Journal of Computer-Aided Molecular Design **18:** 683–696, 2004.
DOI 10.1007/s10822-004-6258-1

Docking studies on NSAID/COX-2 isozyme complexes using Contact Statistics analysis

Giuseppe Ermondi[a,*], Giulia Caron[a], Raelene Lawrence[b] & Dario Longo[c]
[a]*Dipartimento di Scienza e Tecnologia del Farmaco, V.P. Giuria 9, I-10125 Torino, Italy;* [b]*Chemical Computing Group, Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, Quebec, Canada H3A 2R7;* [c]*Bioindustry Park del Canavese, V. Ribes 5, I-10010 Colleretto Giacosa (TO), Italy*

Received 11 August 2004; accepted in revised form 14 November 2004
© Springer 2005

Key words: celecoxib, Contact Statistics analysis, COX-2, docking, flurbiprofen, MOE, nimesulide, NSAIDs, rofecoxib

Summary

The selective inhibition of COX-2 isozymes should lead to a new generation of NSAIDs with significantly reduced side effects; e.g. celecoxib (Celebrex®) and rofecoxib (Vioxx®). To obtain inhibitors with higher selectivity it has become essential to gain additional insight into the details of the interactions between COX isozymes and NSAIDs. Although X-ray structures of COX-2 complexed with a small number of ligands are available, experimental data are missing for two well-known selective COX-2 inhibitors (rofecoxib and nimesulide) and docking results reported are controversial. We use a combination of a traditional docking procedure with a new computational tool (Contact Statistics analysis) that identifies the best orientation among a number of solutions to shed some light on this topic.

Introduction

Nonsteroidal anti-inflammatory drugs (NSAIDs) [1, 2] are of great importance in the treatment of inflammatory diseases. In 1971 [3], it was discovered that COXs are the molecular targets for NSAIDs; thus, inhibition of COXs leads to a decreased production of prostaglandins and tromboxanes, which, in turn, accounts for the beneficial effects of NSAIDs (e.g. anti-inflammatory, antipyretic, analgesic and cardiovascular effects), as well as their undesirable side effects (e.g. gastrointestinal).

In the early 1990s, COX was demonstrated to exist as two distinct isoforms. COX-1 is constitutively expressed as a housekeeping enzyme in nearly all tissues, and mediates physiological response (e.g., cytoprotection of the stomach, platelet aggregation). COX-2, which is expressed by cells that are involved in inflammation (e.g., macrophages, monocytes, synoviocytes), has emerged as the isoform that is primarily responsible for the synthesis of the prostanoids involved in pathological processes, such as acute and chronic inflammatory states [4].

Classical NSAIDs (i.e., aspirin, ibuprofen, naproxen, but not nimesulide) are non-selective inhibitors of both isozymes (IC_{50} for COX-1 is similar to the IC_{50} for COX-2) and can cause gastric and renal failure [1, 5, 6]. Hence, there have been sustained efforts to identify selective COX-2 inhibitors (compounds whose IC_{50} for COX-1 inhibitory activity is significantly above that of COX-2). Recently, the COX-2 selective inhibitors celecoxib (Celebrex®) and rofecoxib (Vioxx®) have been approved by the Federal

*To whom correspondence should be addressed. Phone: +39-011-6707282; Fax: +39-011-6707687; E-mail: giuseppe.ermondi@unito.it

Drug Administration. It has meanwhile been hypothesized that there might be other forms of the COX enzyme yet to be discovered [7]. While this concept is being developed and tested, it has become essential to gain additional insight into the details of the interactions between COX isozymes and NSAIDs.

Although X-ray structures of COX-2 complexed with a small number ligands are available (compounds **1**, **2**, **3** and **4** in Figure 1), experimental data are lacking for two well-known selective COX-2 inhibitors (rofecoxib **5** and nimesulide **6** in Figure 1) and docking results have led to different proposals for their interaction with COX-2 [8, 9], more than one binding-mode being theoretically possible. COX-2 inhibitors are thus a well-known biological system that is suitable for testing new molecular modeling tools.

A number of docking procedures based on different search and scoring algorithms have been proposed [10], but none can treat all biological systems with the same accuracy and efficacy (e.g., Schulz-Gash et al. [11], Wang et al. [12], Ferrara et al. [13]). It is thus advisable, once the biological target has been selected, to set up an adequate system strategy for the system and study goal.

Statistics-based approaches [14, 15] to determine preferred locations of bound atoms from experimental data are a promising new tool, which we have already applied with excellent results in a study on albumin binding [16]. Here two sampling methods are used to generate different protein-ligand configurations (also known as 'poses' [11, 13, 17]). In the first method the ligands are placed manually in the active site and then moved using simple systematic molecule displacements and rotations. In the second a Genetic Algorithm (GA) is used to generate different poses. We evaluate the influence of the pose generation method by comparing the results obtained with the two methods. The poses with the lowest non-bond energy, calculated as the sum of Van der Waals and electrostatic energy, were selected and analyzed using the Contact Statistics tool to choose the most plausible binding mode and to better understand the interaction keys between enzyme and inhibitor.

