Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2007) 494–504
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Three-dimensional QSAR of HPPD inhibitors, PSA inhibitors, and anxiolytic agents: Effect of tautomerism on the CoMFA models

Jian-Wei Zou [a,*], Cheng-Cai Luo [a], Hua-Xing Zhang [a], Hai-Chun Liu [a,b], Yong-Jun Jiang [a], Qing-Sen Yu [a,b]

[a] *Key Laboratory for Molecular Design and Nutrition Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, China*
[b] *Department of Chemistry, Zhejiang University, Hangzhou 310027, China*

Received 22 October 2006; received in revised form 31 January 2007; accepted 7 March 2007
Available online 12 March 2007

Abstract

The present study was design to examine the effect of tautomerism upon the CoMFA results. Three selected data sets involving protropic tautomerism, which are 21 *p*-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors, 35 inhibitors of puromycin-sensitive aminopeptidase (PSA), and 67 anxiolytic agents, were used for this purpose. Atom-by-atom alignment technique was adopted to superimpose the molecules in the data sets onto a template. The structural alignments using different tautomeric forms had no significant difference except the atoms involved in tautomerism, which ensures, to a great extent, that the differences of the CoMFA results result primarily from the tautomerism. All-orientation and all-placement search (AOS-APS) based CoMFA models, in addition to the conventional ones, were derived for each system and proved to be capable of yielding much improved statistical results. In the cases of the data sets of HPPD inhibitors and PSA inhibitors, excellent AOS-APS CoMFA models ($q^2 > 0.8$ with four components for the former and $q^2 > 0.7$ with seven components for the latter) were obtained, and almost no significant difference in statistical quality was observed when using different tautomeric forms to derive the models. However, it was not the case when treating the data set of anxiolytic agents. The keto tautomer, which was the active form of the PBI type inhibitors, produced measurably better results ($q^2 = 0.54$ with eight components) than that the enol one ($q^2 = 0.37$ with five components), indicating the importance of selecting proper tautomer in the CoMFA studies. Furthermore, there existed some substantial differences of the electrostatic field contours between the two different tautomeric forms for all of the three systems considered, whereas the differences in the steric field contour maps were limited. This implies that the resulting new potent ligands may be quite different if one utilizes the CoMFA models of different tautomeric forms for guiding further structural refinements.

Keywords: Tautomerism; 3D-QSAR; CoMFA; Molecular modeling

1. Introduction

Comparative molecular field analysis (CoMFA) has been widely used as a powerful 3D-QSAR tool in the field of medicinal chemistry since it was introduced by Cramer in the late 1980s [1,2]. In CoMFA, electrostatic and steric interaction energies are calculated between a probe atom (most often an sp^3 hybridized carbon atom with a +1 charge) located at each vertex of a space grid and a series of molecules embedded within the grid. Then a statistical analysis using partial least squares (PLS) technique is conducted to establish the relationship between the interaction energies and biological activities of molecules under study. The result of the analysis is usually presented as a set of contour maps which show favorable and unfavorable regions for electronegative (electrostatic field map) and bulky (steric field map) substituents in certain positions.

As a standard field-based 3D-QSAR technique, CoMFA will undoubtedly continue to make a valuable contribution to the drug design. Nevertheless, several problems associated with this methodology itself have persisted [3–5]. First, CoMFA requires the specification of conformations of molecules in the data set and, in many cases, the low-energy conformations derived from theoretical calculations are used as an alternative to bioactive conformations. Second, structural alignments of the molecules under consideration, which reflect approximately their putative binding geometry toward the target receptor, are

* Corresponding author. Tel.: +86 574 88229517; fax: +86 574 88229516.
E-mail address: jwzou@nit.net.cn (J.-W. Zou).

1093-3263/$ – see front matter
doi:10.1016/j.jmgm.2007.03.002

necessarily required. However, there is so far no general rule to align all molecules in a ''correct'' fashion. It seems inevitable that the alignments deviate more or less from the real situation due to lack of sufficient structural information of ligand-receptor complexes. In addition, various options that need to be set in the CoMFA analysis, such as overall orientation of aligned compounds, position of the lattice points, step size, the probe atom type, often significantly influence the results. During the past decade, many CoMFA-related methods have been developed to overcome these disadvantages or improve the quality of CoMFA models [6–8].

Tautomerism is an important phenomenon in chemistry and biology. It is estimated that up to 0.5% of commercially available compounds possess tautomers [9]. In a recent review, Folkers et al. [10] raised the issue of tautomerism in computer-aided drug design through collecting and analysing some examples of log P calculation, similarity index, and the complementarity pattern towards the receptor. Especially, they pointed out the significant impact of considering tautomerism on the prediction of ligand-receptor interactions using various docking approaches [10].

As tautomeric forms of a molecule exhibit difference in functional groups, it might expect that they have different electrostatic and steric field distributions, at least around the involved functional groups, and therefore affect the field-based 3D-QSAR results. This issue, however, was usually neglected in prior 3D-QSAR studies, and in most cases, only the presumed most stable tautomer was used [11]. Although the tautomers in the gas phase or in solution can be determined utilizing modern analytic techniques or advanced quantum mechanical calculations [12,13], little is known about the proper tautomeric form when a ligand binds to the active site of a target protein—an environment quite different from the solution and vacuum. Moreover, in previous reports, the tautomeric state of a ligand, which exists rarely in the gas phase and in solution, has indeed been detected at the enzymatic active sites [14].

The present paper was designed to examine the effect of tautomerism upon the CoMFA results. Three data sets involving protropic tautomerism, which are *p*-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors [15–20], inhibitors of puromycin-sensitive aminopeptidase (PSA) [21], and anxiolytic agents [22–24], have been investigated for this purpose.

2. Methodology

All data sets were taken from the published results [15–24]. The molecular geometries of each compound and its tautomer were optimized with MOPAC 6.0 program implemented in VEGA package [25] using AM1 method (keywords: ''PRECISE'', ''GEO-OK''). CoMFA studies were performed on a Silicon Graphics workstation using the QSAR module in Sybyl 6.8 [26]. Charges were calculated by the Gasteiger-Hückel method [27,28]. The alignment technique based on atom-by-atom least-square fit was used to superimpose the molecules in the data set onto a template. The CoMFA region, probe atom ($C_{sp^3}{}^{+}$), as well as the cutoff energy of steric and electrostatic contributions (30 kcal/mol) were set to the default parameters. A distance dependent dielectric constant was used.

Since CoMFA models are highly sensitive to the different space orientations of the molecular aggregate with respect to the lattice, all-orientation search (AOS) and all-placement search (APS), as described previously [29], were also conducted using the translation and rotation procedures written in SYBYL programming language (SPL). For each orientation and placement of the molecular aggregate, a conventional CoMFA was performed to obtain the optimal model.

PLS analysis was conducted to establish the relationship between the biological activities and the resulting field matrix. Leave one out (LOO) cross-validation was employed to select the number of principal components (NOPC). Equal weights were assigned to steric and electrostatic descriptors using the CoMFA scaling option, and a minimum σ (column filter) value of 2.0 kcal/mol was adopted to minimize the influence of column noise and reduce the computation time. Then PLS analysis was repeated without cross-validation using the optimum NOPC to obtain the final statistical parameters, including the conventional correlation coefficient (r^2), standard deviation (S.D.) and F value, and to derive the CoMFA coefficient contour maps of steric and electrostatic fields.

