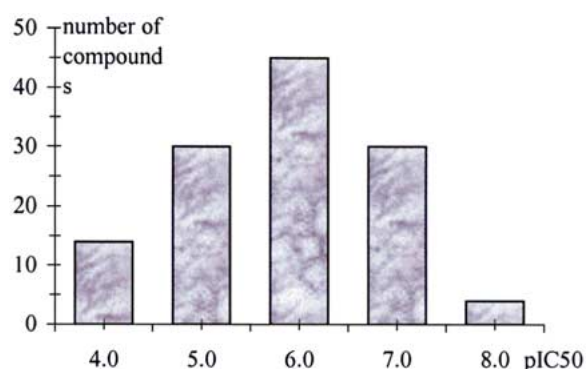


Figure 3. Range of PDE-4 activities.

rats (PD0168787) and are devoid of the adverse effect of emesis [5].

On the basis of the assumption that different series of PDE4 inhibitors could bind into the catalytic center in different fashion, which is also emphasized by the structural discrepancies of the benzodiazepine series with respect to rolipram, xanthine and nitraquazone series, a search for new additional pharmacophores is required in order to improve the accuracy of the binding model of the PDE4 inhibitors.

In the attempt to reach this goal, a quantitative SAR study of a set of 125 benzodiazepine derivatives has been carried out using three different techniques available from SYBYL package [6]. At first, a 2D-QSAR analysis based on the molecular hologram (H-QSAR) techniques [7] has been performed to obtain a preliminary QSAR profile of the benzodiazepine series which provides a perspective on the contribution of the different structural substituents to the com-

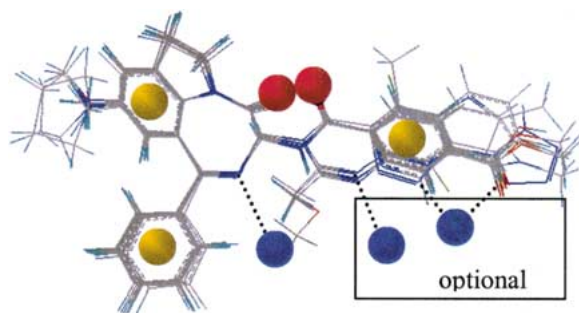


Figure 4. Pharmacophoric alignment.

pound activity. This approach, based on structural fingerprinting, was recently combined with different techniques for statistical analyses such as Multiple Linear Regression (MLR), Principal Component Regression (PCR), Partial Least Square (PLS) and Artificial Neural Network (ANN) to successfully derive QSAR models of the benzodiazepine derivatives for the γ -aminobutyric acid receptor ($GABA_A$) [8]. The main benefit of H-QSAR comes from the speed of the molecular hologram generation process which allows a linear structural representation of compounds in the combinatorial library to be obtained rapidly. Therefore, the derived H-QSAR model could be used to screen a large library of benzodiazepine derivatives. As we routinely used the procedure of parallel synthesis in the field of lead compound optimization, the result of the 2D-QSAR analyses would be an asset to rationalizing and speeding up the screening of new benzodiazepine-based PDE4 inhibitors. Despite the H-QSAR efficiency in the design of combinatorial and parallel libraries, the activity prediction lacks accuracy when structures of the new compounds are greatly different from the training set used to build the 2D-QSAR model (Tanimoto coefficient < 0.85) [9].

On the contrary, 3D-QSAR allows fine predictions for structurally diverse compounds if structure alterations occur within the model's explored volume. However, it requires a time consuming alignment phase which prevent screening of large databases. For this reason, we developed a 2D/3D combined methodology in three steps for the screening of virtual libraries with the aim of obtaining a rough estimation of the contribution of the different molecular fragments of the compounds in the training set to the activity, and providing both a first guide for the structural alignment exercise in the 3D-QSAR analysis and a predictive tool to rapidly screen a large dataset of benzodiazepine derivatives with unknown PDE4 activity. The three

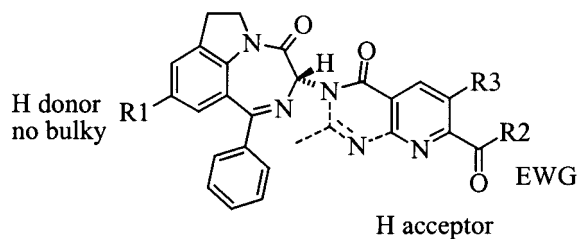
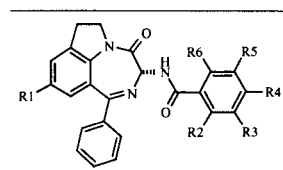




Figure 5. Pharmacophores derived from H-QSAR.

steps are: (1) similarity search: compounds from the virtual library are compared to the H-QSAR training set compounds using the Tanimoto coefficient based on UNITY fingerprints. Hence, compounds having a similarity coefficient higher than 0.85 with at least one compound of the training set are eligible for H-QSAR activity prediction; (2) H-QSAR activity prediction: compounds are predicted for PDE4 IC_{50} using a H-QSAR model, the most potent putative inhibitors (predicted $IC_{50} < 1\mu M$) are eligible for 3D-alignment and 3D-QSAR activity prediction; (3) 3D-QSAR activity prediction: putative inhibitors are aligned using the predefined rules and are predicted for activity using the 3D-QSAR model. In this context, 3D-QSAR analyses have been performed using both the comparative molecular field analysis (CoMFA) [10] and comparative molecular similarity indices analysis (CoMSIA) [11] methods. These two techniques have been shown to be efficient and are widely used to derive 3D-QSAR models of benzodiazepine derivatives on several molecular targets [12], but no result has been reported for the PDE4 enzyme so far. CoMFA, which is the most popular 3D-QSAR method has been chosen because of the renowned robustness of the models it produces. The CoMSIA approach mainly differs from the previous method by the way the molecular fields are calculated and by including an additional lipophilic field. The purpose of using supplementary fields, such as lipophilic and hydrogen bonding potential is not only to increase the significance and predictive power of the 3D-QSAR models, but also to have a better visualisation and interpretation of the obtained correlation in terms of field contribution to the compound activity. As a matter of fact, CoMFA fields such as the electrostatic, might be difficult to interpret if the correlated areas result from several molecular alterations.

The combination of these three techniques of QSAR analysis to the novel series of benzodiazepines not only improves our understanding of the structural basis for their affinity to the catalytic center of the

Table 1. Series A



	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
PD0190172	* 	H	Cl	NH ₂	Cl	H
PD0190171	* NHCOCH ₃	H	Cl	NH ₂	Cl	H
PD0190168	* N(CH ₃) ₂	H	Cl	NH ₂	Cl	H
PD0186505	* NH ₂	H	H	NHCO ₂ -t-Bu	H	H
PD0190830	* CH ₃	H	NO ₂	CH ₃	NO ₂	H
PD0186506	* NH ₂	H	H	NH ₂	H	H
PD0189794	* NH ₂	NHCOCH ₃	H	H	H	Cl
PD0200868	* CH ₃	OCH ₃	H	OCH ₃	H	H
PD0190821	* CH ₃	OH	H	H	H	H
PD0190847	* CH ₃	H	H	OCH ₂ COCH ₃	H	H
PD0190822	* CH ₃	H	H	SO ₂ NHCOCH ₃	H	H
PD0190844	* CH ₃	H	H		H	H
PD0190184	CH ₃	H	Cl	NH ₂	Cl	H
PD0190826	CH ₃	NO ₂	H	OCH ₃	H	H
PD0190849	CH ₃	H	H	CH ₂ NHCO(CH ₂) ₂ CH ₃	H	H
PD0190827	CH ₃	H	H	SO ₂ CH ₃	H	H
PD0190597	CH ₃	NHCONHCH ₂ CH ₃	H	H	H	H

(*) training set.

PDE4 enzyme but also provides a new approach in the design of PDE4 inhibitors.

Materials and methods

Design of the training sets

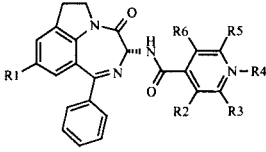
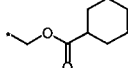
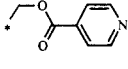
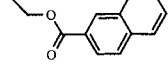
The reliability of the results from the QSAR analyses is sensitive to the structural composition and biological data distribution of the training set. As a matter of fact the training set must be designed such that a maximum of information can be obtained with the minimum number of compounds in order to speed up the analysis process and improve the predictiveness of the model. In addition, the molecules that constitute the testing set may be structurally different but need to retain some similarity in order to obtain reliable QSAR results. The structural similarities in the training set are needed to satisfy the assumption that compounds subjected to QSAR analysis bind into the catalytic center of PDE4 enzyme in a similar fashion. Compounds were selected from three major series of

benzodiazepines (Figure 2). The first series is composed of the benzodiazepine ring substituted by aryl amide groups at position 3, the second type displays a conformationally restricted side chain, whereas the third have aliphatic amide groups.