Methods

Molecular Mechanics (MM) and Dynamics (MD) calculations were run on a Silicon Graphics O2

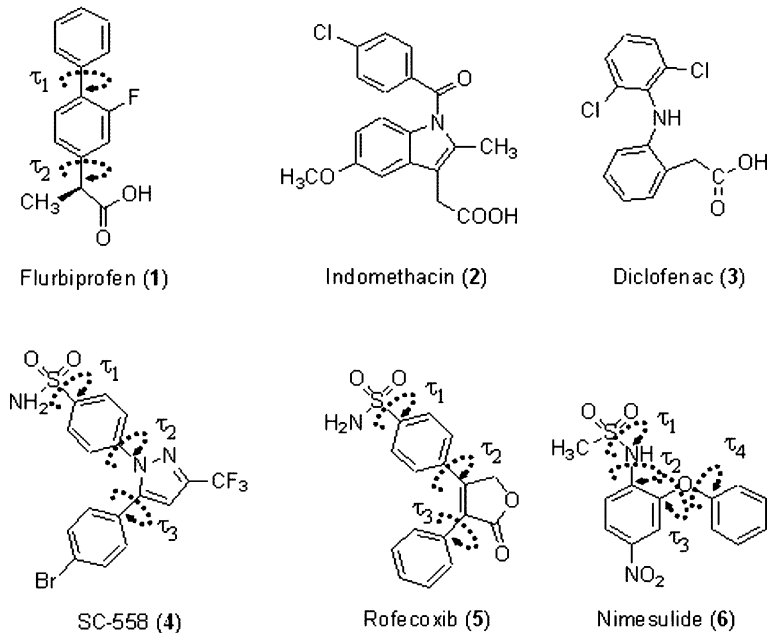

Figure 1. Chemical structures of compounds under study. The relevant torsional angles are shown for compounds under conformational analysis.

workstation and on a Linux based biprocessor Appro 1124 server using the Accelrys molecular modeling package (Accelrys, San Diego, CA, USA) and the MOE molecular modeling package (MOE Version 2002.03, Chemical Computing Group Inc., Montreal, Canada). Docking results were analyzed using MOE.

Conformational analysis of the ligands

Molecular dynamics (MD) methods have been widely used for conformational analysis of small molecules [18, 19]. To enhance conformational sampling, a useful tactic in MD is to apply an elevated temperature to the simulation [20, 21]. A simplified conformational search strategy (Quenched Molecular Dynamics, QMD) that has been tested and reported [16, 22] was adopted [23, 24].

A starting geometry for each ligand was constructed using the Accelrys pre-existing fragment library and fully optimized with the ESFF force field [25] (with force field formal atomic charges) to remove unfavorable steric interactions. High-temperature MD calculations were carried out at 2000 K with the starting velocities calculated from the Boltzmann distribution. Each simulation was run for 100 ps with steps of 1.0 fs, with coordinates saved every 0.05 ps, resulting in 2000 conformations. Finally, 10% of all conformers were randomly selected and saved in a database, ultimately containing 200 conformers.

All conformers in the database were subjected to a two-step energy minimization using the same force field as for the MD calculations. A Steepest-Descent algorithm was first used, with convergence obtained at 0.05 RMSG, followed by a Newton–Raphson algorithm with convergence obtained when the gradient was below 0.001 RMSG.

The conformational similarity of the 200 energy-minimized conformers was investigated by comparing all pairs of conformers. The two comparison criteria were force field energy and root-mean-square distance (RMSD) calculated over all heavy atoms and polar H-atoms. An *ad hoc* Fortran program then calculated the mean and standard deviation of the RMSD values. Two conformers were considered identical when their difference in energy was below 3 kcal/mol and their RMSD less than or equal to the RMSD mean minus the standard deviation. In the case of identical geometries, the higher energy conformer was eliminated from the database.

For each ligand, various starting geometries were used to ensure a proper exploration of conformational space. All minimization and dynamics calculations were performed in the absence of water, assuming a dielectric constant (ε) equal to 1. Finally, two minimum energy conformations with significant geometrical differences were selected for submission to the docking procedure.

Preparation of the COX-2 model

Selection of the COX-2 structures

COX-2 enzyme contains 587 amino acids and one ferric heme group [26]. The X-ray structure of mouse (*Mus musculus*) COX-2 has been resolved as uninhibited, as well as in the presence of four inhibitors (the PDB code and the resolution of the crystal data are reported in Table 1). The amino acid sequence of COX-2 is conserved between species with 85–90% sequence similarity. The structure of human COX-2 is expected to be very similar to the mouse enzyme [6], since the two have 87% identity with strict conservation in the active site [27]. The mouse COX-2 may thus be taken as a model for the human COX-2 enzyme.

The COX-2 protein exists *in situ* as a dimer; however, the reason for dimerization is not known [26]; the monomer structure alone has always been considered in molecular modeling studies [8, 9] assuming the interactions governing the COX-2 inhibitor binding to be reproducible using one monomer.

COX-2 crystallizes with each monomer consisting of three independent folding units:

(i) An N-terminal epidermal growth factor domain (EGF);
(ii) A membrane-binding motif;

Table 1. X-ray structures for COX-2 isoenzymes.