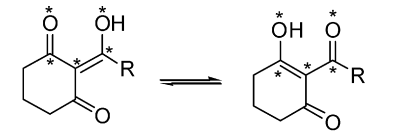
3. Results and discussion

3.1. Data set of HPPD inhibitors

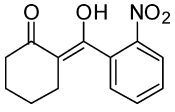
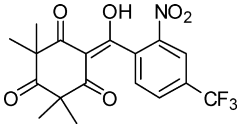
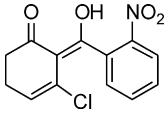
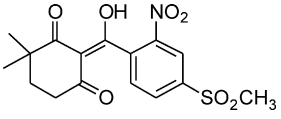
HPPD is a non-heme Fe(II)-dependent enzyme involved in the metabolism of tyrosine in most organisms and in the biosynthesis of tocopherols in plants [30,31]. Potential HPPD inhibitors provide an alternative strategy for treating the life-threatening tyrosinaemia type I disease [30], and several of them, including mesotrione, sulcotrione, and isofluxatole, are currently in use as selective broad leaf herbicides [18,19]. Twenty one compounds containing acyl-cyclohexan-2-one structural fragment and their HPPD inhibitory potencies, expressed as pIC_{50} (M), are listed in Table 1. These compounds, especially those possessing acyl-cyclohexan-1,3-dione skeleton, have many possible tautomeric forms, of which two, as shown in Table 1, have been identified by different techniques to exist [16,32–34]. While theoretical calculations carried out by us [32], crystallography data [16,33] as well as the NMR experiments in $CDCl_3$ solution [34] suggested that the endocyclic enol-tautomer (II) was slightly predominant for compounds of this type, a later NMR study [35] revealed that one of these compounds, 2-[2-nitro-4-(triflouromethyl)benzoyl]-1,3-cyclohexanedione (NTBC, **H9**), existed solely as exocyclic enol-form (I) in aqueous solution at pH 7.0.

Both tautomers I and II were used in the CoMFA study. Using the atoms marked with asterisk (*) in Table 1 (i.e. the β-diketone moiety), and NTBC as the template molecule, all of the 21 molecules were aligned by minimizing the root-mean-square (rms) coordinate difference of the selected atoms. Such an alignment rule was applied because it has been found in the crystal stucture of NTBC-HPPD complex that the β-diketone coordinated bidentately (chelating) with the Fe^{2+} ion of the

Table 1
Structures and activities of HPPD inhibitors



exocyclic enol (I) *endocyclic enol* (II)

	R=	pIC_{50} (M)[a] exptl	pIC_{50} (M)[a] calcd (I)	pIC_{50} (M)[a] calcd (II)
H1	Ph	4.95	5.27	5.09
H2	2-Cl-Ph	6.30	6.16	6.18
H3	2-Br-Ph	6.25	6.12	6.26
H4	2-I-Ph	6.12	6.11	6.27
H5	2-NO_2-Ph	6.80	6.58	6.56
H6	2-CH_3-Ph	5.43	5.34	5.32
H7	2-CF_3-Ph	6.60	6.63	6.73
H8	2-OMe-Ph	4.93	4.92	5.04
H9	2-NO_2-4-CF_3-Ph	7.40	7.51	7.45
H10	2-Cl-4-SO_2CH_3	7.30	7.36	7.42
H11	Methyl	4.95	5.09	5.12
H12	Ethyl	4.75	4.86	4.73
H13	Isopropyl	4.03	4.38	4.24
H14	Cyclopropyl	5.22	4.82	4.58
H15	Cyclohexyl	3.44	3.27	3.49
H16	2-Cl-3-Oet-4-SO_2CH_3-Ph	8.70	8.66	8.53
H17	2-NO_2-4-SO_2CH_3-Ph	7.57	7.59	7.56
H18		6.16	6.27	6.30
H19		7.30	7.36	7.43
H20		6.28	6.25	6.27
H21		7.55	7.46	7.45

[a] Expressed as the logarithm of $1/IC_{50}$ (M) value.

enzyme [36]. The aligned molecules are shown in Fig. 1(a) (tautomer I) and Fig. 1(b) (tautomer II), respectively. The structural alignments show no significant difference between the two tautomeric forms except the atoms involved in tautomerism, implying that the tautomerism has little effect on the overall molecular conformation. For tautomer II, the conformations of compounds **H19** and **H20**, with respect to the R–C(O) bond, differ slightly from those of other benzoylcy-

Table 2
Summary of CoMFA results for HPPD inhibitors

	q^2	S.E.P.	Optimal NOPC	r^2	S.D.	F	Contribution % steric/electrostatic
Exocyclic enol: conventional	0.775	0.700	4	0.970	0.254	131.486	51.6/48.4
AOS_APS	0.836	0.598	4	0.983	0.190	236.757	57.4/42.6
Endocyclic enol: conventional	0.747	0.742	4	0.962	0.286	102.556	50.7/49.3
After AOS_APS	0.839	0.593	4	0.979	0.214	186.088	65.0/35.0

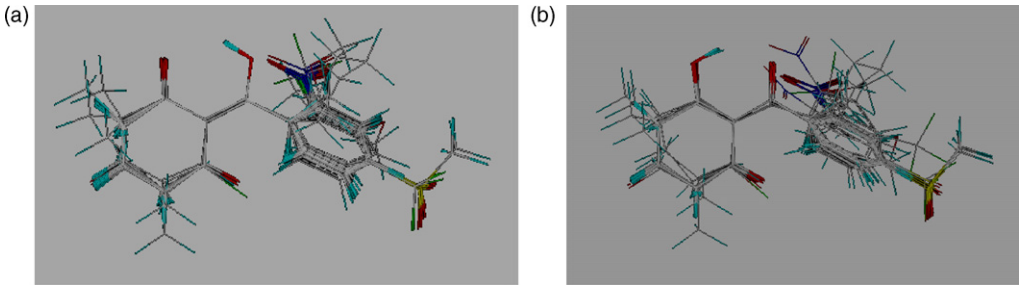


Fig. 1. Structural alignments of all molecules in the data set of HPPD inhibitors: (a) exocylic enol tautomer; (b) endocyclic enol tautomer.

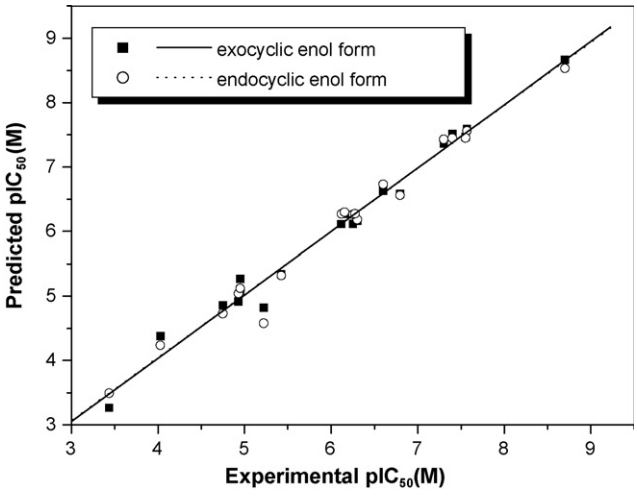


Fig. 2. Plots of the experimental versus predicted pIC_{50} values for the exocylic enol and endocyclic enol CoMFA models of HPPD inhibitors.

clohexanones, which makes Fig. 1(b) look somewhat more cluttered than Fig. 1(a).