The chemical synthesis procedures as well as the SAR description for the second series have been described elsewhere [13]. From our molecular database, 118 compounds from these three benzodiazepines series have been registered with quantitative PDE4 inhibition activity pIC_{50} ($pIC_{50} = -\log[IC_{50}]$) values over 4.0. The compounds displaying a low potency ($pIC_{50} < 4.0$) were then discarded for the QSAR analyses due to the lack of insight on the pIC_{50} values. Figure 3 illustrates the partition of the structures for different ranges of the PDE4 inhibition activities. Noteworthy is the fact that the biological data values in the dataset spreads from $pIC_{50} = 4$ to 9 and each PDE4 inhibition activity range contain at least 4 compounds.

In order to enhance the structural diversity of the training set and reduce the redundancy owing to

Table 2. Series B

		R₁	R₂	R₃	R₄	R₅	R₆
PD0185586	*	CHO	H	H	-	H	H
PD0189658	*	NH ₂	NH ₂	H	-	H	H
PD0188990	*	CO ₂ CH ₂ CH ₂ OH	H	H	-	H	H
PD0186488	*	CH ₃	NHCOCH ₃	H		H	H
PD0188988	*	CH ₂ OCOPh	H	H	-	H	H
PD0188987	*		H	H	-	H	H
PD0188989	*		H	H	-	H	H
PD0189117	*		H	H	-	H	H
PD0188991		CH ₂ OCOCH ₃	H	H	-	H	H
PD0186706		CH ₃	H	H	CH ₃	H	H

(*) training set.

structurally close compounds, a hierarchical clustering analysis was performed on the 118 bendodiazepine derivatives using both UNITY default fingerprint and atompair descriptors.

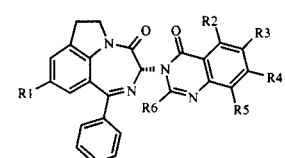
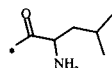
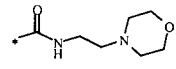
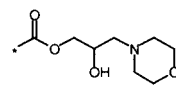
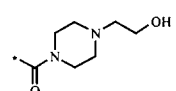
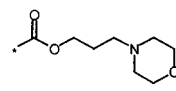
This study provided 19 structural clusters which were defined depending upon the nature of the substituents, regardless of their activity. By randomly selecting compounds within the 19 clusters, a training set containing 74 optimally diverse structures has been obtained. The remaining 44 compounds with high variance of biological data and structural heterogeneity represent the testing set which will be used to assess the quality of the training set and the predictive power of the QSAR model. The structural diversity and the homogeneous repartition of the PDE4 affinities which are present in our training and testing sets of compounds ensure the obtention of meaningful results from either 2D or 3D-QSAR analysis.

2D-QSAR analysis

The molecular hologram techniques have been chosen to obtain a linear representation of the chemical structures in the dataset to speed up the process. This procedure consists of hashing molecular structures in different fragments and storing the structural informa-

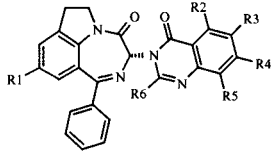
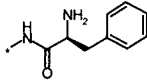

tion into bins. The molecular hologram differs from fingerprint methodologies on the definition of the bins. Rather than using a binary bit string containing either 0 or 1, the molecular hologram retains a count of the number of times each bin is set. Sybyl provides several types of fragment (based on atoms, bonds, connections, hydrogens, H-bond donors and acceptors), as well as parameters to define the smallest and largest size of the fragments, and the length of the hologram for each compound. Statistical analysis of these molecular hologram parameters using the Partial Least Square (PLS) method generates a number of so-called H-QSAR (Hologram-QSAR) models. One of the major issues in the H-QSAR analysis is that the predictive power of the derived model might be biased by the hologram length [12b]. To overcome this issue, several models were generated by using different hologram lengths. This collection of models is referred to as a model ensemble. Statistical analysis of the collection was performed in order to compute the ensemble q^2 (average value of the cross-validated r^2 using the 'leave-one-out' method, Equation 1) and the associated ensemble standard error (SD_e) given by Equation 2, where Nr, Nc and Nm are the number of rows, average number of components and number of models respectively.

Table 3. Series C

						
	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
PD0184268	* CH ₃	CH ₂ OH	H	H	H	CH ₃
PD0184440	* OCH ₃	H	H	CO ₂ H	H	CH ₃
PD0184441	* OCH ₃	H	H	H	H	CH ₃
PD0185583	* NH ₂	H	H	CO ₂ H	H	CH ₃
PD0185584	* NH ₂	H	H	CO ₂ CH ₃	H	CH ₃
PD0185585	* CH ₃	H	Br	H	H	CH ₃
PD0186611	* NH ₂	H	H	H	H	CH ₃
PD0186620	* CH ₃	H	H	H	H	CH ₃
PD0186622		H	H	H	H	CH ₃
PD0186628	* NH ₂	H	CH ₃	H	H	CH ₃
PD0186646	* CH ₃	Cl	H	H	H	CH ₃
PD0186647	* CH ₃	F	H	H	H	CH ₃
PD0186719	* CH ₃	H	H	CH ₂ OH	H	CH ₃
PD0186720	* CH ₃	H	H	CONHCH ₃	H	CH ₃
PD0186721	* CH ₃	H	H	CONH ₂	H	CH ₃
PD0186725	CH ₃	H	H	CO ₂ (CH ₂) ₂ NH(CH ₂) ₂ OH	H	CH ₃
PD0186729	CH ₃	H	H	CONH(CH ₂) ₂ NH ₂	H	CH ₃
PD0186630	* NH ₂	H	CH ₃	H	CH ₃	CH ₃
PD0186731	CH ₃	H	H	CON((CH ₂) ₂ OH) ₂	H	CH ₃
PD0186732	CH ₃	H	H		H	CH ₃
PD0186733	CH ₃	H	H	CO ₂ (CH ₂) ₂ OCH ₃	H	CH ₃
PD0186734	CH ₃	H	H		H	CH ₃
PD0186735	CH ₃	H	H	CO ₂ (CH ₂) ₂ OCOCH ₃	H	CH ₃
PD0186736	CH ₃	H	H		H	CH ₃
PD0186737	CH ₃	H	H		H	CH ₃
PD0186741	* CH ₃	H	H	CON(CH ₃) ₂	H	CH ₃
PD0186744	CH ₃	H	H	CH ₂ OCOCH ₂ OCOCH ₃	H	CH ₃
PD0186745	CH ₃	H	H	CO ₂ CH ₂ CH(OH)CH ₂ OH	H	CH ₃
PD0186747	CH ₃	H	OCH ₃	OCH ₃	H	CH ₃
PD0186754	CH ₃	H	H	OCH ₃	H	CH ₃
PD0189574	* NH ₂	Cl	H	H	H	CH ₃

(*) training set.

Table 3. Continued

		R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
PD0190187			H	H	H	H	CH ₃
PD0194021	* CH ₃		H	H		H	CH ₃
PD0190271	* CH ₃		H	H	H	H	CH ₂ OCH ₃
PD0186755	CH ₃		H	H	CO ₂ H	H	CH ₃
PD0190275	* OCH ₃		H	H	CO ₂ CH ₃	H	CH ₃
PD0186724	OCH ₃		Cl	H	H	H	CO ₂ CH ₂ CH ₃
PD0186718	CH ₃		H	H	H	H	CH ₂ CO ₂ CH ₂ CH ₃

(*) training set.

$$\text{Ens} - q^2 = (\text{Nr} - \text{Nc} - 1)(\text{SD}_e)^2 / \Sigma(Y_{\text{act}} - Y_{\text{mean}})^2, \quad (1)$$

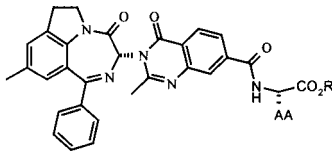
$$\text{SD}_e = \Sigma(\text{SD}) / \text{Nm}. \quad (2)$$

In addition, the selection of one H-QSAR model from the ensemble is based on the least standard error (SD) values corresponding to the optimal length of the hologram. The predictive power of the model is then evaluated by the activity prediction of the testing test.

3D-QSAR analysis

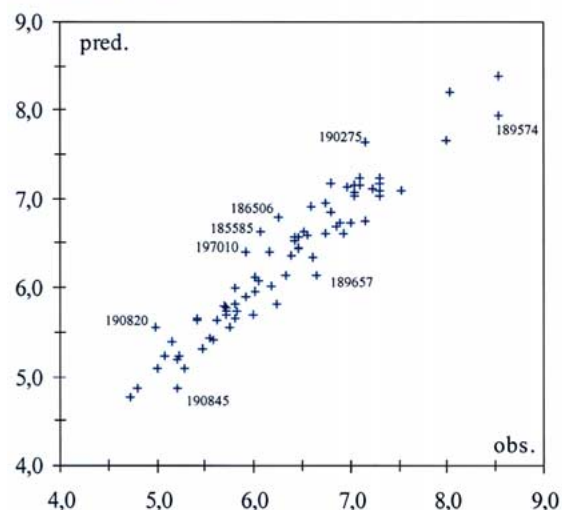
An essential requirement for 3D-QSAR techniques using molecular field analysis is the knowledge of the active conformation of the compound in the data set. In the absence of the structural information on the inhibitor binding mode into the catalytic center of the PDE4 enzyme, conformational analysis of the benzodiazepine derivatives was performed in order to compare compound's potent pharmacophore layout and identify putative active conformations.

