Journal of Computer-Aided Molecular Design, **14:** 277–291, 2000.
KLUWER/ESCOM

Chemometric rationalization of the structural and physicochemical basis for selective cyclooxygenase-2 inhibition: Toward more specific ligands

Enrica Filipponi[a], Violetta Cecchetti[a], Oriana Tabarrini[a], Daniela Bonelli[b] & Arnaldo Fravolini[a,*]

[a]*Istituto di Chimica e Tecnologia del Farmaco, Università di Perugia, via del Liceo 1, I-06123 Perugia, Italy;* [b]*Dipartimento di Chimica, Università di Perugia, via Elce di Sotto 9, I-06123 Perugia, Italy*

Received 6 May 1999; Accepted 13 October 1999

Key words: cyclooxygenase-2, GRID characterization, principal component analysis (PCA), selectivity

Summary

The discovery that proinflammatory prostaglandins are produced by cyclooxygenase-2 (COX-2), an inducible isoform of the constitutive cyclooxygenase-1 (COX-1), opened a new frontier in the treatment of inflammatory diseases, because the selective inhibition of COX-2 can lead to therapeutically effective compounds which do not have the common side effects of classical non-steroidal antiinflammatory drugs (NSAIDs). Different crystallographic structures of both free COX-1 and COX-2 as well as complexes with inhibitors have been solved. Because of the great similarity between the two enzymes, it is difficult to detect the most important structural and physicochemical features that would be useful for designing inhibitors with an improved selectivity. In this paper we describe the application of a chemometric procedure to the study of COX-2 selective inhibition. This method, developed to reveal the most suitable regions of isoenzymes for the design of selective ligands, also has a very practical utility. GRID multivariate characterization of the enzymes and subsequent Principal Component Analysis (PCA) of the descriptor variables allow the identification of chemical groups that could be added to a core template structure to increase ligand selectivity.

Introduction

Non-steroidal antiinflammatory drugs (NSAIDs) are used in the treatment of inflammatory diseases and to alleviate fever and pain. Their activity is due to the inhibition of cyclooxygenase (COX), the enzyme involved in the biosynthesis of pro-inflammatory prostaglandins from precursor arachidonic acid [1].

COX is a bifunctional enzyme. The initial cyclooxygenase reaction, the target of NSAIDs, involves bis-oxygenation of arachidonic acid to PGG_2, while the subsequent peroxidase reaction reduces PGG_2 to PGH_2.

Two isoforms of COX enzyme exist [2]: a constitutive cyclooxygenase-1 (COX-1) and an inducible cyclooxygenase-2 (COX-2). Both COX-1 and COX-2 are homodimeric glycosylated integral membrane proteins, primarily located in the endoplasmic reticulum. The COX-1 enzyme is responsible for maintaining homeostasis (gastric and renal integrity), whereas the cytokine-inducible enzyme COX-2 [3] induces an inflammatory condition [4] in response to inflammatory and mitogenic stimuli, suggesting that COX-1 and COX-2 serve different physiological and pathological functions.

These findings led to the hypothesis that the gastrointestinal and renal toxicity observed with the use of NSAIDs is due to an inhibition of COX-1, while the desired antiinflammatory activity is mediated by inhibition of COX-2. Therefore, a selective inhibition of COX-2 would provide a superior safety profile [5], opening a new era of research in the treatment of inflammatory diseases.

*To whom correspondence should be addressed. E-mail: fravo@unipg.it

Recent reports [6] suggest that the delineation of the roles of COX-1 and COX-2 might not be quite as clear as has been proposed and there are distinct circumstances in which COX-2-derived prostaglandins play an important role in the maintenance of gastrointestinal mucosal integrity. Moreover, the COX-2 inhibitors seem to be therapeutically effective only at doses that also suppress prostaglandin synthesis by COX-1. Even if these studies predict significant limitations of COX-2 inhibitors concerning efficacy and toxicity, it is still possible that these agents will be beneficial for the treatment of some inflammatory conditions.

X-ray structures of ovine [7] COX-1, murine [8] COX-2 and human [9] COX-2 have been solved, devoid of ligands or in complex with different inhibitors. The two enzymes have a sequence identity of about 60% and the overall structures are highly conserved, being practically identical at the catalytic domain. Only one amino acid differentiates the core of the NSAID binding site (Ile523 in COX-1, Val523 in COX-2) and a few other different residues are present in the region around the active site.

Because COX-2 selective inhibitors do not necessarily interact with residues which are unique to COX-2, as evidenced by selective derivatives bound to the human cyclooxygenase active site, and because biochemical and kinetic factors hidden in the X-ray structures could be involved in selective inhibition, *de novo* design or optimization of COX-2 selective inhibitors are not easy tasks.

However, the few structural differences observed play an important role in selective COX-2 inhibition. For example, the NSAID binding site of COX-2 is much larger because of the difference at the 523 residue position. In fact, in COX-2 the smaller Val523 residue allows entry into a second pocket, branching from the main binding site cavity, which is much narrower and inaccessible in COX-1 because of the bulkier Ile523.

Site-directed mutagenesis studies [10, 11] of the active site of COX-2 showed that this simple difference could explain the COX-2 selective activity of the four agents NS-398 [12], nimesulide [13], DuP-697 [14] and SC-58125 [15] (Figure 1).

The X-ray structure of murine COX-2 [8] shows that selectivity of the co-crystallized SC-558 derivative (Figure 1) seems to result from the benzensulfonamide moiety, which binds in the second pocket of COX-2. It is interesting to recall that celecoxib [16] (Figure 1), the first 1,5-diarylpirazole derivative in clinical use as COX-2 selective inhibitor, also belongs to this class of benzensulfonamide derivatives.

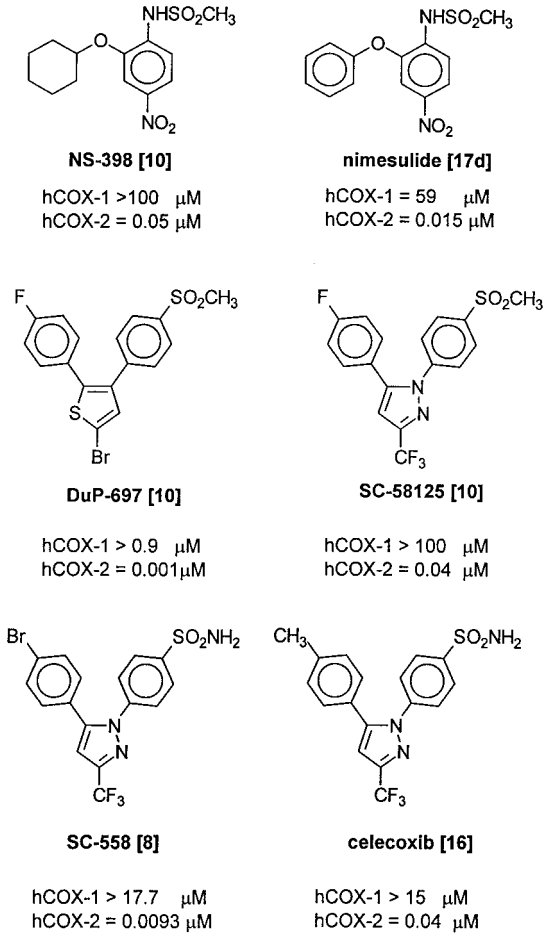

Figure 1. Selective COX-2 inhibitors.

In order to identify other groups capable of selective interaction and to design compounds with improved selectivity, it is important to characterise any other difference between NSAID binding sites of COX-1 and COX-2 in terms of structural and energetic properties, which may influence the intermolecular interaction with a ligand.

Recent attempts based on different procedures have been made to elucidate the structural basis for the mechanism of action of non selective inhibitors and COX-2 selective derivatives [17]. Our strategy follows the methodology already used to clarify other selectivity problems [18]. The approach is based on a multivariate characterisation of the ligand-biomolecule interaction site by means of the GRID program [19, 20]. The information contained in the GRID field variables is then rationalised using Principal Component Analy-

sis (PCA) [21] which highlights the variables which are most related to the factors responsible for selective COX-2 interaction.