The key statistical parameters associated with the CoMFA models for both tautomers are summarized in Table 2. As seen, conventional CoMFAs show good correlation, as reflected by the large cross-validated q^2 values, 0.755 for the tautomeric form I and 0.747 for the tautomeric form II with four components. After the AOS and APS procedures were used, the q^2 values further increase to 0.836 and 0.839, respectively. Also improved by the AOS and APS are the conventional r^2, standard errors of predictions (S.E.P.), standard deviation of estimates (S.D.) and F values of the models (see Table 2). The predicted activities by the AOS-APS CoMFA models are given in Table 1. The excellent quality of these models is also illustrated in Fig. 2, which displays plots of the predicted versus experimental pIC_{50} values for the whole set compounds. In addition, it has been revealed by the both conventional CoMFAs models that the relative contributions of steric and electrostatic fields are about equal, indicating the importance of both terms. The AOS-APS CoMFAs highlight the contribution of steric field, especially for tautomer II.

AOS-APS CoMFAs were selected to construct the contour maps of steric and electrostatic fields for both tautomeric forms. The resulting contour maps as well as the compound **H9** (NTBC) as the reference structure is represented in Fig. 3. Here and hereafter, in the CoMFA steric field, green (80% level

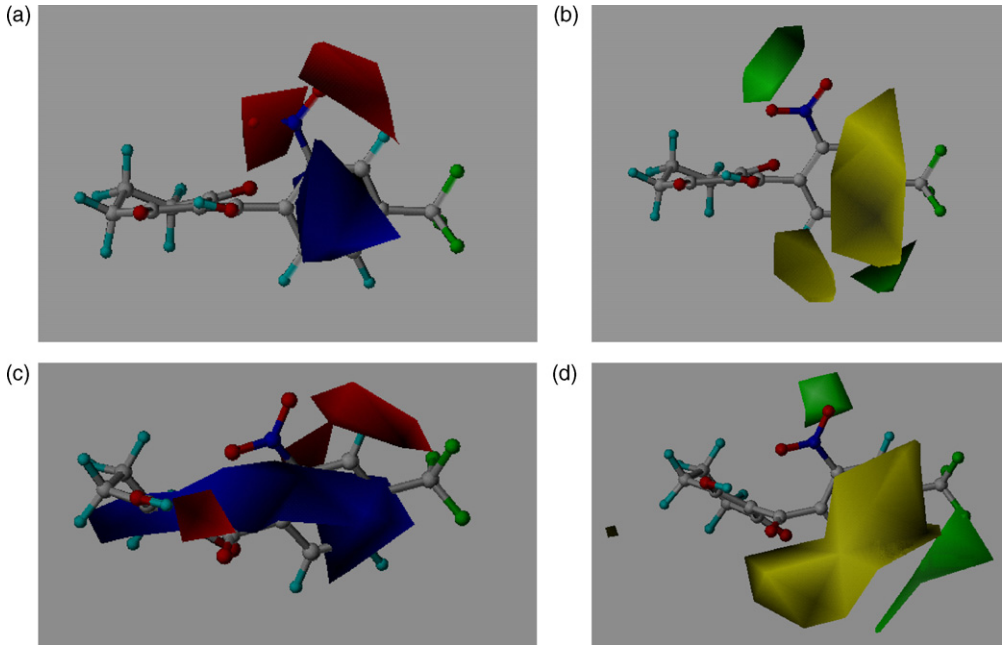
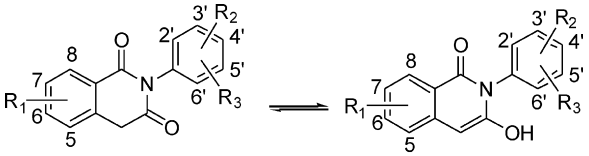


Fig. 3. (a) Electrostatic field and (b) steric field contours generated with the exocyclic enol AOS-APS CoMFA model for HPPD inhibitors. (c) Electrostatic field and (d) steric field contours generated with the endocyclic enol AOS-APS CoMFA model for HPPD inhibitors.

contribution) indicates regions where bulky groups increase activity, whereas yellow (20% level contribution) indicates regions where bulky groups decrease activity. Similarly, in the CoMFA electrostatic field, red (80% level contribution) represents regions where more negatively charged groups increase activity, whereas blue (20% level contribution) represents regions where more positively charged groups increase activity. It has been shown in the electrostatic contour map (Fig. 3(a)) of the exocyclic enol CoMFA model that there are two red regions around the nitro group, implying that an electron withdrawing substituent is beneficial to the activity. This is well consistent with previously proposed structure-activity relationship [18]. The center of the phenyl ring falls into the sole blue region. Since most compounds in the data set possess a benzoyl structural fragment, and an electron-withdrawing substituting can increase the positive potential potentials at the center of the phenyl ring, such a positive charge favorable blue region suggests that an introduction of the electron withdrawing group on the phenyl ring (at the 2′- or other positions, e.g. 4′-CF_3 in the compounds **H9**, **H10**, **H16**, **H18**, **H19**, and **H21**), would also result in an improved activity.

The electrostatic contour map (Fig. 3(c)) for tautomer II is somewhat more complicated. The two red regions in Fig. 3(a), although also occur, deviate somewhat from the position of the nitro group of NTBC, which can be ascribed to the imperfect overlap of the nitro groups in the structural alignment (see

Table 3
Structures and activities of PSA inhibitors[a]



	R_1	R_2	R_3	pIC_{50} (M)[b] exptl	pIC_{50} (M)[b] calcd (enol)	pIC_{50} (M)[b] calcd (keto)
Q1	H	H	H	3.40	3.51	3.55
Q2	H	2′-Me	H	3.78	3.75	3.85
Q3	H	2′-Et	H	4.17	4.18	3.98
Q4	H	3′-Et	H	4.36	4.37	4.30
Q5	H	4′-Et	H	3.42	3.34	3.44
Q6	H	2′-^{i}Pr	H	3.71	3.69	3.64
Q7	H	4′-^{i}Pr	H	3.45	3.45	3.32
Q8	H	2′-OMe	H	4.63	4.62	4.60
Q9	H	3′-OMe	H	4.45	4.51	4.49
Q10	H	4′-OMe	H	4.53	4.58	4.58
Q11	H	2′-SMe	H	5.50	5.37	5.50
Q12	H	3′-SMe	H	4.55	4.50	4.62
Q13	H	4′-SMe	H	4.66	4.66	4.64
Q14	H	2′-Me	6′-Me	4.48	4.51	4.56
Q15	H	2′-Me	5′-Me	4.36	4.36	4.41
Q16	H	2′-Me	4′-Me	4.20	4.03	3.99
Q17	H	2′-Me	3′-Me	3.95	3.99	4.01
Q18	H	3′-Me	4′-Me	4.07	4.15	4.22
Q19	H	3′-Me	5′-Me	5.25	5.13	5.08
Q20	H	2′-Et	6′-Et	6.39	6.41	6.35
Q21	H	2′-Et	5′-Et	4.83	4.80	4.88
Q22	H	2′-Et	4′-Et	4.18	4.25	4.25
Q23	H	2′-Et	3′-Et	3.80	3.90	3.90
Q24	H	3′-Et	4′-Et	4.62	4.63	4.63
Q25	H	3′-Et	5′-Et	5.07	5.03	5.03
Q26	H	2′-^{i}Pr	6′-^{i}Pr	4.96	4.96	4.95
Q27	H	2′-Me	6′-Et	5.81	5.91	5.85
Q28	H	2′-Me	6′-^{i}Pr	5.58	5.59	5.61
Q29	H	2′-^{t}Bu	5′-^{t}Bu	3.54	3.53	3.51
Q30	H	3′-^{t}Bu	5′-^{t}Bu	4.68	4.66	4.61
Q31	H	2′,3′-Fused-Ph		5.23	5.29	5.24
Q32	H	2-Cl	6′-Cl	3.84	3.86	3.80
Q33	H	2-F	6′-F	3.44	3.44	3.50
Q34	7-NH_2	2-Et	6′-Et	4.58	4.51	4.61
Q35	7-NO_2	2-Et	6′-Et	4.08	4.08	4.06

[a] Compounds **34–37** in the original publication (Ref. [21]) are not included due to structural uncertainty.
[b] Expressed as the logarithm of $1/IC_{50}$ (M) value.