Table 4. Serie D

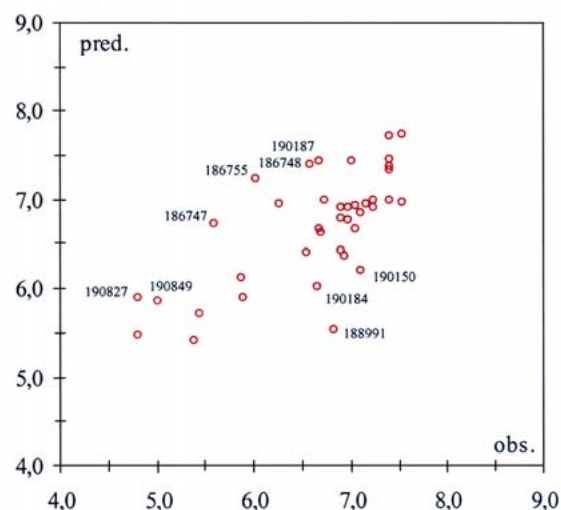
		Amino acid (AA)	R
PD0184270		Glu	H
PD0184271		Pro	H
PD0184272		Val	H
PD0184273		Tyr	H
PD0184274		Ala	H
PD0184275		Phe	H
PD0184276		Gly	H
PD0184277		Leu	H
PD0184278		Asp	H
PD0184279		Lys	H
PD0186748		Phe	CH ₃
PD0189747		Lys	CH ₃

(*) training set.

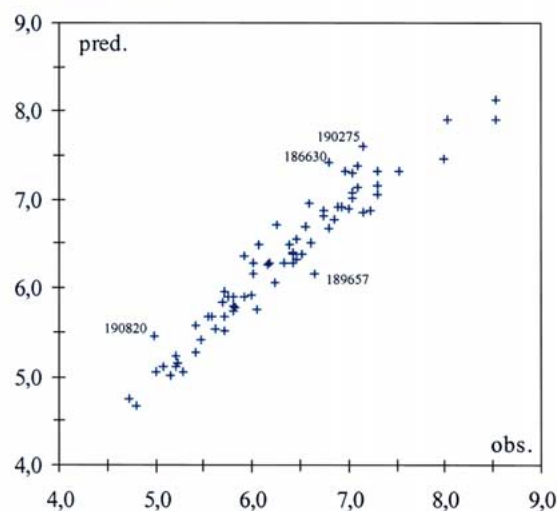
(a) training set



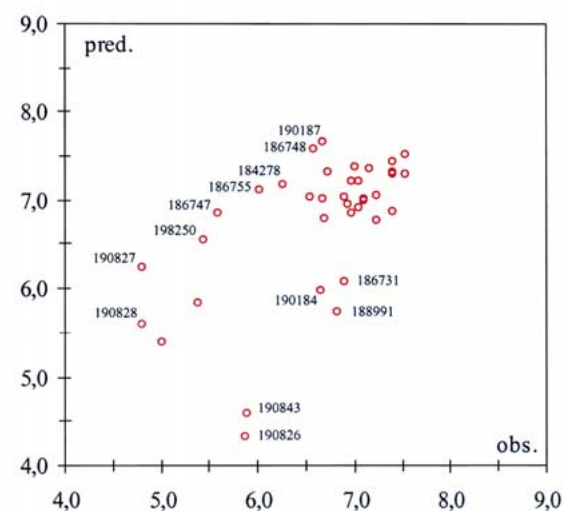
(b) testing sets.

Figure 6. CoMFA: predicted vs. actual pIC₅₀.

(a) training set



(b) testing sets.

Figure 7. CoMSIA: predicted vs. actual pIC₅₀.

Conformational analysis

The puckering of the seven membered ring of the benzodiazepine defines the orientation of the side chain (3-substituent). Different investigations suggest that endo (pseudoaxial) and exo (pseudoequatorial) side chains are energetically favorable to the same degree. However, based on several benzodiazepine X-ray structures [12d], a consensus has been found that the 3-substituent occupies a pseudo-equatorial position (exo). This is consistent with the X-ray structure of the

(*R*)-active enantiomer of the compound PD0186787 in which the seven membered ring adopts a rough boat conformation with a pseudoequatorial (exo) side chain (3-substituent). Moreover, the pyrrolidine ring tends to flatten the benzodiazepine ring system with respect to the usual boat conformation described in the literature [14].

The 118 benzodiazepine derivatives in the dataset were converted to 3D structures using Concord [15]. Compound geometry was then optimized using the MMFF94s force field [16], including electrostatic

terms (MMFF94 partial charges). The method of Powell, available in the Maximin2 procedure was used for energy minimization until the gradient value was smaller than 0.01 kcal/mol Å². A conformational search was then performed using the randomsearch procedure within Sybyl with a maximum hits per conformer set to 10, a r.m.s. threshold of 0.2 Å and an energy cutoff of 8.0 kcal/mol. This protocol was applied to generate different types of low energy conformers which were used to refine the structural alignment for the 3D-QSAR analysis.

Structural alignment procedures

In order to address the question of the active conformation and bypass the lack of atomic coordinates for the PDE4 enzyme in complex with an inhibitor, the ten most structurally diverse derivatives among the most active compounds (pIC₅₀ > 7.0) were first compared using DISCO. DISCO describes compounds by adding pharmacophoric features and measuring distances between them. Comparison of each conformer of each active compound provides insights into common pharmacophores and relative alignments to define common pharmacophores, as well as generating several hypotheses for the active conformation and relative alignment rules. As the selected compounds have several obvious common pharmacophores on the benzodiazepine skeleton and are rather similar, a cutoff of at least six pharmacophoric features of any type was used in order to reduce the number of potent alignments and to thus privilege those displaying the most pharmacophoric features. Eight pharmacophoric points were thus identified (Figure 4), including the three aromatic centroids (yellow), two aromatic normals (not shown for clarity), three hydrogen bond donor sites (blue, including two optional) and two hydrogen bond acceptor atoms (red).

Two active conformational hypotheses, depending on the torsion angle between C-3 and the side-chain nitrogen were depicted for each compound from the DISCO study. The conformers with low H-C3-N-CO torsion angle ($|\theta| < 20^\circ$) are of lower energy than those having large torsion angle ($|\theta| > 160^\circ$). Furthermore, the former conformation corresponds to the X-ray structure of compound PD0186787, and was therefore defined as the active conformation. The 19 cluster lead compounds (cluster's most active compound) were then overlaid to the DISCO model using the appropriate pharmacophores.

It is noteworthy that the pharmacophores highlighted by DISCO are very similar to the maximum common structure. Moreover, the most active compounds are very similar too, therefore the defined pharmacophores are not sufficient enough to align the least active compound side chains. For this reason, a complementary method has been used to align the whole database. Within each derived cluster, each conformer of each compound was overlaid onto its cluster lead compound (cluster's most active compound) using the maximum common structure (MCS, benzodiazepine core). The conformer selection was then carried out by optimizing putative binding interactions within a given cluster by scoring each conformer using its steric and electrostatic CoMFA fields with the Tanimoto coefficient against a reference compound (cluster lead). If several conformations led to the same score, the lowest energy conformer was selected. This automated proprietary procedure optimizes the alignment of similar compounds using the molecular fields, taking into account low energy conformers only. Thus, similar side chains have similar orientations in the final alignment. The geometry of each selected conformer was then optimized using the hamiltonian AM1 (MOPAC 6.0). The Pulay convergence method combined with the MMOK (molecular mechanics correction for amide groups) and EF (Eigenvector Following routine for minimization) [17] keywords were applied in assigning charges as well.

CoMFA and CoMSIA parameters

The CoMFA Steric, Electrostatic, Indicator and Hydrogen bonding molecular fields available within Sybyl were calculated at grid lattice points with the Lennard-Jones and Coulomb potential functions of the Tripos Force Field using a common sp³ carbon probe of charge +1, with a 2 Å grid spacing. The energy cutoff was set to 30 kcal/mol for standard and hydrogen bond steric and electrostatic fields, whereas it was set to 20 and 1 kcal/mol for the indicator steric and electrostatic fields respectively. Those parameters were determined as optimum after several alterations (results not shown). The five CoMSIA similarity indices fields available within Sybyl (steric, electrostatic, lipophilic, and hydrogen bond donor and acceptor) were calculated at grid lattice points using a common probe atom of 1 Å radius, as well as charge, hydrophobicity, and hydrogen bond properties of +1, and an attenuation factor of 0.3 while altering grid spacing from 1 to 2 Å with a 0.5 Å step.

Table 5. Serie D

	R₁	R₂	
PD0186658	*		
PD0190194	* NH ₂		
PD0194016	* NH ₂		
PD0189657	* CH ₃		
PD0189660	* NH ₂		
PD0190823	* CH ₃		
PD0190836	* CH ₃		
PD0190833	* CH ₃		
PD0190842	* CH ₃		
PD0200310	* OCH ₃		
PD0190846	* CH ₃		
PD0190831	* CH ₃		
PD0190832	* CH ₃		
PD0190835	* CH ₃		
PD0190839	* CH ₃		

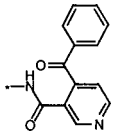
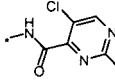
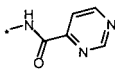
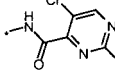
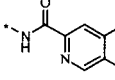
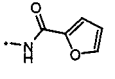
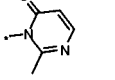
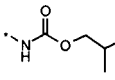
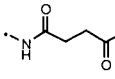
(*) training set.