Materials and methods

Selection of the COX structures and superimposition of their binding sites

In order to model the two COX enzymes, the two structures *1prh*[a] [6] PDB [22] entry for ovine COX-1 (resolution 3.5 Å) and *5cox* [8] PDB entry for murine COX-2 (resolution 3 Å) were used.

Sequence analysis suggests that the cyclooxygenase active sites for human COX-1 and COX-2 enzymes are very similar to the ovine COX-1 active site. Therefore, ovine COX-1 can be considered as a template of human COX-1 enzyme.

Murine COX-2 was used in our study because, although the human enzyme has been solved [9], the X-ray coordinates are not available in the PDB database. The structure of human COX-2 is expected to be very similar to the structure of the murine enzyme, since the two enzymes have an 87% identity with strict sequence conservation in the active site.

The structure of the COX enzymes is made up of three distinct domains: an N-terminal epidermal growth factor domain (EGF), a membrane-binding motif and a C-terminal catalytic domain, which contains the active cyclooxygenase and peroxidase sites.

Pastor et al. [18b] suggest that for this kind of analysis it would be preferable to use the crystallographic structure as they appear in binary complexes with substrate or inhibitors, because the conformation of the enzymes in bound structures is generally more similar to the conformation adopted during ligand interaction. Moreover, the ligand structures could guide the superimposition of the two enzymes at the binding site.

In our particular study, the uncomplexd COX-2 X-ray structure was chosen because the crystallographic structures show that the coordinates of the uncomplexed COX-2 enzyme are generally very similar to those of complexes with inhibitors [9]. There are probably no large-scale conformational changes at the ligand binding site, even though dynamic modifications could be involved in the entry and exit processes.

Moreover, the inhibitors bound to COX-1 and the selective ligands of COX-2 are quite different from a chemical point of view and there is not an obvious superimposition criterion for their structures, which would facilitate the superimposition of the two enzymes. Conversely, the catalytic domains of COX-1 and COX-2 are practically identical and make the amino acidic superimposition easier.

Water molecules were removed from the PDB file and both structures were imported into the SYBYL [23] molecular modelling program. The two structures were superimposed by considering only the backbone of α-carbon atoms of one monomer in the core of the catalytic domain, as defined in the literature [9]. In this way, only the NSAID binding domains of COX-1 and COX-2 were compared, omitting the rest of the enzyme, which is less relevant to our study.

Characterization of the NSAID binding site by means of GRID force field

GRID [19, 20] is a computational procedure originally developed to identify regions of energetically favourable interactions between a molecule and a probe. The interaction energies (E_{xyz}) between the probe and the atoms of the target molecule are calculated at each node of a three-dimensional grid (in which the target molecule is ideally placed) in terms of electrostatic (El), hydrogen-bond (HB) and Lennard-Jones (LJ) potentials. In each grid point of coordinates xyz, the energy value is calculated as indicated in Equation 1

$$E_{xyz} = \sum E_{\mathrm{El}} + \sum E_{\mathrm{HB}} + \sum E_{\mathrm{LJ}} \qquad (1)$$

where the sum is extended to every constitutive atom of the target molecules.

The charges on X-ray structures were neutralised prior to computing GRID maps by the MINIM [20] and FILMAP [20] programmes. These can be used to find the energy minima in a map calculated on all the atoms of the target structures, by using an appropriate counterion as probe, and populate them in the most energetically favourable way, respectively.

GRID allows the computation to be restricted to a selected region of the target molecule. In our study a $25 \times 18 \times 22$ Å grid cage (grid spacing 1 Å) was used, almost centred at the Ser353 residue and containing about 50 residues. The grid axes were chosen so that all the residues of the catalytic region were included.

GRID allows different types of probes to be used, depending upon the chemical information desired. Because a detailed characterisation of the binding site was desirable, the 42 probes listed in Table 1 were used in both the superimposed COX-1 and COX-2 tar-

Table 1. GRID probes used

Probe code	Probe type	Probe code	Probe type
N3+	sp_3 amine NH_3 cation	O=	O sulphate/sulphonamide
N2:	sp_3 NH_2 with lone pair	O1	alkyl hydroxy OH group
N2	neutral flat NH_2 amide	OES	sp_3 ester oxygen atom
N1:	sp_3 NH with lone pair	OS	O sulphone/sulphoxide
N1	neutral flat NH amide	OFU	furane oxygen
N1#	sp NH with one hydrogen	PO4	phosphate dianion
N:=	sp_2 N with lone pair	S1	neutral SH group
N-:	anionic tetrazole N	F	organic fluorine atom
O-	sp_2 phenolate oxygen	Cl	organic chlorine atom
OH	phenol or carboxy OH	Br	organic bromine atom
OC2	ether or furan oxygen	I	organic iodine atom
ON	oxygen of nitro group	COO-	aliphatic carboxylate
N2+	sp_3 amine NH_2 cation	Ar-COO-	aromatic carboxylate
N2=	sp_2 amine NH_2 cation	$CONH_2$	aliphatic neutral amide
N1+	sp_3 amine NH cation	Ar-$CONH_2$	aromatic neutral amide
N1=	sp_2 amine NH cation	CONHR (trans)	aliphatic neutral amide
NH=	sp_2 NH with lone pair	Ar-CONHR (trans)	aromatic neutral amide
N:	sp_3 N with lone pair	AMIDINE	aliphatic cationic amidine
N:#	sp N with lone pair	Ar-AMIDINE	aromatic cationic amidine
NM3	trimethylammonium	M-DIAMINE	meta-diamino benzene
O::	sp_2 carboxy oxygen	C3	methyl group

get molecules, in order to obtain the most diversified chemical information. No metal probe was chosen because there is no experimental evidence that metals are involved in ligand-COX interactions. The probes used give precise energetic and spatial information which can account for specificity and sensitivity of ligand-enzyme interactions. This is particularly important, because they represent chemical groups that could be added to a ligand in order to improve the selective behaviour.

Directive MOVE [20] controls the behaviour of the target. The default value is MOVE = 0 which allows lone pairs and tautomeric hydrogens to move in response to the probe. It also allows the torsion angles of groups like sp^3 hydroxyl or sp^2 amine to change, so that they can twist and make the most favourable hydrogen bonds with the probe. The target responds more flexibly to the probe when MOVE = 1 [20]. Flexible side chains, such as a lysine side chain, respond as the probe is moved from one grid point to the next. For instance, if a carbonyl oxygen probe is used, then the cationic NH_3 of a lysine in the target would tend to move towards it, so that more favourable hydrogen bonds can form and electrostatic interactions can occur. We therefore repeated the calculations using the directive MOVE = 1 to consider the conformational flexibility of residues that could be important, as previously stated, in the entry and exit processes.

Principal component analysis on the GRID field variables

The three-dimensional matrices obtained from each GRID calculation were rearranged in two-dimensional **X** matrices made up of 84 rows (42 probes × 2 target molecules) and 11362 column variables (the energy values calculated at each grid node).

The **X** data matrices were analysed by PCA [21], a multivariate statistical analysis method, which allows the systematic information contained in a data matrix to be extracted and the original number of variables to be reduced into a few factors called Principal Components (PCs). The results of the analysis are given in a few informative diagrams, which permit a simple, straightforward interpretation of the problem under investigation.