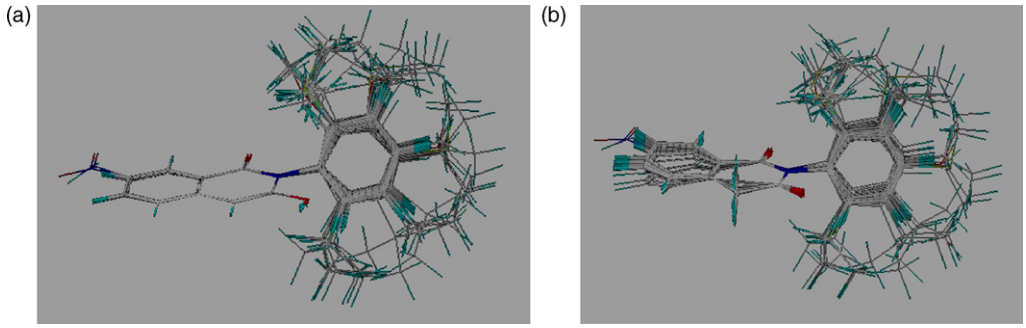


Fig. 4. Structural alignments of all molecules in the data set of PSA inhibitors: (a) enol tautomer; (b) keto tautomer.

Fig. 1(b)). In addition, there is an extra small red region lying between the two oxygen atoms of the β-diketone moiety, and moreover the blue region is not concentrated at the center of the phenyl ring as shown in Fig. 3(a), but extends, along the β-diketone structural unit to the center of the cyclohexanone fragment. In view of poor explanation for these extra contours, the exocyclic enol tautomer seems more likely to be the active form. As for the steric contour maps, it has been shown that there are two small sterically favorable (green) regions surrounding the 2′- and 6′-positions of the phenyl ring, respectively, both for the tautomers I and II (see Fig. 3(b) and (d)). The yellow regions distribute at above the phenyl ring and its 6′-substituent, and the main difference between the two tautomers is that only one large polyhedron has been observed in the exocyclic enol CoMFA contour map (Fig. 3(d)), while in the endocyclic enol one it split into two pieces (Fig. 3(b)). Thus, it can be concluded from above results that the tautomerism has significant effect on the descriptors of the electrostatic field, while the corresponding effect on those of the steric field is limited.

3.2. Data set of PSA inhibitors

A high affinity PSA inhibitor has the potential to serve as an analgesic or an inhibitor of tumor metastasis [37]. Thirty five compounds with a homophthalimide skeleton, together with their PSA inhibitory activities (pIC_{50}), which were taken from ref [21], are given in Table 3. A previous structure-activity relationship study by Komoda et al. [21] indicated that tautomerism of the imidobenzoylketone group in the cyclic imide moiety of the homophthalimide skeleton was important for the inhibitory activity. The authors speculated that the enol tautomer might be the active form in aminopeptidase inhibition, although the keto form is more stable than the enol one.

The CoMFA models have been developed for both tautomeric forms. All molecules were superimposed based on the alignment of homophthalimide skeleton, using compound **Q20** as the template, and the results are shown in Fig. 4(a) and (b), respectively for the enol and keto forms. As indicated, no significant difference was observed in the two aligned plots because the conformational changes resulted from the tautomerization in the rigid homophthalimide skeleton are almost negligible. Thus, it can be expected that the difference of the CoMFA results reported below stems primarily from the tautomerism. Table 4 presents a summary of the statistical results for the CoMFA models. It is noted that, AOS-APS improve remarkably the predictive power of the models, as indicated by the variance of q^2 and S.E.P. values. While the q^2 improves from 0.346 (NOPC = 8) to 0.740 (NOPC = 7) for the enol tautomers, and from 0.322 (NOPC = 8) to 0.726 (NOPC = 7) for the keto tautomers, the S.E.P. decreases by more than 0.1 for either form. The AOS-APS also slightly improved the conventional r^2, S.D. and F values for the enol form, and probably for the keto form if considering the fact that one less component was used.

As for the comparison of the CoMFA models between the two tautomeric forms, our results show that the enol form yields slightly better statistics (Table 4). This seems to support the conclusion derived form the structure–activity relationship study, although further evidences are necessarily required to confirm it. The predicted pIC_{50} values by the AOS-APS CoMFA models for both tautomers are listed in Table 3. The plots of the actual versus predicted activities, which provide a more detailed picture of the predictive quality of the models, are shown in Fig. 5. Both models yield excellent predictive power (with the residuals generally less than 0.2 logarithmic unit).

The contour maps of steric and electrostatic fields generated with the AOS-APS CoMFA models for both tautomeric forms

Table 4
Summary of CoMFA results for PSA inhibitors

	q^2	S.E.P.	Optimal NOPC	r^2	S.D.	F	Contribution % steric/electrostatic
Enol: conventional	0.346	0.667	8	0.989	0.088	285.341	61.8/38.2
AOS_APS	0.740	0.414	7	0.992	0.071	503.647	63.0/37.0
Keto: conventional	0.322	0.687	8	0.989	0.088	285.417	58.0/42.0
AOS_APS	0.726	0.425	7	0.986	0.095	279.880	50.1/49.9

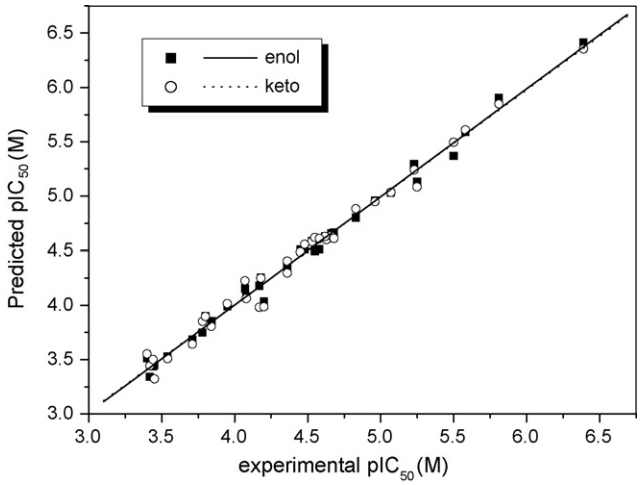


Fig. 5. Plots of the experimental versus predicted pIC_{50} values for the enol and keto CoMFA models of PSA inhibitors.

are depicted in Fig. 6. The most active compound **Q20** is also shown as the reference structure. Whereas the enol CoMFA model favors the contribution of steric field (S/E = 63.0/37.0), almost equal contributions are observed from steric (50.1%) and electrostatic (49.9%) fields in the keto CoMFA model. The steric field contours for both tautomeric forms, as expected, are similar to each other (cf. Fig. 6(b) and (d)) except that there is an additional sterically unfavorable (yellow) region around the 3′-position of the substituted phenyl ring for the keto form. The occurrence of such an additional contour is not reasonable if one noticed the fact that the compounds with more bulky 3′-substituent on the phenyl ring are generally less active (cf. the pIC_{50} values of the compound pairs **Q4/Q1**, **Q12/Q9**, **Q17/Q2** and **Q24/Q5**). On the contrary, the electrostatic field contours of the keto and enol CoMFA models differ significantly form each other. In the enol contour map (Fig. 6(a)), the 5′-substituent of the phenyl ring was found to orient in the positive charge favorable (blue) contour, and the red region was localized around the 6′-substituent. Whereas in the keto one (Fig. 6(c)), only three very small polyhedra, with the two blue ones located respectively at the 2′- and 6′-substituents of the phenyl ring and the red one at the 3′-substituent, have been observed. The remarkable effect of the tautomerism upon the electrostatic field is further evidenced though it is difficult to determine which contour map appears more plausible based on the available data.