Table 5. Continued

PD0190825	* CH ₃	
PD0190850	* CH ₃	
PD0190824	* CH ₃	
PD0190852	* CH ₃	
PD0190838	* CH ₃	
PD0190845	* CH ₃	
PD0190848	* CH ₃	
PD0190150	NH ₂	
PD0190843	CH ₃	
PD0198250	NH ₂	
PD0190828	CH ₃	
PD0190837	CH ₃	
PD0194064	* NH ₂	
PD0199522	* NH ₂	
PD0199521	* NH ₂	
PD0189659	* NH ₂	
PD0190840	CH ₃	

(*) training set.

Table 5. Continued

PD0190854	* CH ₃	
PD0190853	* CH ₃	
PD0186743	* CH ₃	
PD0190834	* CH ₃	
PD0186470	* NH ₂	
PD0190820	* CH ₃	
PD0197010	* CH ₃	
PD0186476	* NH ₂	
PD0186750	CH ₃	

(*) training set.

As in the H-QSAR analysis, the PLS method (partial least square) was used to derive a linear relationship, and cross-validation was performed using the 'leave-one-out' method to check for consistency and predictiveness. The optimal number of components used to derive the non-validated model was defined as the number of components leading to the highest cross-validation r^2 (subsequently call q^2) and the lowest standard error of prediction (SEP), usually corresponding to the highest F ratio. Only models with q^2 over 0.6, and an explained fraction of variance (V) greater than 0.85 for the optimum number of components were further considered. The predictiveness (predictive r^2) is determined using Equation 3, where Y_{mean} is the mean of observed pIC_{50} over the combined training and testing sets.

$$\text{pr}^2 = 1 - \Sigma(Y_{\text{act}} - Y_{\text{mean}})^2 /$$

$$\Sigma(Y_{\text{act}} - Y_{\text{pred}})^2. \quad (3)$$

Results and discussion

2D-QSAR analyses

The combination of the molecular hologram techniques with PLS analyses on the training set of 74 compounds has generated a H-QSAR model with a cross-validation r^2 (q^2) = 0.733 corresponding to $\text{SD}_{\text{cv}} = 0.447$ for an optimum number of components of $n = 3$. The highest q^2 value was obtained at the hologram length of $L = 353$ bits. The rather large size of the hologram string shows that although the training set is composed of benzodiazepine derivatives only, it is still diverse. The structural diversity of the data set is mainly located at the different substituents on the benzodiazepine core structure. These structural features of our training set are among the factors that lead to a high value of q^2 for the cross-validation procedures. In addition to this high value of q^2 , the relatively small number of components ($n = 3$) ensures a reliable non cross-validated H-QSAR model. The derived model presents a r^2 of 0.867 with a standard deviation of $\text{SD} = 0.316$ and an ensemble- $q^2 = 0.691$ ($\text{SD}_e = 0.480$). These parameters enable an assessment of the model quality. On the basis of these QSAR parameters, this model may be used to predict activities with high confidence, as confirmed by the predictive r^2 ($\text{pr}^2 = 0.480$; $\text{SD}_{\text{pred}} = 0.977$) calculated by use of the 44 compounds with known PDE4 affinities defining the testing set. Indeed, with the exception of a few compounds, basically those having functional groups not present in the model (e.g. PD0188991), the H-QSAR model led to good predictions.

Beside the good H-QSAR statistical results, the contribution of the different groups to the activity is of particular interest in order to validate the model by comparison with known SAR and design new active compounds. Regarding the indoline moiety, NH_2 , NHAc and $\text{N}(\text{CH}_3)_2$ tend to increase the potency, whereas methyl, or ethers ($-\text{CH}_2\text{OR}$) lead to decreasing potency. Regarding the linker between the side chain and the seven membered ring, pyridinones (series B) are preferred to amides (series A and C). Even small substituents (methyl, chlorine) on R_3 are detrimental to the activity, whereas 3-pyridyl substituent or the presence of acids, amides, or esters on R_2 , tend to increase the potency. In summary, small H-bond donor groups at R_1 , small substituents at R_2 , and

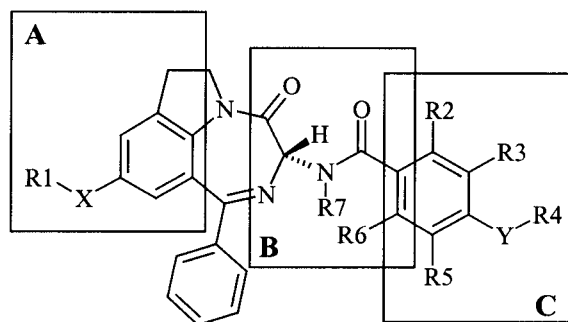


Figure 8. R-group and region definitions.

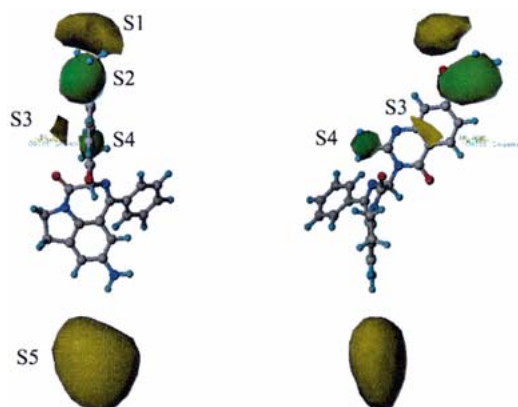


Figure 9. CoMSIA steric 3D-contour map.

the $N=C(CH_3)N$ moiety of compounds of the second series are favourable, whereas any substituent at R_3 is detrimental to affinity (Tables 1–5). Similarly, H-bond acceptor groups (3-pyridyl or 4-carbonyl) on the side-chain enhance the activity (Figure 5).

The present H-QSAR model was used to assess the PDE4 affinity of the 764 benzodiazepines present in the corporate database and previously described as CCK inhibitors, with an average similarity to compounds within the training set of $S_{\text{mean}} = 0.89$ (SD = 0.11). From this virtual screening, 47% of these compounds were found inactive, 48% were predicted to have a micromolar range activity, whereas 5% of the compounds were predicted to be active (100–500 nM). These results are clearly in accord with mass screening data. The compounds which have been predicted to be potent PDE4 inhibitors are among the most active of the 118 compounds composing the data sets of this study.

CoMFA

The obtained CoMFA results were more accurate for a sp^3 carbon probe carrying a charge of +1 for a

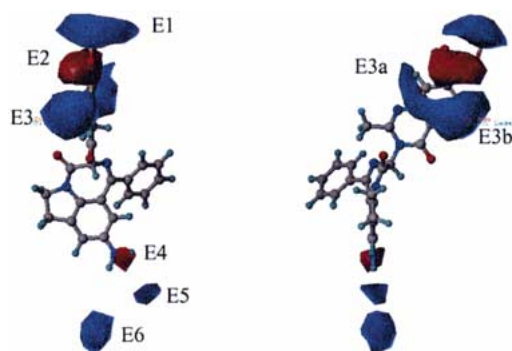


Figure 10. CoMSIA electrostatic 3D-contour map.

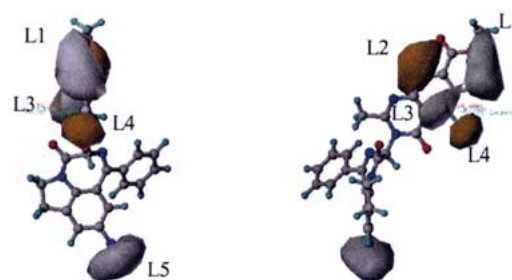


Figure 11. CoMSIA hydrophobic 3D-contour map.

grid spacing of 2 Å, whereas a 1 Å grid spacing did not lead to improvements, but dramatically increased the computation time. A 2 kcal/mol column filtering was applied during the cross-validation process. Indeed, it was found to be a good compromise, optimizing the q^2 /computation time ratio, as compared to 1 kcal/mol or no cutoff (results not shown). As it is often discussed whether the statistical significance of comparative molecular field analysis increases by considering additional property fields, different CoMFA studies using several combinations of different fields (standard steric and electrostatic, H-bond donor and acceptor, and indicator) were performed.