In our study, PCA was performed as implemented in the GOLPE [24] program. According to Pastor et al. [18b] only negative interaction energy values

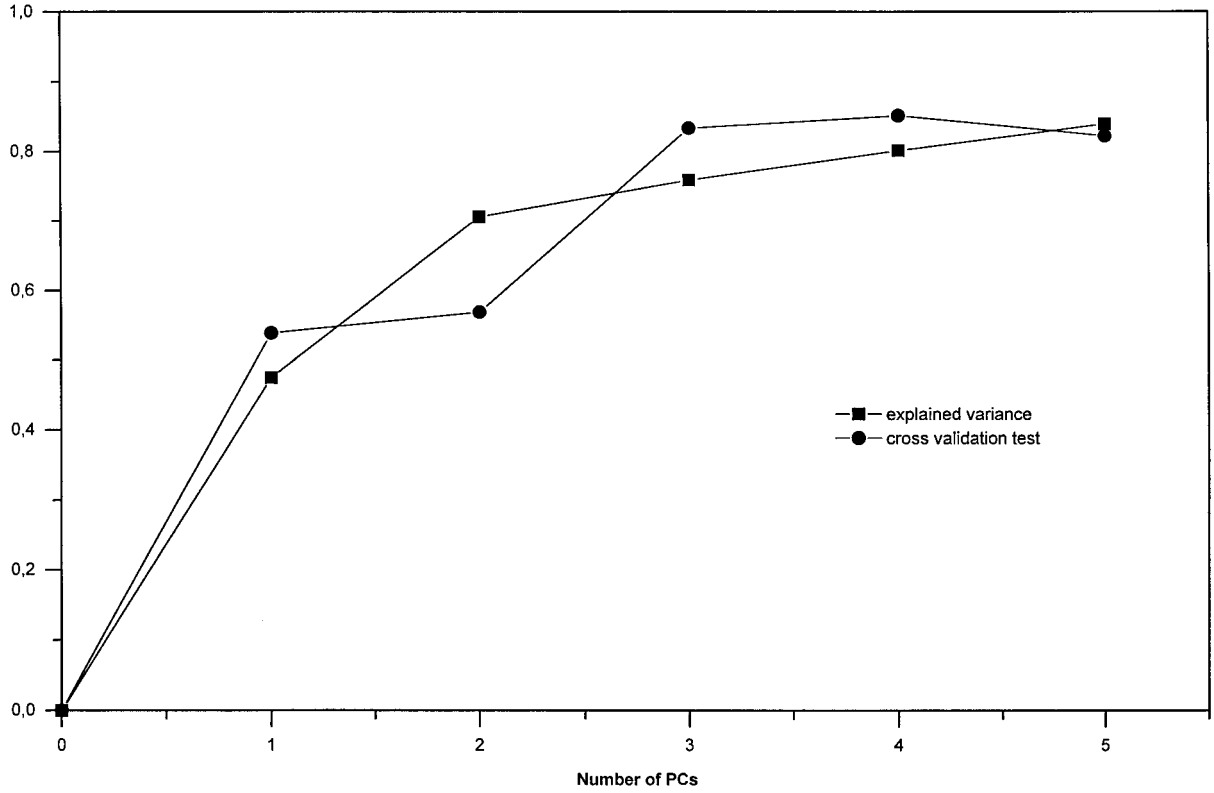

Figure 2. Obtained values of explained variance and cross-validation test using directive MOVE = 0.

should be used in the calculation because they identify favourable interaction regions, while the positive energies are related to repulsive interactions and steric effects. In order to check whether this approach would also be better in this particular instance, we tried to perform calculations including positive interaction energies. These only emphasised the few already known structural differences (especially amino acid at position 523) and did not add anything new to the experimental evidence, while only complicating the interpretation of the data. Therefore the **X** matrices were pretreated, setting all the positive energies to 0 kcal/mol in this particular study also.

Results and discussion

PCA on GRID variables obtained with MOVE = 0 (rigid structures)

PCA gave a model of three statistically significant PCs, which explain 76% of the variance of the data (Figure 2). The number of statistically significant components was chosen according to the cross-validated tests (Figure 2) [25].

Most of the chemical information is contained in the first two PCs as is shown in the score plot of the second versus the first component (Figure 3), where each point represents a different probe-enzyme interaction.

The first component discriminates between the two isoforms of the enzyme and the points corresponding to interactions with COX-2 enzyme are all located at positive values of the first component, while the points corresponding to interactions with COX-1 enzyme are located at negative score values.

From a practical point of view, the more different the behaviour of a probe when it interacts with the two target enzymes, the more selective it is, and the distance between the two points identifying the interaction with COX-1 and COX-2 should be larger in the first PC. In our case, the most selective probes are found at the bottom of the score plot.

The second component distinguishes between the probes, as shown in Figure 3 where each point is labelled by the respective probe code (Table 1). Ac-

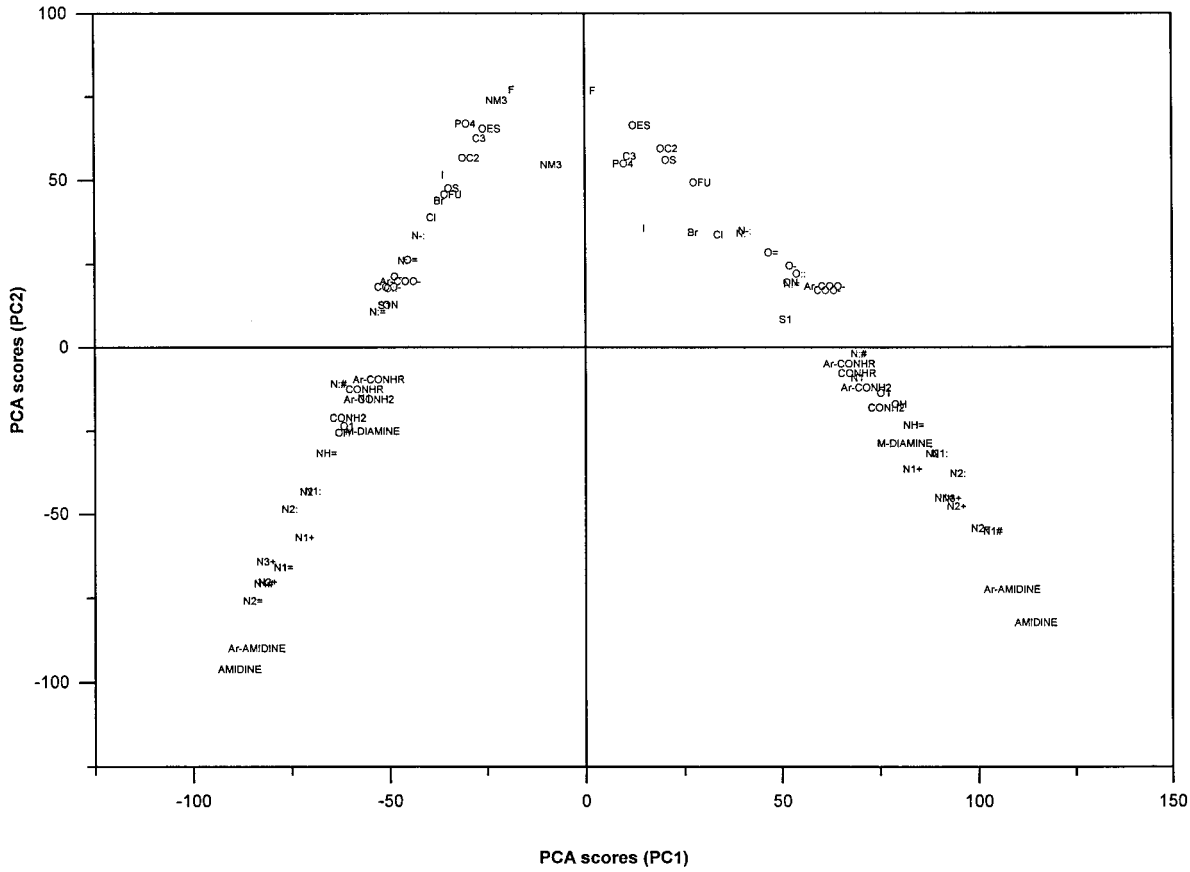

Figure 3. PCA score plot of the second versus the first PC. Each probe-enzyme interaction is indicated by the code for the corresponding probe.

cording to this component, the capacity to donate a hydrogen bond seems to be one of the major determinants for the classification and it decreases from the bottom to the top along the second component; the most selective probes can donate one or more hydrogen bonds.

As shown in previous studies [18], the first component describes the specific behaviour of a probe when it interacts with different regions of the two enzymes, while the second PC describes the intensity of non-selective interactions with common regions of the enzyme. In fact, the considered probes generally have similar score values on the second PC for both COX-1 and COX-2.