3.3. Data set of anxiolytic agents

A group of 67 pyrido[1,2-*a*]benzimidazole (PBI) derivatives and their biological activities (pIC_{50} values measured in the absence of γ-aminobutyric acid, GABA), which were taken from the published results [22–24], are listed in Table 5. This chemical series represents a novel class of potential anxiolytic agents due to their high affinities toward the benzodiazepine binding site on $GABA_A$ receptors. The first compound of the PBI structural type (the ethyl ester of PBI) was reported by Ohta et al. [38] in 1991, and the enol tautomer was assigned in that study as the predominant form through chemical analysis. However, a recent study [39] using NMR spectroscopy and theoretical calculations indicated that this three-membered heterocylic system exists almost exclusively in the keto form (Table 5).

The tricyclic skeletons were used for alignment and the most active compound **P21** was selected as the template molecule.

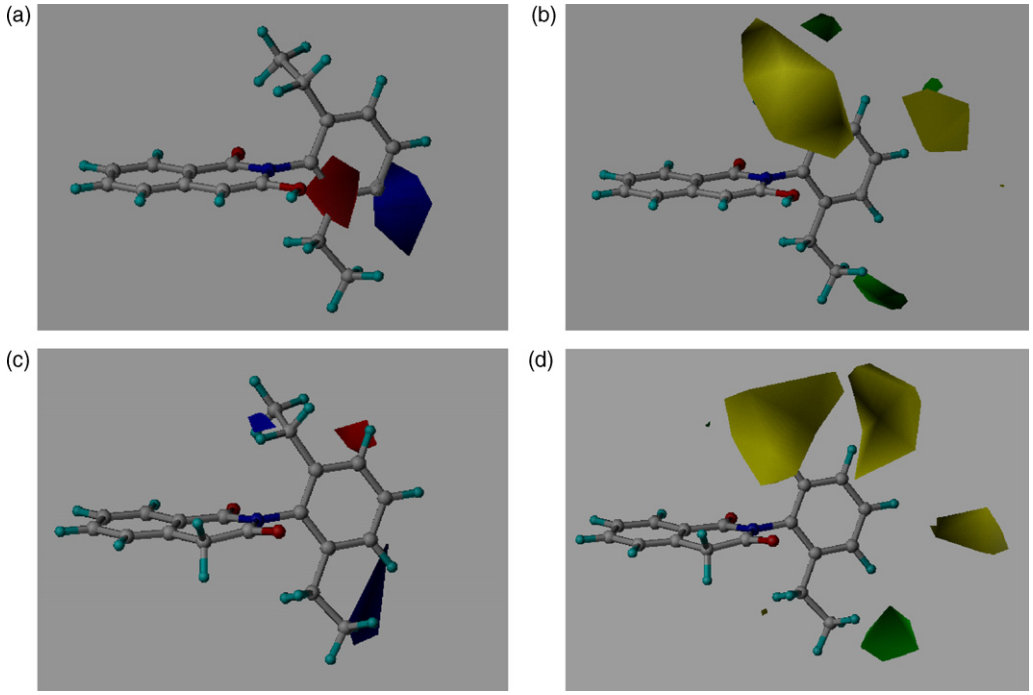
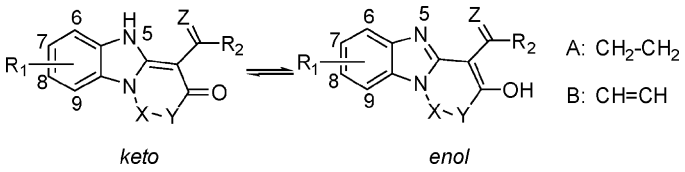


Fig. 6. (a) Electrostatic field and (b) steric field contours generated with the enol AOS-APS CoMFA model for PSA inhibitors. (c) Electrostatic field and (d) steric field contours generated with the keto AOS-APS CoMFA model for PSA inhibitors.

Table 5
Structures and activities of anxiolytic agents



	X–Y	Z	R_1	R_2	pIC_{50} (M)[a] exptl	pIC_{50} (M)[a] calcd (enol)	pIC_{50} (M)[a] calcd (keto)
P1	A	O		NH-Ph	8.04	7.54	7.34
P2	A	O		NH-(4-Cl-Ph)	6.21	6.46	6.52
P3	A	O		NH-(3-Cl-Ph)	6.92	7.06	7.09
P4	A	O		NH-(2-Cl-Ph)	7.92	7.89	7.88
P5	A	O		NH-(2-F-Ph)	8.77	8.22	9.17
P6	A	O		NH-(4-MeO-Ph)	7.39	6.85	7.40
P7	A	O		NH-(3-MeO-Ph)	7.59	7.86	7.53
P8	A	O		NH-(2-MeO-Ph)	6.00	6.02	5.94
P9	A	O		NH-(2,6-Cl_2-Ph)	6.49	7.38	6.58
P10	A	O		NH-(2,6-F_2-Ph)	8.55	8.42	8.46
P11	A	O		NH-(*c*-C_6H_{11})	6.85	6.76	6.58
P12	A	O		NH-(*c*-C_4H_7)	7.52	7.83	7.70
P13	A	O		NH-(*c*-C_3H_5)	7.80	7.78	7.79
P14	A	O		NH_2	5.62	5.46	5.60
P15	A	O		OEt	6.17	6.13	6.28
P16	A	S		NH-Ph	6.44	6.42	6.52
P17	A	O		N(Me)-Ph	5.00	4.95	5.07
P18	A	O		O-Ph	6.70	6.97	6.52
P19	A	O		S-Ph	7.54	7.18	7.84
P20	B	O		NH-Ph	7.62	7.97	7.77
P21	B	O		NH-(2-F-Ph)	9.64	8.86	9.26
P22	A	O		NH-(4-COOH-Ph)	5.00	5.09	4.86
P23	A	O		NH-(4-OH-Ph)	7.34	7.07	7.21
P24	A	O		NH-(4-NH_2-Ph)	4.89	5.84	5.22
P25	A	O		NH-(2-NH_2-Ph)	7.30	6.95	7.21
P26	A	O		NH-(4-NMe_2-Ph)	6.57	6.34	6.23
P27	A	O		NH-(3-NMe_2-Ph)	7.43	7.21	7.55
P28	A	O		NH-(2-NMe_2-Ph)	5.00	4.74	5.04
P29	A	O		NH-(2-F-4-NMe_2-Ph)	6.92	7.01	6.98
P30	A	O		NH-(2-Me-4-NMe_2-Ph)	5.40	5.32	5.21
P31	A	O		NH-(4-pyridinyl)	6.80	7.04	6.96
P32	A	O		NH-(3-pyridinyl)	7.29	7.38	7.45
P33	A	O		NH-(2-pyridinyl)	7.23	7.76	7.59
P34	A	O		NH-(4-(3-Cl-pyridinyl))	6.66	6.33	6.51
P35	A	O		NH-(4-(3-Me-pyridinyl))	5.15	5.34	5.05
P36	A	O		NH-(4-(3-F-pyridinyl))	7.28	6.78	7.34
P37	A	O		NH-(3-(2-Cl-pyridinyl))	7.17	7.78	6.99
P38	A	O		NH-CH_2-(4-pyridinyl)	7.19	7.22	7.23
P39	A	O		NH-CH_2-(3-pyridinyl)	6.82	6.95	6.60
P40	A	O		NH-(2-pyrazinyl)	6.44	6.88	6.53
P41	A	O		NH-(2-pyrimidinyl)	7.27	6.99	7.19
P42	A	O		NH-(4-pyrimidinyl)	6.32	6.24	6.35
P43	B	O		NH-(4-pyridinyl)	6.59	7.45	6.73
P44	A	O	6-OH	NH-(2-F-Ph)	7.98	8.04	7.89
P45	A	O	7-OH	NH-Ph	7.39	7.32	7.48
P46	A	O	9-OH	NH-(2-F-Ph)	7.85	7.99	7.88
P47	A	O	6-OMe	NH-(2-F-Ph)	8.11	7.13	7.77
P48	A	O	6-OMe	NH-(2,6-F_2-Ph)	6.81	7.39	7.20
P49	A	O	7-OMe	NH-Ph	7.40	7.22	7.25
P50	A	O	7-OMe	NH-(2-F-Ph)	8.21	7.98	8.19
P51	A	O	6-Me	NH-Ph	8.00	7.13	7.84
P52	A	O	6-Me	NH-(2-F-Ph)	7.02	7.74	7.53
P53	A	O	7-CF_3	NH-(2-F-Ph)	6.74	6.91	6.82
P54	A	O	6-Cl	NH-(2-F-Ph)	8.15	8.51	8.25
P55	A	O	6-Cl	NH-(2,6-F_2-Ph)	8.96	8.66	8.68
P56	A	O	7-Cl	NH-Ph	7.18	6.96	7.28