The results summarized in Table 6 clearly indicate that standard CoMFA fields (steric and electrostatic, entry 1) provide the highest q^2 (0.703) and a very good r^2 (0.917) as well as the fraction of the explained variance ($V = 0.916$). Both hydrogen bond (entry 2) and indicator fields (entry 3) led to lower results based on these statistical data. Further attempts to combine the latter two fields with the standard steric and electrostatic ones did not lead to any improvements. However, by combining the standard CoMFA and indicator fields (entry 6), the r^2 value was improved up to 0.945 whilst the cross-validation q^2 has slightly decreased down to 0.692 in comparison to the

Table 6. CoMFA results

#	Field	'Leave-one-out' crossvalidation						Predic. r^2	Influence of different fields (%)			
		q^2	n	SD_{cv}	V	r^2	SD		S	E	S	E
1	Standard	0.703	5	0.473	0.916	0.917	0.250	0.649	56.1	43.9		
2	H-bond	0.620	2	0.524	0.720	0.713	0.455	–	46.9	53.1		
3	Indicator	0.670	3	0.492	0.821	0.875	0.302	–	49.5	50.5		
4	Combined 1 + 2	0.658	4	0.504	0.903	0.885	0.293	0.548	25.9	16.0	28.9	29.2
5	Combined 2 + 3	0.645	4	0.519	0.986	0.908	0.267	0.597	29.7	24.4	22.5	23.4
6	Combined 1 + 3	0.692	5	0.481	0.918	0.945	0.204	0.636	20.4	16.0	31.5	32.1

SEP = standard error of prediction, n = number of components, V = fraction of variance, F = F ratio, S = steric, E = electrostatic.

standard CoMFA field alone. The latter r^2 value is probably over estimated as a consequence of the great number of redundancy in the variables from the combination of fields. In order to address this question and assess their predictiveness, each CoMFA model with both q^2 and V (fraction of explained variance) higher than 0.6 and 0.85 respectively were used to predict the activity of compounds in the testing set. Four models (entries 1, 4, 5 and 6) have fulfilled these criteria and their predictiveness (pr^2) were determined using Equation 2. On the basis of the predictive r^2 values (pr^2) in Table 6, the combination of the standard CoMFA with any other fields does not improve the activity prediction of the testing set. Therefore, the CoMFA model which is referred to as standard (steric and electrostatic) fields model was chosen to be used for this study.

Direct plots of CoMFA predicted vs. observed activities using both training and testing sets are shown on Figures 6a and 6b, respectively. As expected from the r^2 (0.917) and the standard deviation values ($SD = 0.250$), a satisfactory linear correlation has been obtained between the observed and computed affinities of the compounds in the training set. It is noteworthy that the predictive r^2 ($pr^2 = 0.649$) value obtained for the testing set on (Figure 6B) is lower than the cross-validated r^2 ($q^2 = 0.703$).

This is the result of the 'random' selection of the testing set compounds among fairly diverse compounds from each 19 clusters. Indeed, after such a selection the testing set may contain compounds displaying features in unexplored regions of the space. However this strategy has been adopted in the attempt to obtain a basic idea on the structural types of compounds for which the derived CoMFA model could be used to predict the PDE4 affinity with an acceptable accuracy.

By analyzing the regression plot of the testing set on Figure 6b, 8 out of 44 compounds appear to be outliers. Of these outliers, 6 compounds are racemic and their PDE4 affinities (pIC_{50}) are overpredicted. This is consistent with the SAR analyses which indicate that the active enantiomers have shown to present higher PDE4 affinities than the racemic compounds in the benzodiazepine series. Therefore, the reliability of the derived CoMFA model would not be affected by the presence of these 6 outliers. However, one racemic compound (PD0188991) of the two remaining outliers is surprisingly under-predicted. By examining the SAR on Table 2 (Serie B), the substitution by a bulky group at the position R_1 reduces the PDE4 affinity to less than $pIC_{50} = 5.5$ except for this compound (PD0188991) for which the pIC_{50} increases to 6.82. This compound could be assumed as a singleton in terms of biological affinity with respect to the compound cluster bearing a bulky group at R_1 and having a pIC_{50} less than 5.5. This could explain why the derived CoMFA model was not able to predict the activity of PD0188991 with accuracy. Similarly, the last outlier PD0190150 turns out to be also a singleton since its biological affinity is higher than the compounds belonging to the same series. On the basis of these analyses, only two out of 44 compounds could be considered as outliers which suggest that the derived CoMFA model remains a reliable tool for predicting the PDE4 affinity of new structures in the benzodiazepine series.

CoMSIA

The CoMSIA results were obtained from using the same structural alignment and the same training and testing sets as defined in the CoMFA analyses. In contrast to the previous CoMFA studies, the 1 Å grid spacing was found to provide the optimum results with

Table 7. CoMSIA results with 1 kcal/mol column filtering

#	Field	'Leave-one-out' crossvalidation						Predic. r^2	Influence of different fields (%)			
		q^2	n	SD _{cv}	V	r^2	SD		S	E	L	H
1	Steric/electrostatic	0.588	5	0.557	0.852	0.844	0.343	0.454	20.8	79.2		
2	Lipophilic	0.483	8	0.647	0.858	0.855	0.342	-			100.0	
3	H-bond acc./don.	0.594	2	0.541	0.711	0.682	0.479	-				100.0
4	Combined 1+2	0.581	3	0.554	0.832	0.804	0.379	-	9.2	28.3		62.4
5	Combined 1+3	0.663	5	0.504	0.909	0.874	0.308	0.384	13.9	43.0	43.1	
6	Combined 1+2+3	0.641	7	0.534	0.940	0.923	0.247	0.352	8.0	24.1	21.5	45.3

SEP = standard error of prediction, n = number of components, V = fraction of variance; F = F ratio; S = steric; E = electrostatic; L = lipophilic; H = donor and acceptor H-bond fields.

respect to the 2 Å grid resolution. It is noteworthy that using a 1 kcal/mol columns filtering cutoff vs. 2 kcal/mol slightly improved the CoMSIA q^2 , whereas no improvement was observed in the CoMFA analyses while decreasing the cutoff (results not shown). The effect of different combinations of five fields on the CoMSIA results were also investigated. The results of these studies are summarized on Table 7.

A comparative study of the data on Table 6 and Table 7 displays that the cross-validation r^2 (q^2) from CoMSIA are lower than those of CoMFA for the same type of field combinations. This can be interpreted as a lower ability of CoMSIA in data extrapolation and is even more evident in the predictiveness scores (*vide supra*). Interestingly, the highest q^2 has been obtained for the overall combined CoMSIA model (entry 6), although the r^2 was slightly lower than without the lipophilic field (entry 5). However, as the overall combined fields model is the one providing the most descriptive information, the subsequently called CoMSIA model will refer to entry 6.

Similarly to the procedure applied in the previous CoMFA studies, each CoMSIA model with q^2 and the fraction of the variance (V) over 0.5 and 0.85, respectively, were used to predict the activity of the compounds in the testing set. In agreement to the observations on the q^2 values, the predictive r^2 obtained using the CoMSIA model are lower than the CoMFA ones (Table 7). Furthermore, the highest predictive r^2 is obtained with combined steric and electrostatic fields which correspond to the standard CoMFA.

Direct plots of CoMSIA predicted vs. observed activities using both training and testing sets are shown on Figures 7a and 7b respectively. Based on the data on Figure 7a and the r^2 value (0.923), the CoMSIA model provides a linear regression plot as satisfactory as that observed in the CoMFA studies for the

training set ($r^2 = 0.917$). Regarding the activity prediction of the testing set, the CoMSIA model displays a trend of linear correlation despite the low value of the predictive r^2 (0.352). Noteworthy is that most of the outlier compounds which have been observed in the previous CoMFA analysis of the testing set also occur whilst using the CoMSIA model. As previously discussed, the compounds for which the PDE4 affinity are over predicted are mainly racemic compounds but those with underpredicted pIC₅₀ are biologically singletons. Furthermore, for those compounds, the error of prediction is clearly higher using the CoMSIA model, affecting dramatically the predictive r^2 (0.352 vs. 0.649 in CoMFA). This is actually the result of the insufficiency in CoMSIA extrapolation accuracy, as compared to CoMFA. Of interest are two particular compounds (PD0190843, PD0190826) for which PDE4 affinities were well predicted by the CoMFA model but became outliers in using CoMSIA. The main structural feature of these compounds is the presence of a nitro group as benzodiazepine substituent. As the derived CoMSIA model has been built by the combination of steric, electrostatic, lipophilic and H-bond fields, it could be assumed that the addition of the two latter fields would have a negative effect on the affinity prediction on these nitro benzodiazepine derivatives. This emphasizes that the standard CoMFA fields provide the best prediction activity tools in the benzodiazepine series, whereas CoMSIA provides the best descriptive tools to design new active compounds.

Both CoMFA (steric and electrostatic) and CoMSIA (overall) models are predictive enough to derive 3D-contour maps. However, the CoMSIA contour maps (Figures 9–12) are smoother and display more information owing to the additional hydrophobic and are therefore easier to interpret than the corresponding CoMFA maps (not shown).