Ligand	Resolution [Å]	PDB code	Reference
Uninhibited	3.0	5COX	[6]
Flurbiprofen (**1**)	2.5	3PGH	[6]
Indomethacin (**2**)	2.9	4COX	[6]
Diclofenac (**3**)	2.9	1PXX	[36]
SC-558 (**4**)	3.0	1CX2	[6]
SC-558 (**4**)	2.8	6COX	[6]

(iii) A C-terminal catalytic domain, which contains the active cyclooxygenase and peroxidase active sites.

The superposition of the α-carbons of the entire enzyme for all residue pairs of the considered models (inhibited and uninhibited), yields a RMSD ranging from 0.35 to 0.56 Å. When the superposition is limited to the subset of residues closely surrounding the inhibitor molecule (residues within 5 Å of the inhibitor), the RMSD ranged from 0.35 to 0.58 Å. A visual inspection of the models revealed that the inhibition does not produce significant changes in the conformation of the residues around the inhibitors; though the conformation of Arg120 appeared to be strongly influenced by the inhibitor interaction. In the complex SC558/COX-2, the X-ray data do not allow unique definition of the conformation of the sulphonamide [6], and the conformation depends on the X-ray structure (6COX and 1CX2).

The binding of inhibitors does not greatly perturb the resting structure [6], as is shown by the fact that the greatest RMSD deviation for all Cα atoms between the X-ray structures reported in Table 1 is 0.56 Å. On this basis, the choice of model to use as starting point for the study does not appear to be relevant, and we therefore used the structure 6COX, since the molecule SC558 (a typical COX-2 selective inhibitor) can be used as a validation system.

Preparation of the enzyme structure

The mouse COX-2 structure complexed with **4** was used as enzyme template. All the structures crystallize as dimers and thus the A chain alone was selected for the calculations. The residues of *N*-acetyl-D-glucosamine (NAG) and *β*-octylglucoside (*β*-OG) were eliminated, as they are not relevant for the enzyme binding properties. Bound ligands were also removed from the template.

Adding the hydrogen atoms in their standard geometry completed the structures of the two enzymes. The complete system was then minimized following the procedure described below.

Docking protocols

Docking protocols can be described as a combination of two components: a search strategy and an evaluation of the docking results (the definition of a scoring function [10]). The search algorithm should generate an optimum number of poses that include the experimentally determined binding mode. We applied two different procedures to evaluate the influence of the search method on the final results.

Electrostatic interactions were treated with a distance-dependent dielectric constant 4ε as reported in the literature [28].

Systematic perturbative protocol

Search procedure. For calculations we employed a general force field implemented in the Accelrys Discovery program and we developed a simple docking procedure that can be summarized in three steps:

(i) The ligand obtained from QMD (two conformations, see above) is manually positioned in the binding site;
(ii) To obtain different poses, the ligand is slightly and rigidly moved (2 Å in each direction) within the binding site by translating its center of mass along the *x*, *y* and *z* axes; at each position the ligand is rotated around its principal axes in 60° steps;
(iii) The non-bonded energy (electrostatic energy + Van der Waals energy) was evaluated for each orientation and the 10 poses with the lowest non-bonded energies were selected and further minimized.

Minimization procedure. Energy minimization of proteins is a cpu-demanding process and the demand increases rapidly with the number of atoms. Thus, a number of minimization strategies were tested, considering either the complete enzyme structure (optimal accuracy) or a subset of residues (best speed).

After a number of assays, the following compromise was found: minimization of the starting structure prepared from PDB data was complete, but minimization of the 10 candidates obtained from the search procedure applied to the optimized model was carried out using the residue subset for which at least one atom in the residue was within 20 Å of the ligand atoms, a cut-off used by other workers [29].

To avoid artificial movements away from the original structure due to large initial forces, the

protein model was gradually relaxed [30]. Regardless of the number of atoms considered, the minimization procedure was stepwise: (a) geometry optimization of all hydrogen atoms, keeping the rest of the structure fixed; (b) geometry optimization of hydrogen atoms and side-chain atoms, keeping the backbone atoms fixed; (c) geometry optimization of all atoms except for Cα carbons; (d) complete geometry optimization without constraint.

Selection of poses. The final poses obtained by minimizing the ten complexes generated by the search procedure were checked for their equivalence; in particular, two poses were considered equivalent if the RMSD of heavy atoms was less than 0.3 Å [31].

Genetic Algorithm protocol

Search procedure. Energy calculations were performed using the MMFF94x [32] force field implemented in MOE. A modified MOE-Dock module, kindly provided by the Chemical Computing Group, was used to perform an automatic docking exploration for different positions and orientations. A Genetic Algorithm search procedure was used with MOE default parameters (1500 generations per run, three mutation frequencies, seven birthrates). The ligand flexibility was taken into account, and a random initial orientation was used.

Selection of poses. The final geometries were sorted using an energy criterion, and the RMSD for all atoms was calculated between poses. Two poses were considered identical when their difference in RMSD was below 0.3 Å.