Table 5 (*Continued*)

	X–Y	Z	R_1	R_2	pIC_{50} (M)[a] exptl	pIC_{50} (M)[a] calcd (enol)	pIC_{50} (M)[a] calcd (keto)
P57	A	O	7-Cl	NH-(2-F-Ph)	8.29	7.91	7.92
P58	A	O	7-Cl	NH-(2,6-F_2-Ph)	8.20	8.15	8.32
P59	A	O	8-Cl	NH-Ph	6.76	6.60	6.74
P60	A	O	7-F	NH-(2-F-Ph)	8.72	8.17	8.32
P61	A	O	7-F	NH-(2,6-F_2-Ph)	8.20	8.49	8.66
P62	A	O	6,7-F_2	NH-(2-F-Ph)	8.85	8.71	8.71
P63	A	O	6,7-F_2	NH-(2,6-F_2-Ph)	8.96	9.03	8.98
P64	A	O	6,8-F_2	NH-(2-F-Ph)	7.70	7.84	7.56
P65	A	O	6,8-F_2	NH-(2,6-F_2-Ph)	7.52	8.03	7.64
P66	A	O	8,9-F_2	NH-(2-F-Ph)	6.07	6.64	6.50
P67	A	O	8,9-F_2	NH-(2,6-F_2-Ph)	7.18	6.84	6.79

[a] Expressed as the logarithm of $1/IC_{50}$ (M) value.

Table 6
Summary of CoMFA results for anxiolytic agents

	q^2	S.E.P.	Optimal NOPC	r^2	S.D.	F	Contribution % steric/electrostatic
Enol: conventional	0.227	0.920	1	0.370	0.831	38.215	40.3/59.7
AOS_APS	0.367	1.013	5	0.851	0.417	69.858	55.2/44.8
Keto: conventional	0.281	0.909	4	0.790	0.491	58.371	46.4/53.6
AOS_APS	0.536	0.755	8	0.951	0.246	139.875	52.7/47.3

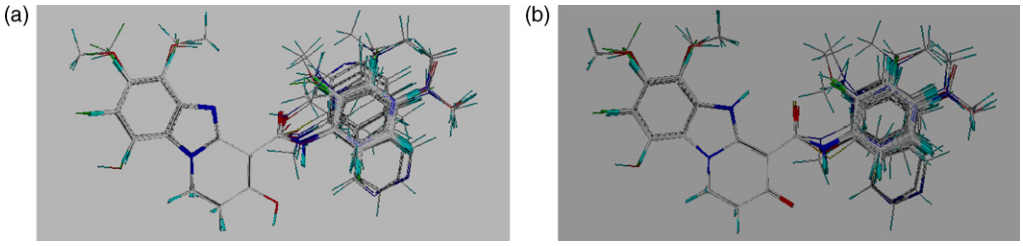


Fig. 7. Structural alignments of all molecules in the data set of anxiolytic agents: (a) enol tautomer; (b) keto tautomer.

The superimposed graphs of all 67 molecules for the enol and keto forms are illustrated in Fig. 7(a) and (b), respectively. Again, the structural alignments exhibit great similarity, which may minimize, to a great extent, the effect of conformation on the difference of the CoMFA results between the two tautomeric forms.

Table 6 collects the statistical results of the conventional and AOS-APS CoMFA models of both tautomers. As one can see, the quality of the CoMFA models is far less satisfactory than that for the HPPD and PSA inhibitors. The conventional CoMFA models produce low q^2 value, 0.227 for the enol tautomer and 0.281 for the keto one, both of which did not even exceed the generally accepted criterion, 0.3 [40]. For the enol tautomer, only one ONPC was given by the model even though the initial number of components increased to 20. AOS-APS improves significantly the q^2 values to 0.367 for the enol tautomer and to 0.536 for the keto tautomer. Also improved considerably are the conventional r^2 and S.D. values. The former increases from 0.370 to 0.851, and from 0.790 to 0.951, for the enol and keto form, respectively, whereas the latter is reduced from 0.831 to 0.417 (enol form), and from 0.491 to 0.246 (keto form). The contribution ratio of the steric versus electrostatic fields is 40.3/59.7 (enol form) and 46.4/53.6 (keto form) with respect to the conventional CoMFA models, and the importance of steric component is slightly enhanced after AOS-APS (steric/electrostatic = 55.2/44.8 and 52.7/47.3, respectively for the enol and keto model).

The predicted biological activities by both the enol and keto CoMFA models are given in Table 5. Fig. 8 depicts graphically the

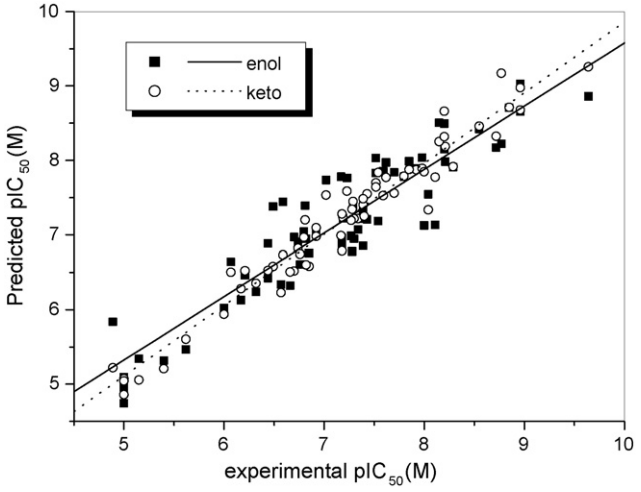


Fig. 8. Plots of the experimental versus predicted pIC_{50} values for the enol and keto CoMFA models of anxiolytic agents.