Table 8. PDE4 observed and predicted affinities for the training set ($\text{pIC}_{50} = -\log[\text{IC}_{50}]$)

Regid	pIC_{50} obs.	CoMFA		CoMSIA		H-QSAR		REGID	pIC_{50} obs.	CoMFA		CoMSIA		H-QSAR	
		Pred.	Err	Pred.	Err	Pred.	Err			Pred.	Err	Pred.	Err	Pred.	Err
PD0184268	6.43	6.57	0.14	6.29	-0.15	6.75	0.32	PD0190172	7.15	6.76	-0.39	6.85	-0.31	6.83	-0.32
PD0184440	6.96	7.14	0.18	7.33	0.37	7.34	0.38	PD0190194	6.92	6.61	-0.31	6.92	0.00	6.82	-0.10
PD0184441	6.74	6.61	-0.13	6.87	0.13	7.25	0.51	PD0190271	7.05	7.03	-0.02	7.09	0.04	7.04	-0.01
PD0185583	8.03	8.21	0.18	7.92	-0.11	7.91	-0.12	PD0190275	7.15	7.64	0.49	7.61	0.46	7.38	0.23
PD0185584	8.54	8.40	-0.14	8.13	-0.41	7.96	-0.58	PD0190820	4.99	5.56	0.57	5.46	0.47	5.61	0.62
PD0185585	6.08	6.63	0.55	6.50	0.42	6.93	0.85	PD0190821	5.81	6.00	0.19	5.74	-0.07	5.58	-0.23
PD0185586	5.83	5.74	-0.09	5.78	-0.05	5.65	-0.18	PD0190822	5.42	5.67	0.25	5.28	-0.15	5.52	0.10
PD0186470	7.00	6.74	-0.26	6.90	-0.10	6.88	-0.12	PD0190823	5.16	5.40	0.24	5.01	-0.15	5.63	0.47
PD0186476	6.47	6.46	-0.01	6.32	-0.15	6.48	0.01	PD0190824	5.59	5.43	-0.16	5.69	0.10	5.62	0.03
PD0186488	5.41	5.64	0.23	5.58	0.17	5.32	-0.09	PD0190825	5.69	5.81	0.12	5.85	0.16	5.87	0.18
PD0186505	6.43	6.53	0.10	6.42	-0.01	6.50	0.07	PD0190830	6.39	6.38	-0.01	6.49	0.10	6.07	-0.32
PD0186506	6.26	6.79	0.53	6.71	0.45	6.50	0.24	PD0190831	5.93	5.89	-0.04	5.91	-0.02	5.93	0.00
PD0186611	8.00	7.67	-0.33	7.45	-0.55	7.83	-0.18	PD0190832	5.81	5.83	0.02	5.80	-0.01	5.70	-0.11
PD0186620	6.80	6.85	0.05	6.68	-0.12	6.99	0.19	PD0190833	6.33	6.15	-0.18	6.28	-0.05	5.99	-0.34
PD0186628	7.10	7.23	0.13	7.37	0.27	7.56	0.46	PD0190834	5.72	5.70	-0.02	5.96	0.24	5.59	-0.13
PD0186630	6.80	7.18	0.38	7.42	0.62	7.33	0.53	PD0190835	5.81	5.67	-0.15	5.91	0.10	5.90	0.09
PD0186646	7.10	7.17	0.07	7.14	0.04	7.13	0.03	PD0190836	6.42	6.57	0.15	6.39	-0.03	6.24	-0.18
PD0186647	7.22	7.11	-0.11	6.88	-0.34	7.05	-0.17	PD0190838	5.24	5.23	-0.01	5.15	-0.09	5.64	0.40
PD0186658	7.05	7.08	0.03	7.03	-0.02	7.25	0.20	PD0190839	5.72	5.79	0.07	5.68	-0.04	5.78	0.06
PD0186719	7.30	7.24	-0.06	7.05	-0.25	6.78	-0.52	PD0190842	6.24	5.82	-0.42	6.07	-0.17	5.76	-0.48
PD0186720	7.30	7.18	-0.12	7.32	0.02	7.05	-0.25	PD0190844	4.81	4.88	0.07	4.67	-0.14	5.09	0.28
PD0186721	7.30	7.04	-0.26	7.15	-0.15	7.10	-0.20	PD0190845	5.22	4.87	-0.35	5.24	0.02	5.30	0.08
PD0186741	7.05	7.15	0.10	7.30	0.25	7.06	0.00	PD0190846	6.06	6.08	0.02	5.76	-0.30	6.09	0.03
PD0186743	5.72	5.74	0.02	5.51	-0.21	5.91	0.19	PD0190847	5.75	5.56	-0.19	5.90	0.15	5.45	-0.30
PD0188987	5.22	5.19	-0.04	5.12	-0.10	5.21	-0.01	PD0190848	5.08	5.23	0.15	5.10	0.02	5.57	0.49
PD0188988	5.29	5.10	-0.19	5.06	-0.23	5.15	-0.15	PD0190850	5.63	5.64	0.01	5.53	-0.10	5.69	0.06
PD0188989	5.01	5.09	0.08	5.04	0.03	4.98	-0.03	PD0190852	5.55	5.43	-0.12	5.67	0.12	5.75	0.20
PD0188990	5.48	5.31	-0.17	5.41	-0.07	5.59	0.11	PD0190853	6.01	5.96	-0.05	6.29	0.28	5.86	-0.15
PD0189117	4.73	4.77	0.04	4.75	0.02	4.82	0.09	PD0190854	6.02	6.13	0.11	6.16	0.14	5.55	-0.48
PD0189574	8.54	7.95	-0.59	7.91	-0.63	7.97	-0.57	PD0194016	6.89	6.74	-0.15	6.91	0.02	6.88	-0.01
PD0189657	6.64	6.14	-0.50	6.17	-0.47	6.28	-0.36	PD0194021	7.52	7.10	-0.42	7.32	-0.20	7.09	-0.43
PD0189658	6.55	6.59	0.04	6.70	0.15	6.28	-0.27	PD0194064	7.30	7.10	-0.21	7.15	-0.15	6.76	-0.54
PD0189659	6.46	6.57	0.11	6.55	0.09	6.45	-0.01	PD0197010	5.93	6.42	0.49	6.37	0.44	6.04	0.11
PD0189660	6.59	6.91	0.32	6.95	0.36	7.12	0.53	PD0199521	6.52	6.63	0.11	6.39	-0.13	6.73	0.21
PD0189794	6.17	6.41	0.24	6.28	0.11	6.55	0.38	PD0199522	6.74	6.95	0.21	6.82	0.08	6.74	0.00
PD0190168	6.62	6.35	-0.27	6.50	-0.12	6.74	0.12	PD0200310	6.19	6.03	-0.16	6.30	0.11	6.10	-0.09
PD0190171	6.85	6.69	-0.16	6.77	-0.08	6.68	-0.17	PD0200868	6.00	5.71	-0.30	5.91	-0.09	5.35	-0.65

The 3D-QSAR contour maps revealing the contribution of the CoMSIA steric field to the compound activity are shown in Figure 9. Sterically unfavorable regions are depicted in yellow, whereas favorable regions are in green. Striking is that large substituents in region **A** (see Figure 8) are detrimental to the activity indicated by the unfavorable steric contour maps S5. This confirms the assumption that a singleton com-

pound in the testing set such as PD0190814 bearing a bulky group at this region but exhibiting an unexpected experimental pIC_{50} high value (6.8) would have its potency predicted with less accuracy.

By moving along the benzodiazepine structure, a small favorable steric area S4 has been found in the region **B** which is the result of the methyl group of the pyridinone series producing higher activity for com-

Table 9. PDE4 observed and predicted affinities for the testing set ($\text{pIC}_{50} = -\log[\text{IC}_{50}]$)