To solve our classification problem it is important to know which probes show a selective behaviour as well as to determine which regions of the considered enzyme are the most suitable for the selective interactions. In PCA a loading is associated with each variable for each PC. The loadings represent the coefficients of the original variables in the linear combination which define a component and are therefore related to the importance of the considered variables in determining that PC. The two-dimensional loading plot of Figure 4 shows which energy variables mainly influence each component.

Because each variable has a precise grid location, the corresponding three-dimensional loading plots for each component (Figures 5 and 6) define in a more intuitive way the regions of the target molecules responsible for specific and non-specific interactions.

Since only negative interaction energies are used, energy variables characterised by negative loading values are responsible for probe-enzyme interactions identified by positive scores and vice versa.

The two-dimensional loading plot of Figure 4 can be divided into three main regions: **A**, **B** and **C**. The loadings belonging to region **A** are mainly responsible for the position in the score plot of the most selective probe-COX-2 interactions. In the corresponding three-dimensional loading plot for the first PC, these loadings define a region (grid surface Figure 5) within

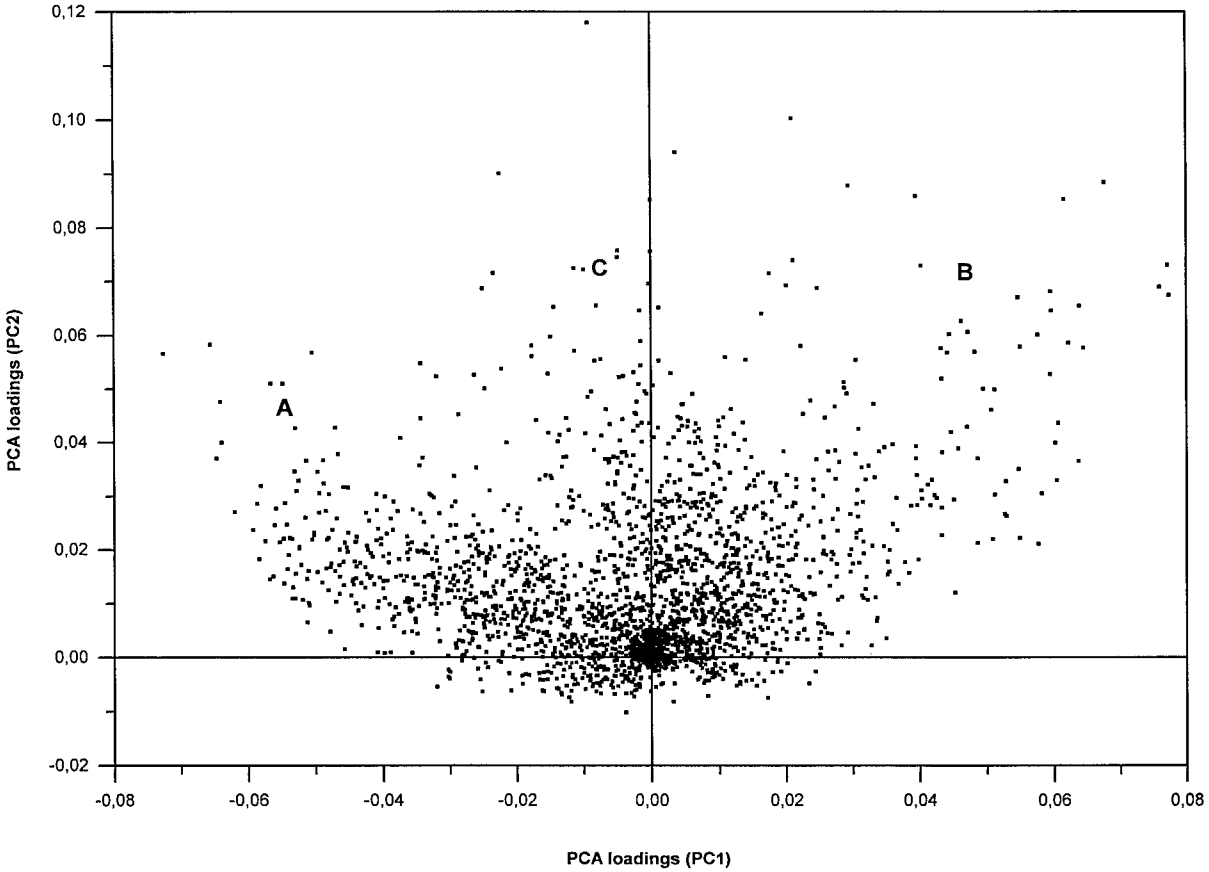

Figure 4. PCA loading plot of the second versus the first PC (MOVE = 0) where each point corresponds to a grid field variable.

the second pocket of COX-2 which is determined by Val523.

The entry gate to this second pocket is quite hydrophobic, due to the presence of the side chains of Leu352, Tyr355, Val523, and Phe518. Conversely, the inner part is relatively polar and inside this region the hydrogen-bond donor probes could interact favourably with the side chain of Gln192[b], with the heterocylic nitrogens of His90[c], with the side chain of Arg513 and finally, with the carbonyl oxygen of Phe518.

The most selective probes, which can donate more than one hydrogen, have a high degree of affinity for COX-2 in this region. In particular the AMIDINE and Ar-AMIDINE probes (Figure 7), which show a positively charged polar head bound to a hydrophobic group, are identified as the most selective.

It is interesting to observe that these bulky multiatom probes (particularly Ar-AMIDINE) look somewhat like the benzensulfonamide moiety of the previously mentioned COX-2 inhibitors (Figure 1) and the supposed selective behaviour takes on significance because of the mechanism of action proposed for the reported analogues. In fact, the aromatic or aliphatic chain should be surrounded in the first part of the pocket by the hydrophobic residues but the polar head of the probes can also make hydrogen bonds with polar residues of the inner part. This pattern is practically absent in COX-1 enzyme, which retains only the last polar part of the second pocket but this is not accessible. Therefore the interactions between COX-1 and the probes are shifted to the opposite side of the first PC.

The amidine group is not an extensively used substituent in the design of COX-2 selective inhibitors. We found an indication about the activity of a COX-2 selective molecule substituted with an amidine-like moiety [26]. The molecule shown in Figure 8 has an $IC_{50} > 100$ μM against human COX-1, while it has an $IC_{50} = 4.2$ μM against human COX-2.

A docking experiment by the GRID/GROUP[d][20] docking procedure was performed showing that the compound of Figure 8 can interact with the COX-2

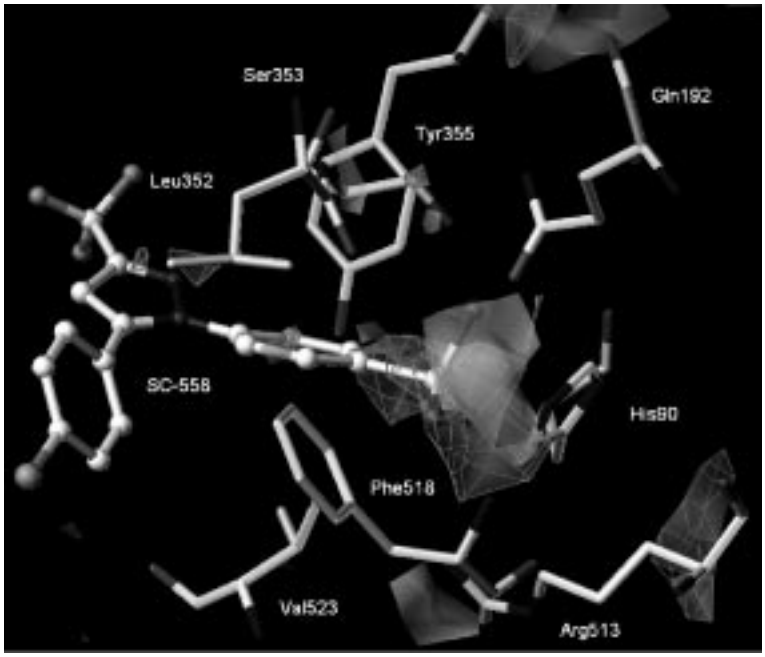

Figure 5. Three-dimensional loading plot of the first PC contoured at the −0.04 level (grid surface) and three-dimensional loading plot of the second PC contoured at the +0.04 level (shadow surface). The figure blows up the residues surrounding the second pocket. The two contour surfaces almost design the internal surface of the second pocket in COX-2 (loadings of PC1) and in COX-1 (loadings of PC2). For better orientation the SC558 ligand was added, as it appears bound to the X-ray structures of murine COX-2 [8] (6*cox* PDB entry). To obtain this figure, the complexed enzyme was superimposed over the uncomplexed one.