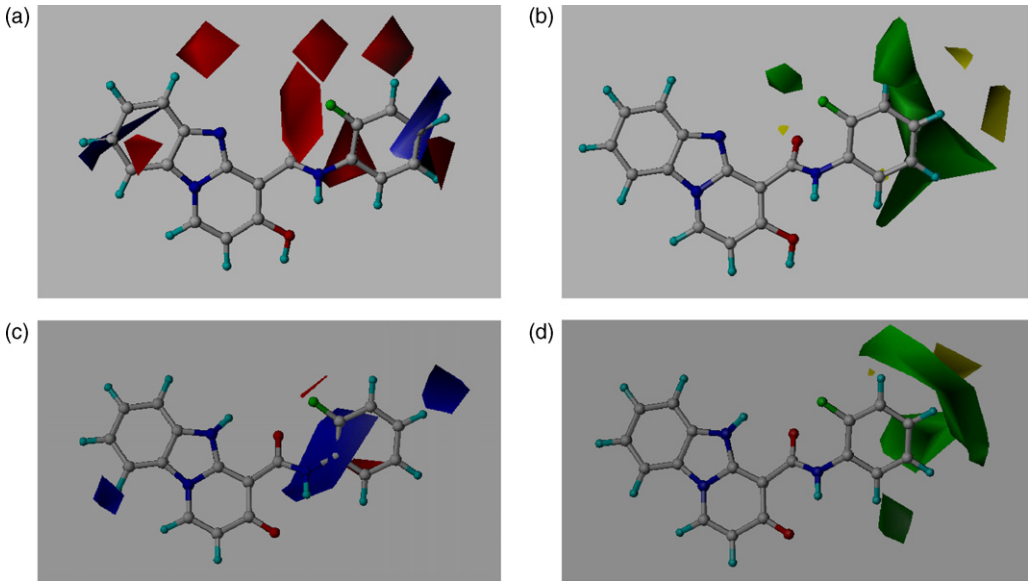


Fig. 9. (a) Electrostatic field and (b) steric field contours generated with the enol AOS-APS CoMFA model for anxiolytic agents. (c) Electrostatic field and (d) steric field contours generated with the keto AOS-APS CoMFA model for anxiolytic agents.

plots of the predicted versus measured activities. This, together with the comparison of the statistical quantities listed in Table 6 (especially the cross-validated q^2) between the two tautomeric forms indicate clearly that the keto CoMFA model represents measurably higher predictive ability than the corresponding enol one. In this sense, the keto form of the PBI-type compounds is more likely to be adopted when interacting with the $GABA_A$ receptor. Indeed, a later study [41] conducted by the same group revealed that substituting other groups for H at the N5-site (see Table 5) of the PBI derivatives, which prevents the formation of enol tautomer, did not result in a reduced activity.

Fig. 9 illustrates the contour maps of steric and electrostatic fields for both tautomeric forms generated with the AOS-APS CoMFA models. Compound **P21** is also shown as the reference structure. Due to relative low quality of the models, these underlying coefficient maps seemingly fail to provide a satisfactory interpretation for the known structure–activity relationships. Nevertheless, it has once again been shown definitively that there exists a distinct difference in the electrostatic field contour map between the two different tautomeric forms, whereas the steric field contour maps represent, to a great extent, similar spatial orientations.

4. Summary and conclusion

3D-QSAR studies on three selected systems, which are *p*-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors, inhibitors of puromycin-sensitive aminopeptidase (PSA), and anxiolytic agents have been conducted using the CoMFA methodology. Particularly, the effect of tautomerism upon the CoMFA results has been examined. For each system, the structural alignments using different tautomeric forms show no significant difference except the atoms involved in tautomerism, which ensures that the differences of the CoMFA results result primarily from the tautomerism. In the case of the data set of HPPD inhibitors, excellent AOS-APS CoMFA models ($q^2 > 0.8$ with four components) have been obtained, whether using the endocyclic enol or the exocyclic enol forms. Good results ($q^2 > 0.7$ with seven components) have also been achieved when treating the data set of PSA inhibitors. As for the comparisons of the CoMFA results between the two tautomers considered, both systems give almost the same statistical quality. However, a similar analysis using the data set of 67 anxiolytic agents results in quite different consequence for different tautomers. The CoMFA model of the keto tautomer, which is more likely to be the active form of the PBI type inhibitors, represents reasonably higher predictive ability ($q^2 = 0.54$ with eight components) than that of the corresponding enol tautomer ($q^2 = 0.37$ with five components). Furthermore, it has been shown definitively for all three systems that there exists a substantial difference in the electrostatic field contour maps between the two different tautomeric forms, whereas the steric field contours have, to a great degree, similar spatial distribution. This implies that the resulting new potent ligands may be quite different when using the CoMFA models of different tautomeric forms to guide further structural modification.

Acknowledgements

The authors are grateful to the Natural Science Foundation of Zhejiang (no. Y406042) and PhD Fund of Ningbo (no. 2004A610010) for financial support.

References

[1] R.D. Cramer III, D.E. Patterson, J.D. Bunce, Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steriods to carrier proteins, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[2] U. Norinder, Recent progress in CoMFA methodology and related techniques, Perspect. Drug Discov. Des. 12–14 (1998) 25–39.