Regid	pIC ₅₀ obs.	CoMFA		CoMSIA		H-QSAR		%	REGID obs.	pIC ₅₀ Pred.	CoMFA Err	CoMSIA Pred.	CoMSIA Err	H-QSAR Pred.	H-QSAR Err	%	
		Pred.	Err	Pred.	Err	Pred.	Err										
PD0184270	6.66	6.76	0.10	6.89	0.23	7.07	0.41	5	PD0186737	7.40	7.12	−0.28	7.26	−0.14	6.99	−0.41	5
PD0184271	6.54	6.49	−0.05	6.95	0.41	7.10	0.56	9	PD0186744	7.22	7.07	−0.15	6.83	−0.39	6.70	−0.52	2
PD0184272	6.96	6.99	0.03	7.06	0.10	7.08	0.12	5	PD0186745	7.52	7.75	0.23	7.46	−0.06	7.08	−0.44	1
PD0184273	7.05	6.78	−0.27	6.81	−0.24	6.78	−0.27	6	PD0186747	5.58	6.86	1.28	6.89	1.31	6.82	1.24	3
PD0184274	6.72	7.09	0.37	7.22	0.50	7.01	0.29	2	PD0186748	6.57	7.50	0.93	7.47	0.90	6.74	0.17	7
PD0184275	6.96	6.87	−0.09	6.73	−0.23	6.77	−0.19	5	PD0186750	5.38	5.55	0.17	5.91	0.53	5.73	0.35	5
PD0184276	7.15	7.08	−0.07	7.33	0.18	7.04	−0.11	1	PD0186754	6.89	7.05	0.16	7.01	0.12	6.90	0.01	1
PD0184277	7.10	6.93	−0.18	6.87	−0.23	7.09	−0.01	5	PD0186755	6.01	7.38	1.37	7.14	1.13	7.08	1.07	0
PD0184278	6.26	7.09	0.83	7.07	0.81	7.02	0.76	5	PD0188991	6.82	5.57	−1.26	5.59	−1.23	5.14	−1.68	0
PD0184279	6.92	6.45	−0.48	6.86	−0.06	7.07	0.15	4	PD0189747	7.40	7.41	0.01	6.77	−0.63	7.06	−0.34	5
PD0186622	7.40	7.53	0.13	7.38	−0.02	7.82	0.42	3	PD0190150	7.10	6.59	−0.51	6.89	−0.21	6.76	−0.34	3
PD0186706	4.53	5.73	1.20	5.57	1.04	5.63	1.10	5	PD0190184	6.64	6.11	−0.53	6.00	−0.65	5.86	−0.78	0
PD0186718	6.68	6.74	0.06	6.72	0.04	7.01	0.33	7	PD0190187	6.66	7.31	0.65	7.59	0.93	7.55	0.89	4
PD0186724	7.22	7.02	−0.20	7.01	−0.22	6.85	−0.37	13	PD0190597	4.76	5.73	0.97	5.54	0.78	5.65	0.89	3
PD0186725	7.40	7.47	0.07	7.29	−0.11	7.10	−0.30	4	PD0190826	5.87	6.13	0.26	4.32	−1.55	5.67	−0.21	7
PD0186729	7.05	7.09	0.04	7.16	0.11	7.02	−0.03	1	PD0190827	4.81	6.02	1.21	5.60	0.79	5.59	0.78	4
PD0186731	6.89	6.64	−0.25	6.14	−0.75	7.00	0.11	4	PD0190828	4.80	5.56	0.76	5.84	1.04	5.77	0.97	23
PD0186732	7.52	7.06	−0.46	7.24	−0.28	6.99	−0.53	5	PD0190837	4.37	5.55	1.18	5.75	1.38	5.79	1.42	14
PD0186733	7.40	7.44	0.04	7.29	−0.11	7.06	−0.34	0	PD0190840	4.68	5.70	1.02	5.75	1.07	5.85	1.17	3
PD0186734	7.40	7.51	0.11	7.40	0.00	7.04	−0.36	6	PD0190843	5.88	5.91	0.03	4.58	−1.30	5.85	−0.03	24
PD0186735	7.00	7.54	0.54	7.34	0.34	7.00	0.00	1	PD0190849	5.01	5.84	0.83	5.48	0.47	5.45	0.44	2
PD0186736	6.89	6.94	0.05	7.09	0.20	6.99	0.10	8	PD0198250	5.44	6.19	0.75	6.29	0.85	7.01	1.57	16

pounds from this structural subfamily. In the region **C**, an unfavorable area (S1) which is located slightly outside of the benzodiazepine side chain, indicates that large substituents in this area significantly reduce the PDE4 affinity. Conversely, medium size substituent at the edge of the side chain enhances the binding affinity as supported by the presence of the favorable green contour map S2. These results show that a steric effect is needed at the region S2 but its size should not expand up to reaching the unfavorable region S1. Furthermore, the presence of a small unfavorable yellow contour S3 in the vicinity of the benzodiazepine side chain suggests that this part of the molecule should be constituted by a 'thin' structural moiety such as aromatics for an optimum potency. This assumption is supported by the fact that flattening the benzodiazepine side chain by a conformational constraint (cyclisation between R_6 and R_7) in the region **B** and **C** improves the PDE4 affinity as observed in the sub-series **C** (Table 3). The contribution of the CoMSIA electrostatic field is illustrated by the contour maps in Figure 10. Electrostatic positive favorable (or negative unfavorable) regions correspond to the blue areas, and

the negative favorable (or positive unfavorable) is represented by the red areas. Note however that, as for the CoMFA electrostatic fields, these regions might result from unfavorable interactions rather than from favorable ones. Striking is the importance of negative charge (E4) in the close vicinity of the aromatic moiety (region **A**), owing to the high affinity of the primary amino derivatives. Similarly, the two blue areas E5 and E6 are related to the low affinity of the esters and amides for E5 and aromatic groups for E6 on R_1 . This indicates that negatively charged residue in these regions reduce PDE4 affinity. Similarly, a blue area is located close to the side-chain tip (E1), which may be attributed to the methyl ester carbonyl group (PD0185584) in which electron withdrawing character is probably responsible for the arising blue area (E3). It is noteworthy that region E3a is thinner than E3b, owing to the affinity of 3-pyridyl derivatives, revealing a probable hydrogen bond acceptor feature in this area, or the metal chelation site. Indeed, the presence of the metal binding site in the PDE4 catalytic center has been investigated and confirmed by the X-ray structure of the PDE4 enzyme [18]. A model describing the

interaction of some rolipram-based PDE4 inhibitors with a metal was already proposed in the literature [19]. These reported data are consistent to the assumption that the 3-pyridyl moiety in the benzodiazepine favors the metal interaction for an optimum potency. At the region **B**, the CoMSIA analysis was not able to provide any information on the electrostatic field contribution.

The hydrophobic field contributions are depicted in Figure 11. Lipophilic favorable regions are colored in orange, and hydrophilic favorable regions are colored in cyan. In agreement with the electrostatic fields, the main hydrophilic contributions are from the two polar regions at both molecule tips (L1 and L5). Two lipophilic favorable regions appear in the immediate vicinity of the aromatic side chain (L2 and L4), owing to the good affinity of PD0189574 (2-Cl) and PD0190172 (2,5-Cl-4-amino) respectively. However, the presence of the lipophilic area L2 might also result from the decrease affinity of the diazo derivative PD0186658. It is noteworthy that although the lipophilic field did not improve the predictiveness of the model, the 3D-maps clearly correlate with the biological data, and therefore confirms a probable redundancy in the descriptor variables.

The hydrogen bond field contributions are shown in Figures 12 and 13. Favorable hydrogen bond acceptor and donor features (on the ligand) are colored in red and blue respectively, whereas the unfavorable regions are in yellow. The hydrogen donor field confirms the importance of the 1-amino group. Although both directions seem important (see areas D4 and D5), according to both steric and hydrogen bond acceptor fields, a single hydrogen bond is involved, allowing single non-hydrogen bond acceptor, medium size substituents on the nitrogen. The red contour maps A2a and A2b confirm the importance of the hydrogen bond acceptor feature in the neighborhood of the benzodiazepine side chain as observed on the electrostatic (E2a, E2b) field contribution maps.

Despite the lack of structural diversity at the region **B**, an unfavorable hydrogen bond acceptor area (A7) has been found. Noteworthy is that in the contour map, A7 is out of plane with respect to the conformationally restricted side chain. This suggests that when the carbonyl group in the side chain is not coplanar to its own phenyl moiety, the optimal compound potency is not reached. This provides explanations why the amide side chains in which the carbonyl groups are in a slightly staggered conformation with respect to the phenyl moiety are less active than pyridone. By exam-

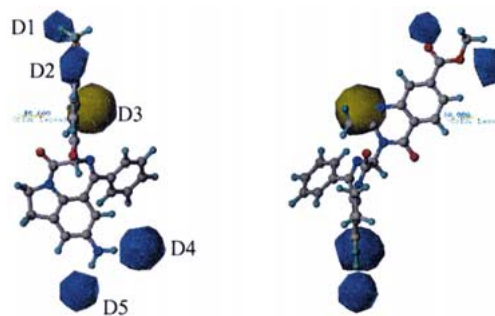


Figure 12. CoMSIA H-bond donor 3D-contour map.

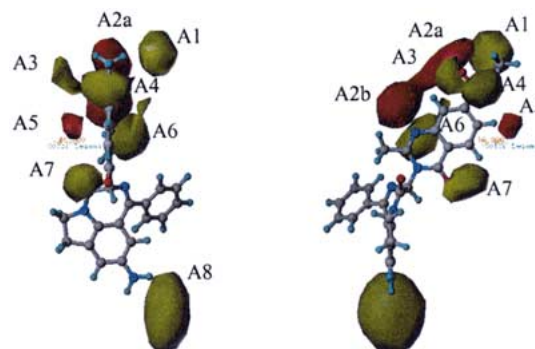


Figure 13. CoMSIA H-bond acceptor 3D-contour map.

ining the hydrogen bond acceptor field produced by the substituent at R_1 , the unfavorable contour map A8 confirms how carbonyl substituents lead to decreased affinity in this region. As previously mentioned for the lipophilic field, most of the H-bond depicted areas confirm the steric and electrostatic indicators and are therefore mostly the result of redundant variables.