Energy minimization. Despite the cpu request, in this case we minimized the complete protein to better differentiate the two search procedures.

Energy minimization was carried out following a stepwise procedure:

(i) geometry optimization of all hydrogen atoms while keeping the rest of the structure fixed;
(ii) geometry optimization of hydrogen and side chain atoms while keeping the backbone atoms fixed;
(iii) geometry optimization of all atoms except Cα carbons;
(iv) geometry optimization of protein areas with energy gradient components greater than 1 kcal mol^{-1} $Å^{-1}$;
(v) full geometry optimization without constraints.

Contact Statistics analysis

The docking results obtained with the above two search protocols were analyzed using a combination of MOE applications [16].

An empirical scoring tool was first applied to evaluate the strength of the hydrogen bonds and salt bridges [33]. Each hydrogen bond contact shown in the MOE has an attached label, e.g. '90%, 2.9 Å', which indicates 'hydrogen bond at 90% strength and a distance of 2.9 Å'. The strength of a contact is given relative to an 'ideal' hydrogen bond of that type, as computed by the scoring function.

MOE's Contact Statistics tool [33] was then used to determine the preferred locations of the ligand atoms. The Contact Statistics tool calculates, from the 3D atomic coordinates of the protein atoms, preferred locations for hydrophobic and hydrophilic ligand atoms based on statistics obtained from the crystallographic structures of receptor-ligand complexes [14, 15]. In MOE, the preferred locations of bound atoms are estimated from a collection of all crystallographic structures in the Protein Data Bank with a resolution of 2.0 Å or better [34].

The Contact Statistics tool produces two preference maps, (a) a green surface which is a 90% probability isocontour for hydrophobic atoms, and (b) a red surface, which is a 90% probability isocontour for polar atoms. This means that a ligand fits the maps well if its hydrophobic atoms are located in proximity to the green surfaces and its polar atoms are located in proximity to the red surfaces. When the ligand matches the preference maps, a good interaction may be expected.

Results and discussion

To facilitate reading, the X-ray data for COX-2 inhibitors complexes (see above) were inspected to find the interaction keys, as reported in the first part of this section; the same X-ray data were then used to validate the Contact Statistics tool.

The poses generated by the search procedure in the Contact Statistics analysis were ranked with the aim of selecting the most probable ones, as reported below. In particular, this procedure was applied to two experimentally determined NSAID/COX-2 complexes, flurbiprofen/COX-2 and SC558/COX-2. To determine how closely the predicted pose resembles that observed in the experimental complex structure (an efficient search procedure should correctly reproduce X-ray data [12]) the RMSD of non-hydrogen atomic positions of the ligands was used, this being the most straightforward method to evaluate a prediction in terms of docking accuracy [35, 17].

Finally, the study was extended to two important NSAIDs, nimesulide and rofecoxib, for which the X-ray structures have not been resolved, and computational studies have indicated different proposals for their interactions with COX-2 [26].

Analysis of X-ray data

Definition of the interaction keys

An examination of the five X-ray structures of the inhibited COX-2 enzyme (Table 1) reveals some important interaction keys for the binding of COX-2 inhibitors (Figure 2). At the outset, three main cavities can be described:

(i) The NSAIDs considered are bound at varying depths within a cavity (*Cavity A*) formed by Phe381, Tyr385, Leu384, Trp387 and Phe518;