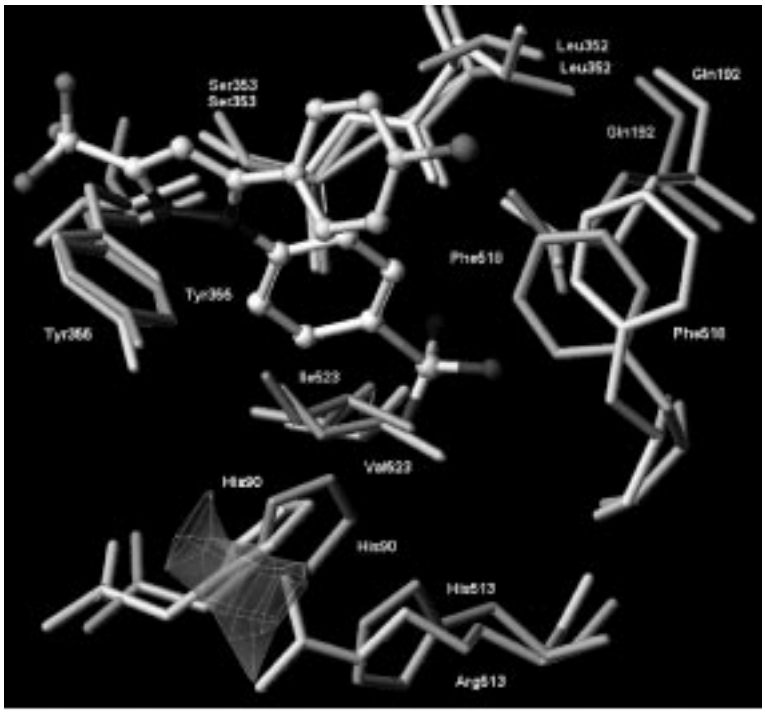

Figure 6. Three-dimensional loading plot of the first PC contoured at the +0.05 level. The loadings define an important region around the polar extremity of Arg513 which can probably interact with the hydrogen-bond network at the bottom of the NSAID binding site. For better orientation the SC558 ligand was added, as it appears bound to the X-ray structures of murine COX-2 [8] (6*cox* PDB entry). For comparison, we added both superimposed COX-1 (dark grey) and COX-2 (light grey) enzymes.

enzyme when the amidine-like moiety binds inside the second pocket. The interaction seems to be possible when the amidine group is positively charged, when both directive MOVE = 0 and MOVE = 1 are considered. As observed in the X-ray structure of murine COX-2 complexed with SC-558, the two oxygen atoms can form hydrogen bonds with His90 and Arg513 and the positively charged amidine group can form a hydrogen bond with the carbonyl oxygen of Phe518.

The PCA model also seems to explain why methyl sulfone derivatives are usually less active than their sulphonamide counterpart [16], at least within this particular class of 1,5-diarylpyrazole derivatives. According to the PCA results, the oxidised sulphur atom is not selective itself (OS probe in the score plot of Figure 3). Therefore it seems reasonable that the notable COX-2 selectivity of the sulphonamide derivatives could mainly be due to the amide nitrogen (N2 probe), while the methyl of methyl sulfone moiety does not show any selectivity (C3 probe in the score plot of Figure 3).

The loadings belonging to region **B** of the two-dimensional loading plot are also responsible for the position of the most selective probe-COX-1 interactions in the score plot. They define an interaction region around residue Arg513 (surface of Figure 6) which represents an important difference between COX-1 and COX-2. In fact, His513 in COX-1 replaces Arg513 in COX-2. The difference in this position could influence the interaction with COX-1 and COX-2 isoenzymes in two different ways. First, the long side chain of Arg513 extends far enough to interact directly with a probe along the whole second pocket, while the imidazole ring of His513^c does not. Moreover, the Arg513 side chain could participate in the hydrogen-bond network made up of Arg120, Tyr355 and Glu524 at the bottom of the NSAID binding site. These residues play an important role in the entry process to the catalytic cavity. It is clear from crystallographic studies [9] that alternate open/closed conformations involving interaction of Glu524 with Arg513 could explain some kinetics aspects of the COX-2 selective inhibition, because the shorter histidine side chain of COX-1 cannot directly participate in this hydrogen-bond network.

The loadings belonging to region **C** define the contribution of the variables responsible for the second PC. It is interesting to note that the most selective probes (according to the first PC) should be able to interact simultaneously with regions common to both

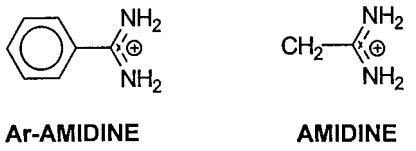

Figure 7. AMIDINE and Ar-AMIDINE probes.

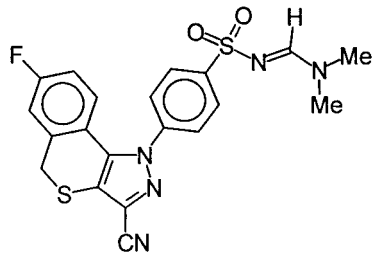

Figure 8. Example of molecule with an amidine-like moiety, showing selective behaviour.

COX-1 and COX-2 (according to the second PC). Therefore, the probes which are located at the bottom of the score plot should be able to increase both specific and non-specific interactions with the target enzymes, while the probes at the top of the score plot should not be able to increase either specific or non-specific interactions.

Because the second component describes the intensity of non-specific interactions, the corresponding three-dimensional contour plot defines the common area in which ligands interact with both COX-1 and COX-2.

The loadings of the second PC delineate a small region which almost corresponds to the polar head of the second pocket in COX-2 and define the whole pocket in COX-1 (shadow surface of Figure 5). Practically, the second component again describes the interaction of polar probes with residues Phe518 and Gln192. Since we know that the entry into the second small pocket of COX-1 is not allowed, increasing the affinity of a group in this region will not facilitate interaction with COX-1. On the contrary, increasing the affinity of polar groups for the polar head of the second pocket in COX-2 could lead to ligands with improved selectivity.

PCA on GRID variables obtained with MOVE = 1 (flexible structures)

As previously stated, the GRID directive MOVE = 1 takes into account the conformational flexibility of the amino acidic side chains. The two descriptions of the binding site cavity of COX-2 enzyme, obtained

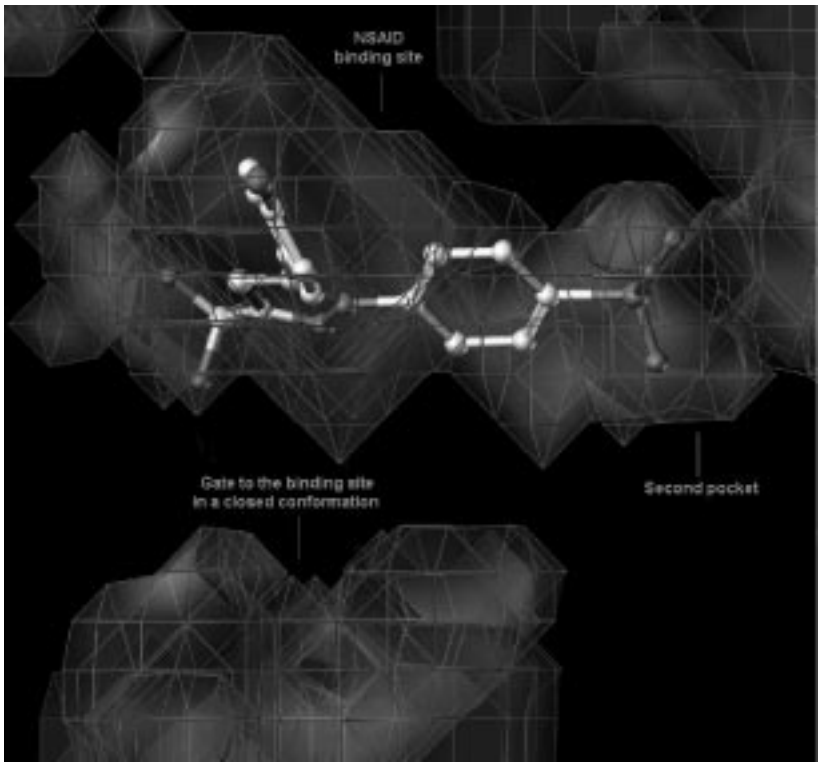

Figure 9. Description of the binding site by means of GRID with MOVE = 0 (rigid structures).