By combining the five 3D-CoMSIA contour maps, several pharmacophores for PDE4 inhibitors can be identified (Figure 14): one H-bond donor, three H-bond acceptors, two polar regions and one hydrophobic sites. Two hydrophobic sites have been included in the PDE4 inhibitor pharmacophores in order to represent the phenyl and indolenine moieties at the regions A and B despite the fact that no precise information on the structural requirements has been obtained for these areas due to the lack of local structural diversity. Based on the hydrophobicity of these two groups, it could be assumed that these two groups are involved with the PDE4 catalytic center by hydrophobic interactions. As compared to the previously reported pharmacophore models of PDE4 inhibitors [4], it is noteworthy that our model has three hydrophobic regions, whereas others have one and two respectively. Similarly, the hydrogen bond acceptor feature distribution is slightly

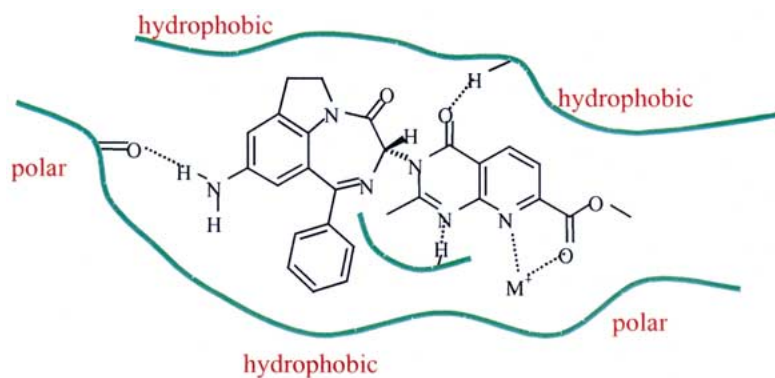


Figure 14. Derived pharmacophoric model.

different from those of previously reported models. It is therefore more likely that our series of compounds have a different binding mode in PDE4.

Further comparison of this model with the X-ray crystal structure of PDE4 active site, led to the conclusion that the polar area located at the vicinity of the benzodiazepine side chain tips defined by either 3-pyridyl or carboxyl group should be involved in the metal chelation, whereas the indolenine core should mimic cAMP purine moiety and the phenyl group would fit the hydrophobic pocket S2.

Conclusion

A series of benzodiazepine, inhibitors of PDE4 were successfully used to build a 2D-QSAR (H-QSAR) and two 3D-QSAR (CoMFA and CoMSIA) models. These three models gave new insights into the pharmacophores involved in PDE4 inhibition and were used to design new PDE4 inhibitors and to screen a large library of in-house benzodiazepine derivatives.

Comparison of the 3D-QSAR models led to the conclusion that CoMSIA is less alignment sensitive than CoMFA and it produces smoother 3D-QSAR contour maps which therefore enhance the understanding of the model. Its predictiveness however is significantly less accurate for compounds with substituents outside the regions explored by the model. Indeed, the use of lattice points closer to the molecule area in CoMSIA, as compared to CoMFA, not only leads to smoother contour maps, but also contributes to dropping variables corresponding to lattice points located far from the molecule areas. In contrast, CoMFA enlarges the explored region and therefore reduces the number of extrapolated points. Furthermore, CoMSIA usually requires a smaller grid spacing

which improves both q^2 , and r^2 , but dramatically increases the calculation time. In addition, we found that although both lipophilic and hydrogen bond fields within CoMSIA are mostly redundant with the steric and electrostatic fields their resulting contour maps are easy to interpret and very interesting.

Therefore, CoMSIA appears to be a valuable tool to help understanding of structure-activity relationships. On the other hand, CoMFA is very good in predicting activity of relatively diverse compounds within a series. The benefits of the two approaches are therefore complementary, and combined use of H-QSAR and CoMFA models has provided a valuable tool in the field of screening large libraries. H-QSAR has been used to screen out the less potent PDE4 inhibitors, whereas the remaining most potent ones were aligned and predicted for PDE4 activity using the CoMFA model. Application of the CoMSIA approach adds a molecular understanding (the pharmacophoric model) of these results.

Comparison of the CoMSIA 3D-contour maps with the recently reported structure of the PDE4 active site led to the conclusion that the polar area located at the tip ends of the benzodiazepine side chain (C3-substituent) should be involved in the metal chelation site. Based on the size of the PDE4 active site which is relatively large (440 \AA^3) [3], the different series of PDE4 inhibitors could exhibit different binding modes. In consequence, comparative studies of the pharmacophoric models from different series would generate results with lack of insights.

The next step of our study will be focused on the validation of the present model by docking analysis of the benzodiazepine series into the crystal structure of the PDE4 active site.

In the meantime, the obtained 2D and 3D QSAR models can be used to predict activity of new structures in the benzodiazepine series. In addition, it has been shown that the combination of the CoMFA model with the crystal structure of the molecular target provides an activity prediction tool better than the traditional docking scoring function alone. Hence, even when the X-ray data of the PDE4 enzyme is available, the CoMFA model will remain a powerful complementary tool for the design of new PDE4 inhibitors.

References

1. Torphy, T.J., Stadel, J.M., Burman, M., Cieslinski, L.B., McLaughlin, M.M., White, J.R. and Livi, G.P. *J. Biol. Chem.*, 267 (1992) 1798.
2. Barnette, M.S., Grous, M., Cieslinski, L.B., Burman, M., Christensen, B. and Torphy, T.J. *J. Pharmacol. Exp. Ther.*, 273 (1995) 1396.
3. Xu, R.X., Hassell, A.M., Vanderwall, D., Lambert, M.H., Holmes, W.D., Luther, M.A., Rocque, W.J., Milburn, M.V., Zhao, Y., Ke, H. and Nolte, B.T. *Science*, 288 (2000) 1822.
4. a. Polymerooulos, E.E. and Höfgen, N.A. *Quant. Struct.-Act. Relat.*, 16 (1999) 231
b. Segarra, V., Crespo, M.I., Pujol, F., Beleta, J., Domenech, T., Miralpeix, M., Palacios, J.M., Castro, A. and Martinez, A. *Bioorg. Med. Chem. Lett.*, 8 (1998) 505
c. Crespo, M.I., Pages, L., Vega, A., Segarra, Lopez, M., V., Domenech, T., Miralpeix, M., Beleta, J., Ryder, H. and Palacios, J.M. *J. Med. Chem.*, 41 (1998) 4021–4035.
d. Polymerooulos, E.E. and Höfgen, N.A. *Quant. Struct.-Act. Relat.*, 18 (1999) 543.
5. Pascal, Y., Andrianjara, C.R., Auclair, E., Avenel, N., Bertin, B., Calvet, A., Feru, F., Lardon, S., Moodley, I., Ouagued, M., Payne, A., Pruniaux, M.P. and Szilagyi, C. *Bioorg. Med. Chem. Lett.*, 10 (2000) 35.
6. Sybyl 6.6 ed., Tripos Associates Ltd., Missouri, USA, 1992.
7. a. Heritage, T.W. and Lowis, D.R. *Molecular hologram QSAR in Rational Drug Design: Novel Methodology and Practical Applications*. ACS Symposium Series Vol. 719, in press.
b. Tong, W., Lowis, D.R., Perkins, R., Chen, Y., Welsh, W.J., Goddette, D.W., Heritage, T.W. and Sheehan D.M. *J. Chem. Inf. Comput. Sci.*, 38 (1998) 669.
8. QSAR GABA au moyen des ANN.
9. Sybyl 6.6, H-QSAR Documentation, Tripos Associates Ltd.
10. Cramer, R.D. III, Patterson, D.E. and Bunce, J.D. *J. Amer. Chem. Soc.*, 110 (1998) 5959.
11. Klebe, G., Abraham, U. and Mietzner, T., *J. Med. Chem.*, 37 (1994) 4130.
12. a. Burden, F.R. and Winkler, D.A. *J. Chem. Inf. Comput. Sci.*, 39 (1999) 236.
b. Winkler, D.A., Burden, F.R. and Watkins, A.J.R. *Quant. Struct.-act. Relat.*, 17 (1998) 14.
c. Sinha, J., Kurup, A., Paleti, A., Gupta, S.P. *Bioorg. Med. Chem.*, 000 (Author please check) (1999) 11.27.
d. Tokarski, J.S. and Hopfinger, A.J. *J. Med. Chem.*, 37 (1994) 3639.
13. Burnouf, C., Auclair, E., Avenel, N., Bertin, B., Bigot, C., Calvet, A., Chan, K., Durand, C., Fasquelle, V., Féru, F., Gilbertsen, R., Jacobelli H., Keksi, A., Lallier, E., Maignel, J., Martin, B., Milano, S., Ouagued, M., Pascal, Y., Pruniaux, M.-P., Puaud, J., Rocher, M.-N., Terrasse, C., Wrigglesworth, C. and Doherty A.M. *J. Med. Chem.*, 43 (2000) 4850.
14. Blount, J.F., Fryer, I.R., Gilman, N.W., Todaro, L.J. *Mol. Pharmacol.* 24 (1983) 425–428.
15. a. Pearlman, R.S. *Chem. Des. Auto. News*, 2 (1987) 1.
b. CONCORD Manual, Tripos, 1995.
16. Halgren, T. J. *Am. Chem. Soc.*, 112 (1990) 4710.
17. Baker, J. J. *Comp. Chem.*, 7 (1986) 385.
18. a. Omburo, G.A., Jacobitz, S., Torphy, T.J. and Colman, R.W. *Cell. Signalling*, 10 (1998) 491.
b. Jacobitz, S., Ryan, M.D., McLaughlin, M., Livi, G.P., De-Wolf, W.E., Jr. and Torphy, T. J. *Mol. Pharmacol.*, 51 (1997) 999.
19. Kleinman, E., Campbell, E. Giordano, L., Cohan, V., Jenkinson, T., Cheng, J., Shirley, J., Pettipther, E.R., Salter, E.D., Hibbs, T.A., DiCapua, F.M. and Bordner, J. *J. Med. Chem.*, 41 (1998) 266.