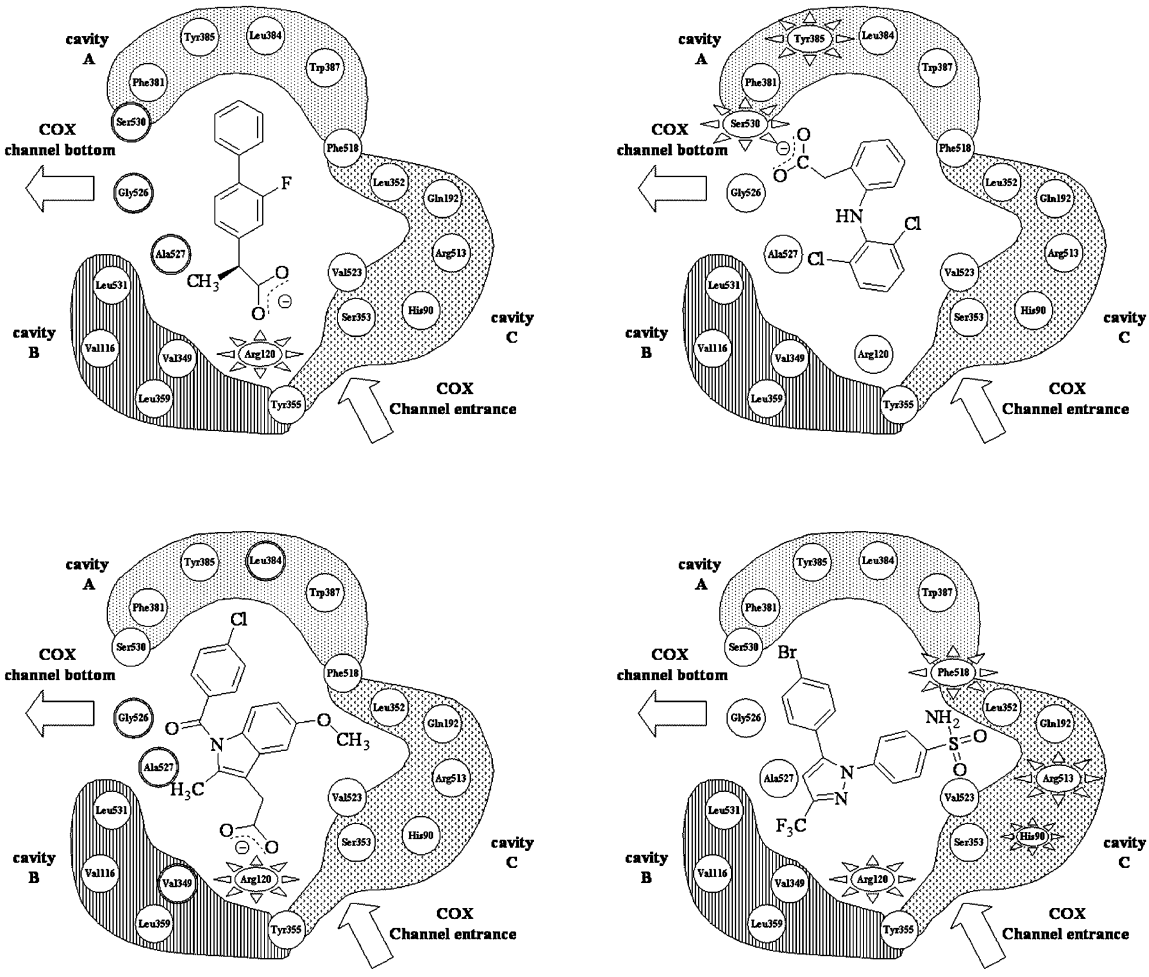

Figure 2. Schematic representation of the ligand/COX-2 keys as inferred from X-ray data for **1**, **3**, **2** and **4**.

(ii) The carboxylate group of flurbiprofen, indomethacin and the trifluomethyl group of SC-558 are located in a cavity (*Cavity B*) formed by Leu531, Val116, Val349, Leu359 and Tyr355;
(iii) Val523 permits the location of the SC-558 phenylsufonamide and of the six-member ring of indomethacin indole moieties in a cavity (*Cavity C*) formed by Phe518, Leu352, Val523, the backbone of Ser353 and Tyr355.

Furthermore, there are also some important interactions:

(i) A salt bridge formed by the carboxylate of **1** and **3** with Arg120 gives traditional NSAIDs the same anchor point in COX-1 and COX-2 isozymes and thus limits their selectivity due to the ligand's decreased freedom of movement;
(ii) In **1**, a hydrogen bond with Tyr355 reinforces the interaction of the inhibitor with the enzyme;
(iii) Diclofenac binds to COX-2 in an inverted conformation with its carboxylate group hydrogen-bonded to Tyr385 and Ser530;
(iv) **1** and **4** interact with residues that are part of the COX channel wall, such as Gly526 and Ala527;
(v) Beyond Cavity C, the SC-558 sulphonamide group interacts with His90, Arg513 and Gln192 forming a hydrogen bond. This interaction is a key point for COX-2 selectivity.

In line with previous observations, ligand–enzyme interaction keys are therefore defined as interactions occurring in cavities A, B and C, plus the salt bridge described above. Throughout this report, analysis and interpretation of docking results are based on these interaction keys.

Contact Statistics analysis applied to X-ray data

The Contact Statistics tool was initially applied to X-ray data to validate the procedure in the case of COX-2 inhibitors.

In the flurbiprofen complex (Figure 3) only the distal phenyl ring is well surrounded by a hydrophobic Contact Statistics grid, whereas the carboxylic group is located in a preferred hydrophilic Contact Statistics region. The carboxylic group can form a salt bridge with Arg120 and a hydrogen bond with Tyr355, as in the COX-1 enzyme. The remaining structural moieties (the methyl and fluorophenyl groups) are not involved in other interactions. This analysis suggests that in the flurbiprofen/COX-2 complex, **1** does not take full advantage of all points of contact with the receptor.

Indomethacin is a stronger COX-2 inhibitor than flurbiprofen [6] and this is likely due to an increment in some non-specific interactions (Figure 3). The carboxylic group can form a salt bridge (as **1**) though the bromophenyl ring is located deeper within Cavity A than the flurbiprofen distal phenyl ring, and is surrounded by the hydrophobic preferred Contact Statistics grid. Some carbon atoms of the six-member indole ring are located in a region of hydrophobic preferred contacts, producing additional favorable interactions. Finally, indomethacin penetrates, though not deeply, in Cavity C, but the methoxy group is located in non-favorable regions.