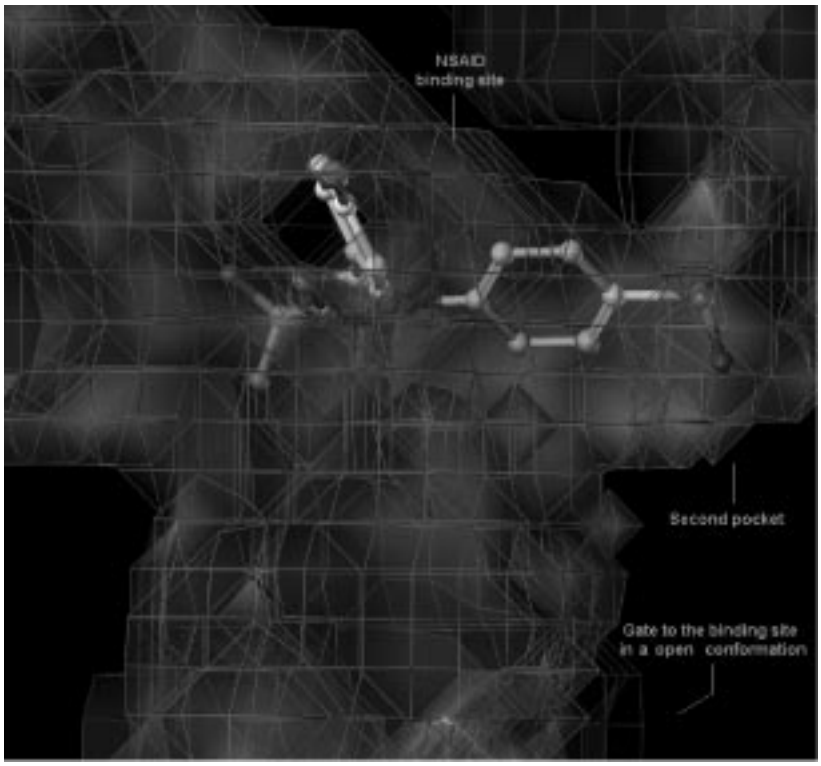

Figure 10. Description of the binding site by means of GRID with MOVE = 1 (flexible structures).

by using the same methyl probe but with the two different directives MOVE = 0 (Figure 9) and MOVE = 1 (Figure 10), show that the cavity is much larger when conformational flexibility is considered. Moreover, even if starting from a conformation in which the entry gate to the NSAID binding site is in a closed conformation, we can consider the possibility of an open conformation occurring in the binding process.

Therefore, in order to further investigate the influence of the conformational flexibility of the side chains, all the calculations were repeated using directive MOVE = 1. This second PCA analysis yielded a model of three statistically significant components, which explains 89% of the variance of the data (Figure 11). Although the explained variance is not too different from that obtained with the previous analysis, the interpretation of the model seems to be quite different. It is always true that the first component discriminates between the targets, while the second component ranks the probes, but they are classified in a different way (Figure 12).

This new classification identifies some differences in the influence of variables on determining the components.

Analogously to the previous analysis, the two-dimensional loading plot can be divided into three different regions **A**, **B** and **C**.

Also in this case, the variables represented by the loadings of region **A** are principally responsible for the position in the score plot of the most selective probe-COX-2 interactions, while the variables represented by loadings of region **B** are principally responsible for the most selective probe-COX-1 interactions. The corresponding three-dimensional loading plots identify the same regions highlighted in the previous analysis, therefore leading to an analogous interpretation. They confirm the role played by the second pocket of COX-2 in the selective behaviour of a probe and the importance of residue Arg513. The variables represented by the loadings belonging to region **C** are also in this case responsible for the distribution of the probes along the second PC.

Polar single-atom probes like N2:, N1:, NH=, N:#, O1, N2 or multi-atom probes like Ar-CONHR and CONHR can favourably interact inside the second cavity of COX-2 and this can explain their selective behaviour. At first glance it is not clear why the Ar-AMIDINE and AMIDINE, which were previously indicated as the most selective probes, do not show a selective behaviour along with other charged probes, previously indicated as selective. The corresponding three-dimensional loading plot for the second component (Figure 13) identifies an extensive region delimited by the amino acidic triad Tyr355, Glu524 and Arg120 at the bottom of the NSAID binding site.

The previous model also identified this region but it was less important according to the loading values for the second component.

The interaction of charged groups inside this region should be particularly strong, because of the presence of the negatively charged side chain of Glu524 and the positively charged side chain of Arg120.

It is important to recall that the COX-1 and COX-2 structures used in this study show a closed conformation where Arg120 maintains its salt-bridge with Glu524. In COX-2 the closed conformation can be transformed into an open one thanks to the interaction of Arg120 with Arg513 that is not allowed in COX-1 which shows an His513 residue at the same position.

Practically, polar groups, like the selective probes mentioned, could reach the catalytic site and favourably interact inside the second pocket of COX-2 only if there is an open gate determined by Arg513 but they cannot modify the closed conformation. In contrast, charged groups can destroy the salt-bridge between Arg120 and Glu524 in both COX-1 and COX-2, which probably leads to the same conformational state at the bottom of the NSAID binding site for the two isoforms.

These results are in agreement with the experimental evidence that non-selective cyclooxygenase inhibitors like flurbiprofen, indomethacine and even aspirin generally show a carboxylate moiety. The structures of ovine COX-1 bound to aspirin [27] and of murine [8] COX-2 bound to flurbiprofen and indomethacine clearly show that the carboxylate group forms a salt-bridge with the guanidinium group of Arg120 and a hydrogen bond to Tyr355.

It is interesting to note that the non-selective behaviour of Ar-COO- and COO- probes was already identified by the previous analysis with directive MOVE = 0 and it is confirmed by this second analysis. The carboxylate moiety probably competes efficiently to interact with the Arg120 residue because of the presence of a unitary negative charge, which strongly alters the conformational features at the bottom of the NSAID binding site. This second model also supports the experimental evidence that sulphonamide derivatives are generally more active than methyl sulfone congeners. The only difference is that in this second analysis the OS probe seems to show a selective behaviour.