[3] H. Kubinyi, QSAR and 3D QSAR in drug design. Part 2. Applications and problems, Drug Discov. Today 2 (1997) 538–546.
[4] Y.C. Martin, 3D QSAR: current state, scope and limitations, Perspect. Drug Discov. Des. 12–14 (1998) 3–23.
[5] K.H. Kim, G. Greco, E. Novellino, A critical review of the recent CoMFA applications, Perspect. Drug Discov. Des. 12–14 (1998) 257–315.
[6] S. Khanna, M.E. Sobhia, P.V. Bharatam, Additivity of molecular fields: CoMFA study on dual activators of $PPAR_{\alpha}$ and $PPAR_{\gamma}$, J. Med. Chem. 48 (2005) 3015–3025.
[7] M. Arakawa, K. Hasegawa, K. Funatsu, Application of the novel molecular alignment method using the Hopfield neural network to 3D-QSAR, J. Chem. Inf. Comput. Sci. 43 (2003) 1396–1402.
[8] S.J. Cho, A. Tropsha, M. Suffness, Y.-C. Cheng, K.-H. Lee, Antitumor agents. 163. Three-dimensional quantitative structure-activity relationship study of 4′-*o*-demethylepipodophyllotoxin analogs using the modified CoMFA/q^2-GRS approach, J. Med. Chem. 39 (1996) 1383–1395.
[9] S.V. Trepalin, A.V. Skorenko, K.V. Balakin, A.F. Nasonov, S.A. Lang, A. Andrey, A.A. Ivashchenko, N.P. Savchuk, Advanced exact structure searching in large databases of chemical compounds, J. Chem. Inf. Comput. Sci. 43 (2003) 852–860.
[10] P. Pospisil, P. Ballmer, L. Scapozza, G. Folkers, Tautomerism in computer-aided drug design, J. Recept. Signal Transduct. 23 (2003) 361–371.
[11] For a recent example, see: M.L. López-Rodríguez, M. Murcia, B. Benhamú, A. Viso, M. Campillo, L. Pardo, Benzimidazole derivatives. 3. 3D-QSAR/CoMFA model and computational simulation for the recognition of 5-HT_4 receptor antagonists, J. Med. Chem. 45 (2002) 4806–4815.
[12] Z. Rapport, The Chemistry of Enols, Wiley, Chichester, U.K, 1990.
[13] B. Hernandez, F.J. Luque, M. Orozco, Tautoermism of xanthine oxidase substrates hypoxanthine and allopurinol, J. Org. Chem. 61 (1996) 5964–5971.
[14] H. Brandstetter, F. Grams, D. Glitz, A. Lang, R. Huber, W. Bode, H.W. Krell, R.A. Engh, The 1.8-Å crystal structure of a matrix metalloproteinase-8 barbiturate inhibitor complex reveals a previously unobserved mechanism for collagenase substrate recognition, J. Biol. Chem. 276 (2001) 17405–17412.
[15] Y.-L. Lin, C.-S. Wu, S.-W. Lin, D.-Y. Yang, SAR studies of 2-*o*-substituted-benzoyl- and 2-alkanoyl-cyclohexane-1,3-diones as inhibitors of 4-hydroxyphenylpyruvate dioxygenase, Bioorg. Med. Chem. Lett. 10 (2000) 843–845.
[16] C.-S. Wu, J.-L. Huang, Y.-S. Sun, D.-Y. Yang, Mode of action of 4-hydroxyphenylpyruvate dioxygenase inhibition by triketone-type inhibitors, J. Med. Chem. 45 (2002) 2222–2228.
[17] S.-W. Lin, Y.-L. Lin, T.-C. Lin, D.-Y. Yang, Discovery of a potent, non-triketone type inhibitor of 4-hydroxyphenylpyruvate dioxygenase, Bioorg. Med. Chem. Lett. 10 (2000) 1297–1298.
[18] D.L. Lee, C.G. Knudsen, W.J. Michaely, H.L. Chin, N.H. Nguyen, C.G. Carter, T.H. Cromartie, B.H. Lake, J.M. Shribbs, T. Fraser, The structure–activity relationships of the triketone class of HPPD herbicides, Pestic. Sci. 54 (1998) 377–384.
[19] G. Mitchell, D.W. Bartlett, T.E.M. Fraser, T.R. Hawkes, D.C. Holt, J.K. Townson, R.A. Wichert, Mesotrione: a new selective herbicide for use in maize, Pest Manag. Sci. 57 (2001) 120–128.
[20] M.K. Ellis, A.C. Whitfield, L.A. Gowans, T.R. Auton, W. McLean Provan, E.A. Lock, D.L. Lee, L.L. Smith, Characterization of the interaction of 2-[2-nitro-4-(trifluoromethyl)benzoyl]-4,4,6,6-tetramethylcyclohexane-1,3,5-trione with rat hepatic 4-hydroxyphenylpyruvate dioxygenase, Chem. Res. Toxicol. 9 (1996) 24–27.
[21] M. Komoda, H. Kakuta, H. Takahashi, Y. Fujimoto, S. Kadoya, F. Kato, Y. Hashimoto, Specific inhibitor of puromycin-sensitive aminopeptidase with a homophthalimide skeleton: identification of the target molecule and a structure–activity relationship study, Bioorg. Med. Chem. 9 (2001) 121–131.
[22] B.E. Maryanoff, W. Ho, D.F. McComsey, A.B. Reitz, P.P. Grous, S.O. Nortey, R.P. Shank, B. Dubinsky, R.J. Taylor, J.F. Gardocki, Potential anxiolytic agents. Pyrido[1,2-*a*]benzimidazoles: a new structural class of ligands for the benzodiazepine binding site on GABA-A receptors, J. Med. Chem. 38 (1995) 16–20.
[23] B.E. Maryanoff, D.F. McComsey, W. Ho, R.P. Shank, B. Dubinsky, Potential anxiolytic agents. 2. Improvement of oral efficacy for the pyrido[1,2-*a*]benzimidazole (PBI) class of GABA-A receptor modulators, Bioorg. Med. Chem. Lett. 6 (1996) 333–338.
[24] B.E. Maryanoff, S.O. Nortey, J.J. McNally, P.J. Sanfilippo, D.F. McComsey, B. Dubinsky, R.P. Shank, A.B. Reitz, Potential anxiolytic agents. 3. Novel A-ring modified pyrido[1,2-*a*]benzimidazoles, Bioorg. Med. Chem. Lett. 9 (1999) 1547–1552.
[25] A. Pedretti, L. Villa, G. Vistoli, VEGA: a versatile program to convert, handle and visualize molecular structure on Windows-based PCs, J. Mol. Graph. 21 (2002) 47–49.
[26] SYBYL Version 6.8, Tripos Associates, St. Louis, MO, 2001.
[27] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, Tetrahedron 36 (1980) 3219–3228.
[28] A. Streitwieser, Molecular Orbital Theory for Organic Chemists, Wiley, New York, 1961.
[29] R. Wang, Y. Gao, L. Liu, L. Lai, All-orientation search and all-placement search in comparative molecular field analysis, J. Mol. Model. 4 (1998) 276–283.
[30] D.-Y. Yang, 4-Hydroxyphenylpyruvate dioxygenase as a drug discovery target, Drug News Perspect. 16 (2003) 493–496.
[31] G.R. Moran, 4-Hydroxyphenylpyruvate dioxygenase, Arch. Biochem. Biophys. 433 (2005) 117–128.
[32] M.-L. Huang, J.-W. Zou, D.-Y. Yang, B.-Z. Ning, Z.-C. Shang, Q.-S. Yu, Theoretical studies on tautomerism of benzoylcyclohexane-1,3-dione and its derivatives, J. Mol. Struct. (THEOCHEM) 589–590 (2002) 321–328.
[33] E.V. Borisov, A.I. Verenich, A.A. Govorova, A.S. Lyakhov, T.S. Khlebnikova, The structure of benzoyldimedone from X-ray diffraction analysis, Russ. Chem. Bull. 49 (2000) 1068–1070.
[34] N.N. Shapetko, Y.S. Bogachev, J.L.D.N. Radushnovaand Shigorin, Study of the type of proton potential energy function in strong intramolecular hydrogen bonds by proton, deuterium-2, and carbon-13 NMR and MO LCAO methods in the CNDO/2 approximation, Dokl. Akad. Nauk SSSR. 231 (1976) 409–412.
[35] M. Kavana, G.R. Moran, Interaction of (4-hydroxyphenyl)pyruvate dioxygenase with the specific inhibitor 2-[2-nitro-4-(trifluoromethyl)benzoyl]-1,3-cyclohexanedione, Biochemistry 42 (2003) 10238–10245.
[36] J.M. Brownlee, K. Johnson-Winters, D.H.T. Harrison, G.R. Moran, Structure of the ferrous form of (4-hydroxyphenyl)pyruvate dioxygenase from *Streptomyces avermitilis* in complex with the therapeutic herbicide, NTBC, Biochemistry 43 (2004) 6370–6377.
[37] H. Fujii, M. Nakajima, T. Aoyagi, T. Tsuruo, Inhibition of tumor cell invasion and matrix degradation by aminopeptidase inhibitors, Biol. Pharm. Bull. 19 (1996) 6–10.
[38] S. Ohta, T. Yuasa, Y. Narita, I. Kawasaki, E. Minamii, M. Yamashita, Synthesis and application of imidazole derivatives. Synthesis of pyrido[1,2-*a*]benzimidazolone derivatives, Heterocycles 32 (1991) 1923–1931.
[39] A.B. Reitz, D.A. Gauthier, W. Ho, B.E. Maryanoff, Tautomerism and physical properties of pyrido[1,2-*a*]benzimidazole (PBI) GABA-A receptor ligands, Tetrahedron 56 (2000) 8809–8812.
[40] M. Clark, R.D. Cramer III, The probability of chance correlation using partial least squares (PLS), Quant. Struct. Act. Relat. 12 (1993) 137–145.
[41] A.D. Jordan, A.H. Vaidya, D.I. Rosenthal, B. Dubinsky, C.P. Kordik, P.J. Sanfilippo, W.-N. Wu, A.B. Reitz, Potential anxiolytic agents. Part 4. Novel orally-active N^5-substituted pyrido[1,2-*a*]benzimidazoles with high GABA-A receptor affinity, Bioorg. Med. Chem. Lett. 12 (2002) 2381–2386.