Diclofenac is more active against COX-2 than several other carboxylic acid-containing NSAIDs [36]. Furthermore, it exhibits a distinctive binding mode compared to several other arylacetic acid-type inhibitors when complexed to COX-2 (Figure 3). This compound binds in Cavity A with its carboxylate group adjacent to Tyr385 and Ser530: the carboxylate is located in a region of preferred hydrophilic contacts and forms hydrogen bonds with the OH groups of Tyr385 and Ser530. The dichlorophenyl group projects down into the main channel of the active site and is surrounded by hydrophobic preferred Contact Statistics regions, as is the benzyllic acid group, the remaining nitrogen seems not to be involved in particular interactions. In the COX-2 binding site diclofenac is able to optimize all the interactions that it can form, hence it is not surprising that this compound is the most active non-selective NSAID against COX-2 of the three considered.

Finally, Contact Statistics analysis was applied to the inhibitor SC558, a highly selective COX-2 inhibitor closely analogous to celecoxib. The bromophenyl ring is located in Cavity A (Figure 3), and is closely surrounded by preferred hydrophobic Contact Statistics grids, as shown by indomethacin and flurbiprofen. The fluoromethyl group faces Arg120 and one of the fluorine atoms is located in a hydrophilic region. The pyrazole ring is located with a nitrogen adjacent

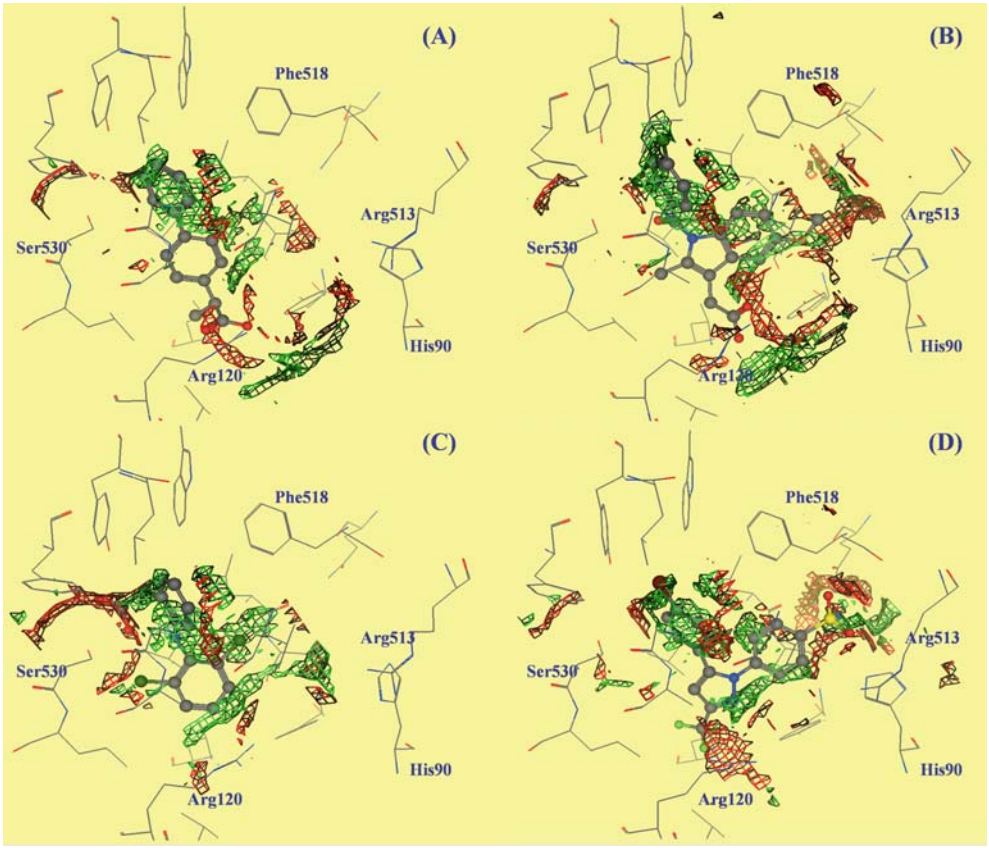

Figure 3. The analysis of crystallographic data of four COX-2 inhibitors bound to the enzyme is reported. The preference maps are also reported: the hydrophilic preference map (at 90% probability) is colored in red, the hydrophobic preference map (at 90% probability) is colored in green. (A) Flurbiprofen (3PGH); (B) Indomethacin (4COX); (C) Diclofenac (1PXX); (D) SC-558 (6COX).

to a preferred hydrophobic Contact Statistics region. A hydrophobic preferred Contact Statistics region surrounds the aromatic ring of the phenylsulphonamide group. Finally, the sulphonamide group is located in a hydrophilic preferred Contact Statistics region, and can form hydrogen bonds between the sulphonamide nitrogen and the nitrogen of the Gln192 amide group, whereas a sulphonamide oxygen can form another hydrogen bond with the nitrogen of Phe518. The conformation of the SC558 phenylsulphonamide group is not unique: the angle τ_1 of the sulphonamide group in the crystal model 1CX2 is rotated by about 120° compared to the model 6COX and is less available for hydrogen bonding.

In conclusion, the use of Contact Statistics analysis can be useful to interpret experimental data. Comparison between different molecules is difficult, because it is not possible to quantify the weight of favorable and unfavorable interactions: however, it is possible to evaluate whether the molecule can potentially exploit all its interaction capacity, as in the case of diclofenac.