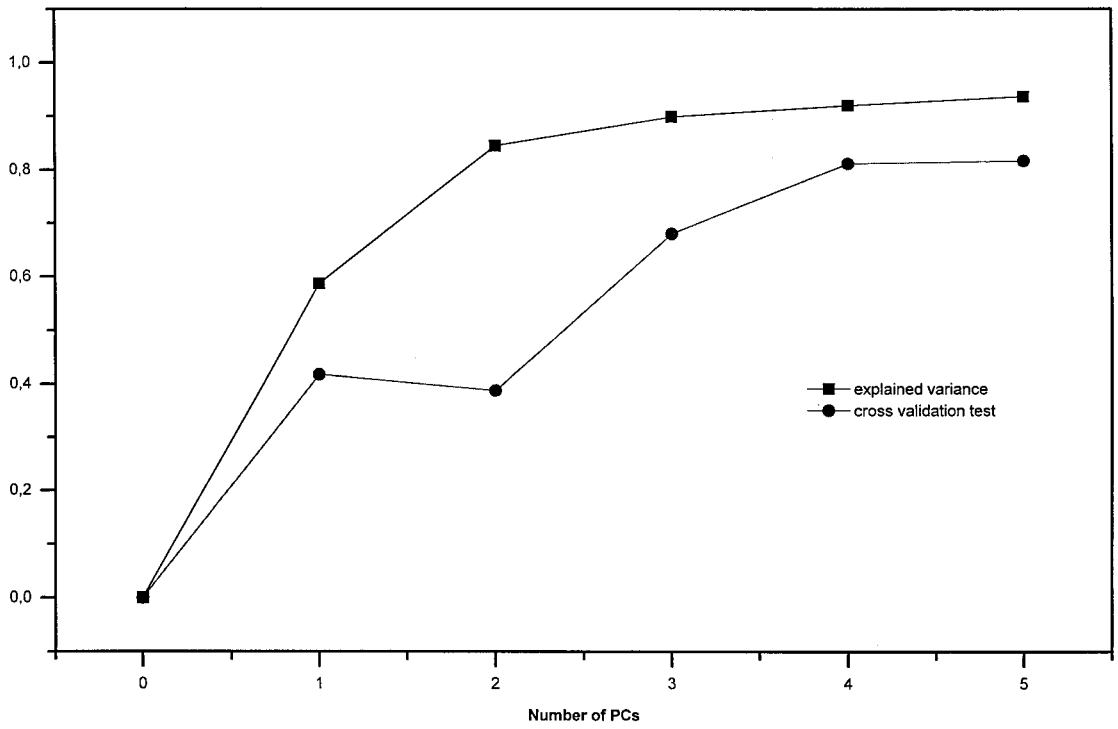

Figure 11. Obtained values of explained variance and cross-validation test using directive MOVE = 1.

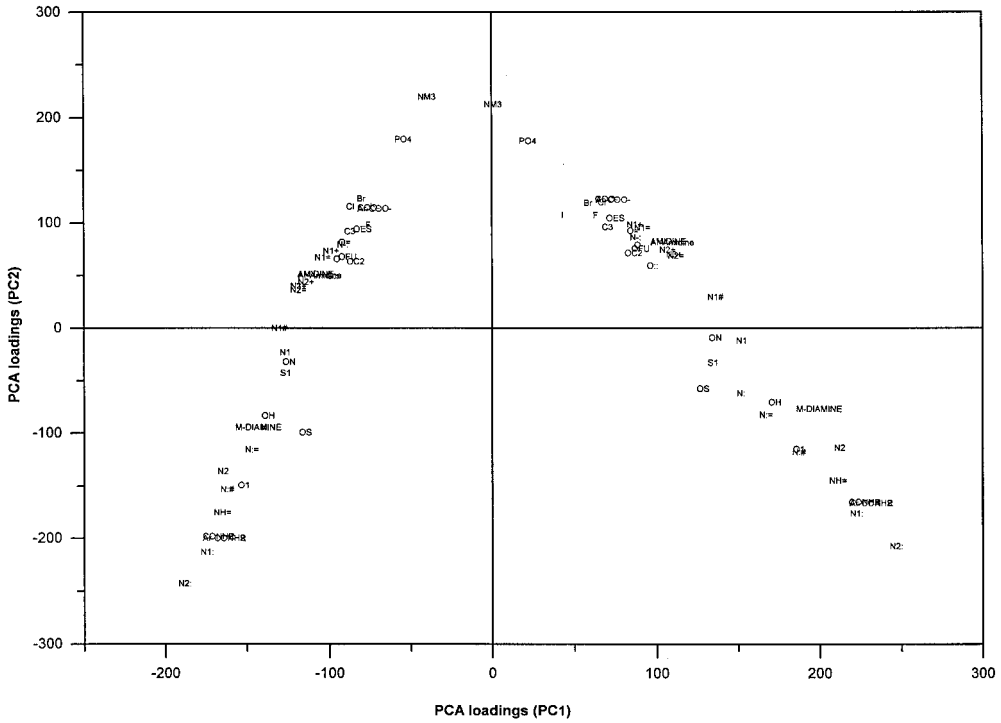

Figure 12. PCA score plot of the second versus the first PC (model with MOVE = 1). Each probe-enzyme interaction is indicated by the code for the corresponding probe.

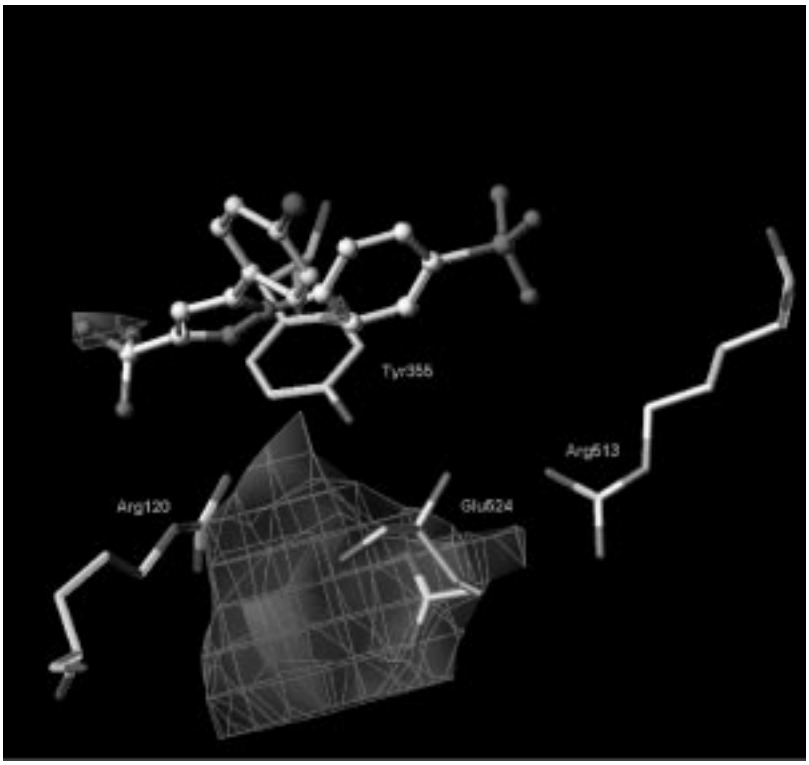

Figure 13. Three-dimensional loading plot of the second PC contoured at the +0.03 level (model with MOVE = 1). For better orientation the SC558 ligand was added.

We can state that when calculations are performed with MOVE = 0, the results show the partial effect of the interactions inside the second pocket. When directive MOVE = 1 is used the results account for the global effect inside the second pocket and at the bottom of NSAID's binding site, leading to a better model in terms of consistency with experimental data.

Conclusions

In this work the most important structural features of a selective COX-2 inhibition have been analysed. The GRID/PCA procedure has highlighted the same regions of the enzyme that other authors have previously identified by a comparative analysis of COX-1 and COX-2 X-ray structures, and the obtained results are compatible with earlier findings [17]. The importance of the Val523 residue in enlarging the active site and allowing the access to the second pocket in COX-2 has been noted. This cavity plays a major role in ligand selectivity: probes with a hydrophobic core and a polar head can selectively interact inside this second cavity, which shows a hydrophobic entry gate and a polar head.

The PCA results show the well-established role played by the Arg513 residue. The long side chain of the residue provides hydrogen bond interactions inside the whole second pocket of COX-2. Moreover, the residue can be involved in alternate conformations of residues which make up the hydrogen-bond network at the bottom of the entry channel which leads to the NSAID binding site. Arg513 can therefore influence the affinity inside the second pocket and control the kinetics of the ligand entry process.

The results confirm the reliability of the method and may help to identify selective inhibitors with docking procedures or 3D-QSAR analyses.

The most interesting advantage of the method is that it allows all the possible ligand-enzyme interactions to be evaluated quantitatively and to assign their relative importance in an objective way. GRID maps calculated on target molecules seem to be particularly useful for this kind of analysis (especially with MOVE = 1), because they give detailed spatial information which accounts for affinity and specificity, the probes

used being representative of chemical groups which could be added to a ligand structure.

PCA can distinguish between different kinds of interactions. The results support some experimental evidence, such as the high selectivity of molecules bearing a benzensulfonamide moiety [16] with respect to their methyl sulphone analogues.

The method suggests which chemical group could be added to a template structure in order to increase the ligand COX-2 selectivity. The details of the possible substitution of the sulphonamide moiety with an amidine group have been discussed. According to the model, other interesting groups could be added to a core structure like an aromatic amide moiety or a meta-diamino benzene.

It has been discussed how the protonation and the charge of the considered molecules can influence selective interaction with the two target enzymes. It appears that a charged group may increase the affinity of a ligand for both target molecules, to the detriment of selectivity.

Notes

[a]The crystals contain a non-steroidal anti-inflammatory drug 'flurbiprofen' and the drug binding site has been identified, although regularised and refined coordinates for protein-bound drug are not available.

[b]If the X-ray structures were obtained with a >2.5 Å resolution, no clear differentiation between Gln192 side chain heavy atoms is possible. In the case of these structures, the orientation found in COX-1 structure is very similar to the orientation found in COX-2 structure. Moreover, the conformation of the Gln192 side chain in 1prh has been confirmed by a recent COX-1 structure solved at 3.0 Å resolution (1cqe).

[c]The GRID force field takes into account two tautomeric positions for hydrogen on the His ring. Other kinds of parameterisation are allowed but this one is recommended under normal physiological conditions. For more detailed information see GRID16 manuals, Molecular Discovery Ltd.

[d]Programme GROUP resembles GRID but can deal with bigger probes. It can also be used to fit a molecule of known structure into a set of GRID maps.

References

1. a. Vane, J.R., Nature (New Biol.), 231 (1971) 232.
   b. Smith, J.B. and Willis, A.L., Nature (New Biol.), 231 (1971) 235.
2. Vane, J.R., Mitchell, J.A., Appleton, I., Tomlinson, A., Bishop-Bailey, D., Croxtail, J. and Willoughby, D.A., Proc. Natl. Acad. Sci. USA, 91 (1994) 2046.
3. a. Hla, T. and Neilson, K., Proc. Natl. Acad. Sci. USA, 89 (1992) 7384.
   b. Xie, W., Chipman, J.G., Robertson, D.L., Erokson, R.L. and Simmons, D.L., Proc. Natl. Acad. Sci. USA, 88 (1991) 2692.
   c. Kujubu, D.A., Fletcher, B.S., Varnum, B.C., Lim, R.W. and Herschman, H.R., J. Biol. Chem., 14 (1971) 1171.
4. a. Masferrer, J.L., Zweifel, B.S., Manning, P.T., Hauser, S.D., Leahey, K.M., Smith, W.G., Isakson, P.C. and Seibert, K., Proc. Natl. Acad. Sci. USA, 91 (1994) 3228.
   b. Seibert, K. and Masferrer, J.L., Receptor, 4 (1994) 17.
   c. Meade, E.A., Smith, V.L. and DeWitt, D.L., J. Biol. Chem., 268 (1993) 6610.
   d. Mitchell, J.A., Akarasereemont, P., Thiemerman, C., Flower, R.J. and Wane, J.R., Proc. Natl. Acad. Sci. USA, 90 (1993) 11693.
   e. Harada, Y., Hatanaka, K., Saito, M., Majima, M., Ogina, M., Kawamura, M., Ohno, T., Yang, Q., Katori, M. and Yamamoto, S., Biomed. Res., 15 (1994) 127.
5. a. Vane, J.R., Nature, 367 (1994) 215.
   b. Vane, J.R. and Botting, R.M., Inflamm. Res., 44 (1995) 1.
   c. DeWitt, D.L., Meade, E.A. and Smith, V.L., Am. J. Med., 95 (1993) 2A-40S.
6. Wallance, J.L., TIPS, 20 (1999) 4.
7. Picot, D., Loll, P.J. and Garavito, M., Nature, 367 (1994) 243.
8. Kurumbail, R.G., Stevens, A.M., Gierse, J.K., McDonald, J.J., Stegeman, R.A., Pak, J.Y., Gildehaus, D., Miyshiro, J.M., Penning, T.D., Seibert, K., Isakson, P.C. and Stallings, W.C., Nature, 384 (1996) 644.
9. Luong, C., Miller, A., Barnett, J., Chow, J., Ramesha, C. and Browner, M.F., Nature Struct. Biol., 3 (1996) 927.
10. Gierse, J.K., McDonald, J.J., Hauser, S.D., Rangwala, S.H., Koboldt, C.M. and Seibert, K., J. Biol. Chem., 271 (1996) 15810.
11. Quipeng, G., Lee-Ho, W., Ke-He, R. and Richard, J.K., J. Biol. Chem., 271 (1996) 19134.
12. Futaki, N., Takahashi, S., Yokoyama, M., Arai, I. and Higuchi, S., Prostaglandins, 47 (1994) 55.
13. Swingle, K.F., Moore, G.G.I. and Grant, T.J., Arch. Int. Pharmacodyn., 221 (1976) 132.
14. Gans, K.R., Galbraith, W., Roman, R.J., Haber, S.B., Kerr, J.S., Schmidt, W.K., Smith, C., Hewes, W.E. and Ackerman, N.R., J. Pharmacol. Exp. Ther., 254 (1990) 180.
15. Seibert, K., Zhang, Y., Leahy, K., Hauser, S., Masferrer, J., Perkins, W., Lee, L. and Isakson, P., Proc. Natl. Acad. Sci. USA, 91 (1994) 12013.
16. Penning, T.D., Talley, J.J., Bertenshaw, S.R., Carter, J.S., Collins, P.W., Docter, S., Matthew, J.G., Lee, L.F., Malecha, J.W., Miyashiro, J.M., Rogers, R.S., Rogier, D.J., Yu, S.S., Anderson, G.D., Burton, E.G., Cogburn, J.N., Gregory, S.A., Koboldt, C.M., Perkins, W.E., Veenhuizen, A.W., Zhang, Y.Y. and Isakson, P.C., J. Med. Chem., 40 (1997) 1347.
17. a. Pouplana, R., Perez, C., Sanchez, J., Lozano, J.J. and Puig-Parellada, P., J. Comput.-Aided Mol. Design, 13 (1999) 287.
   b. Gabriel, F.F., Vasantha, P. and Kuppuswamy, N., Bioorg. Med. Chem., 6 (1998) 2337.
   c. Pedretti, A., Villa, A.M., Villa, L. and Vistoli, G., Il Far-

maco, 52 (1997) 487.
d. Stewart, K.D., Loren, S., Frey, L., Otis, E., Klinghofer, V. and Hulkower, K., Bioorg. Med. Chem. Lett., 8 (1998) 529.
18. a. Cruciani, G. and Goodford, P.J., J. Mol. Graphics, 12 (1994) 116.
b. Pastor, M. and Cruciani, G., J. Med. Chem., 38 (1995) 4637.
c. Matter, H. and Schwab, E., J. Med. Chem., 42 (1999) 4506.
19. Goodford, P.J., J. Med. Chem., 28 (1985) 849.
20. GRID v.16 Manuals, Molecular Discovery Ltd.
21. Wold, S., Esbensen, K. and Geladi, P., Chemom. Intell. Lab. Syst., 2 (1987) 37.
22. Brookhaven National Laboratories Protein Databank.
23. SYBYL 6.3 Molecular Modeling Software, Tripos, Inc.
24. Baroni, M. and Pastor, M., GOLPE v. 4.0.2, Multivariate Infometric Analysis 1998.
25. a. Cruciani, G., Baroni, M., Clementi, S., Costantino, G., Riganelli, D. and Skagerberg, B., J. Chemometrics, 6 (1992) 335.
b. Wold, S., Technometrics, 20 (1978) 397.
26. WO96/09304 PCT/US95/11403.
27. Loll, P.J., Picot, D. and Garavito, R.M., Nat. Struct. Biol., 2 (1995) 639.