Flurbiprofen and SC558 docking

Flurbiprofen

Flurbiprofen (**1**), a profen-derivative, is a classic NSAID [1] bearing a carboxylic group. **1** is a COX-1 selective inhibitor, pharmacologically endowed with undesired gastrointestinal (GI) side-effects. The profens S-enantiomers preference for COXs was already investigated by Llorens et al. [8] and does not require further comment.

Because of its pK_a ($pK_a = 4.27$ [37]) the electrostatics responsible for interaction of **1** with enzymes is unknown, and it is also impossible to

deduce it from X-ray data (see above). Thus, both neutral and anionic forms of flurbiprofen were submitted to conformational analysis and docking studies, but no significant differences emerged.

Flurbiprofen has essentially two degrees of freedom: the torsional angle between the two phenyl moieties (τ_1 in Figure 1) and the dihedral angle between the isopropyl moiety and the phenyl ring (τ_2 in Figure 1). Four conformations were obtained from conformational analysis regardless of the protonation state: two forms have τ_1 equal to 39°, the other two have τ_1 equal to −39°. The two conformations with the same τ_1 differ in the orientation of the carboxylic group, which is either directed toward the aromatic fluorine atom or flipped by 180°. The four conformations have roughly the same energy and it is reasonable to suppose that all of them are accessible. For docking studies the two conformers (see Experimental section) at lowest energy were selected.

Five poses (P1, P2, P3, P4 and P10) were found within 1 kcal/mol of the minimum and can be grouped into three clusters. In P10, the pose with the lowest non-bonded energy, the isopropionic moiety occupies Cavity C. Pose P4, which is at higher energy (about 0.6 kcal/mol), correctly reproduces the experimental pose, despite the RMSD from the X-ray solution (1.2 Å). Finally, in the last three poses (P1, P2 and P3) the isopropionic group is located in Cavity A.

Contact Statistics analysis enabled two of the three poses to be eliminated from the third cluster: in P1 and P2, the carboxylic group is clearly located in a region with preferred hydrophobic interactions and is not able to form hydrogen bonds.

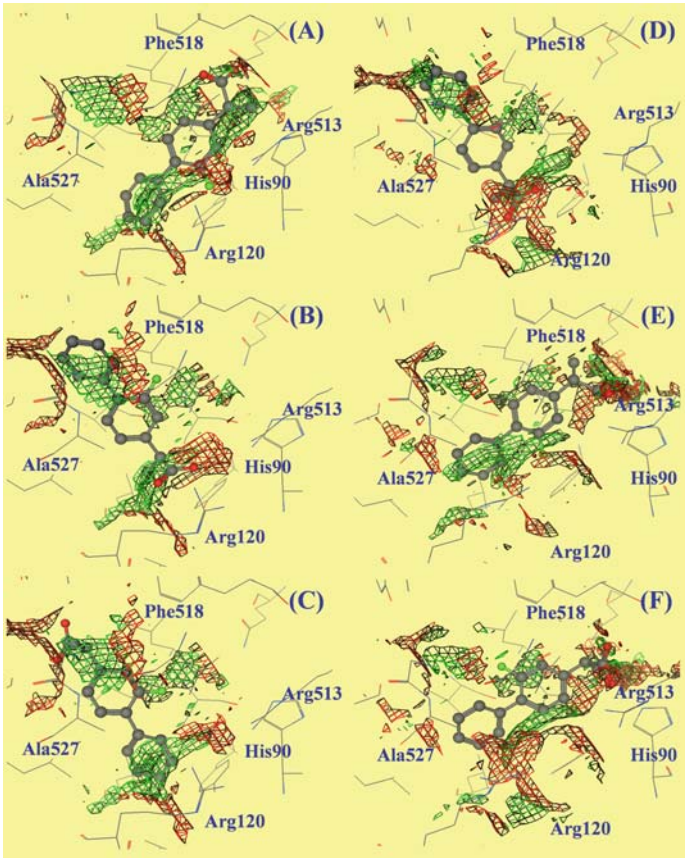

Figure 4. The results of the two search protocols applied to the flurbiprofen/COX-2 complex are reported. The preference maps are also reported: the hydrophilic preference map (at 90% probability) is colored in red, the hydrophobic preference map (at 90% probability) is colored in green. (A) Pose P10 obtained with the systematic perturbative protocol; (B) pose P4 obtained with the systematic perturbative protocol; (C) pose P3 obtained with the systematic perturbative protocol; (D) pose PGA5 obtained with the GA protocol; (E) pose PGA3 obtained with the GA protocol; (F) pose PGA4 obtained with the GA protocol.

In pose P10 (Figure 4), the isopropionic moiety forms hydrogen bonds with the nitrogen of Ile517 and the methyl group interacts with His90. The fluorine atom of the fluorophenyl group is close to Tyr355 and Arg120, whereas the distal phenyl is close to the backbone of Val116 and the hydroxyl of Tyr355. This is the most unusual pose, although the carboxylic moiety is located in a hydrophilic region, as is the case with the sulphonamide moiety of the COX-2 selective ligand SC558. Contact Statistics analysis shows that the distal phenyl and part of the isopropionic group are closely surrounded by a green surface indicating a favorable hydrophobic interaction. The carboxylic moiety and the fluorine atom are located in a red surface.

The second acceptable solution (P4) correctly reproduces the interaction keys (Figure 4); the high RMSD is likely due to the shift of Arg120, which permits formation of the salt bridge as in the X-ray data. The hydrogen bond with Tyr355 is also present. Contact Statistics analysis gives essentially the same results as analysis of the X-ray data.

In the last docked solution (P3), the isopropionic carboxylic group can form two hydrogen bonds with Ser530 and Tyr385. The distal phenyl ring is close to Tyr355 hydroxyl whereas the fluorophenyl ring is located in a region without any particular interactions. Contact Statistics analysis confirms this picture: the methyl and carboxylic isopropionic groups are located in favorable regions as part of the distal phenyl group. The fluorine atom is close to a green region. Finally, the aromatic ring of the fluorophenyl group is distant from any region. It is noteworthy that in the crystallographic solution of the complex diclofenac/COX-2 the carboxylic moiety of the ligand is located in this region and interacts with Ser530.

Other poses exist with non-bond energy that differs more than 1 kcal/mol from the minimum, but Contact Statistics analysis rejected them. This occurred for all complexes, and thus only poses within 1 kcal/mol of the minimum were considered.

In conclusion, the first docking procedure can explore extensive regions of the binding site. In addition, in all poses two out of three cavities were occupied by flurbiprofen, which can bind COX-2 in different ways without occupying the site efficiently.

The analysis obtained by generating the poses with the GA protocol produced five poses (PGA5, PGA1, PGA3, PGA4 and PGA10) within 5 kcal/mol of the minimum. The poses with non-bond energy that differed by more than 5 kcal/mol from the minimum were rejected, because Contact Statistics revealed unfavorable contacts for all complexes considered. Poses P1 and P2 were among the rejected poses. The rejected PGA13 is similar to P3 but the molecule is shifted and the carboxylic group is located in a green region and cannot form a hydrogen bond.

Pose PGA10 placed the methyl and the carboxylic group in unfavorable Contact Statistics regions and was thus rejected.

The pose with the lowest non-bond energy (PGA5) is that with the lowest RMSD with regard to the X-ray solution (0.75 Å) and the Contact Statistics conclusions are the same as those obtained with pose P4 (Figure 4). Pose PGA1 differs from PGA5 in the orientation of the fluorine atom: the fluorophenyl ring is rotated by 180° with respect to the X-ray solution. Contact Statistics shows that the fluorine atom is located near a red region, so this pose might be less favorable than PGA5 [38].

Poses PGA3 and PGA4 (Figure 4) are close to pose P10 with the isopropionic group located in Cavity C. In PGA3 and PGA4 the carboxylic group is located near Arg513 and His90 in a region with red, favorable, and green, unfavorable, grids. The carboxylic moiety can form two hydrogen bonds with Arg513 and His90 in PGA4 and one hydrogen bond with Arg513 in PGA4.

In conclusion, the GA protocol coupled with Statistics Analyses is able to reproduce the X-ray solution and confirms that **1** can bind at the COX-2 site in different ways.

Comparison of the two search methods suggests that the X-ray solution can be found independently of the search protocol, whereas the alternative binding modes may depend on the choice of search protocol and on the choice of force field. In our opinion, in this case the lack of independence may be ascribed to the low specificity of flurbiprofen, which can occupy two of the three cavities, producing reasonable interactions as indicated by the Contact Analyses.

SC558

SC-558 (**4** in Figure 1) is a highly selective COX-2 inhibitor belonging to the vicinal diaryl heterocyclic class [1]. These compounds are characterized

by a central carbocyclic or heterocyclic ring system bearing two vicinal aryl moieties. For optimal activity, one aromatic ring must have a methylsulphone or a sulphonamide substituent in the para position [39]. The central ring is responsible for the appropriate orientation of the aromatic rings in space and ultimately for binding to the enzyme [1]. A wide variety of heterocycles can serve as templates for the COX-2 inhibitor, but pyrazole and cyclopentenone now appear to be the most appropriate to achieve COX-2 selectivity [1].

SC-558 has three degrees of freedom: the two torsional angles separating the three rings (τ_2 and τ_3 in Figure 1) and the angle between the sulphonamide group and the phenyl ring (τ_1).

The QMD procedure applied to SC-558 yields two conformers with the same energy (−41.4 kcal/mol). In the first conformer (similar to the X-ray data) the phenyl carrying the sulphonamide group is roughly coplanar with the pyrazole ring ($\tau_2 = -18.6°$); the sulphonamide moiety is roughly perpendicular to the plane of the pyrazole heterocyle ($\tau_1 = 102.3°$). Finally, the bromophenyl group forms an angle τ_3 equal to −45.8 with the heterocycle. The second conformer can be viewed as the mirror image of the first ($\tau_1 = -102.3°$; $\tau_2 = 18.6°$; $\tau_3 = -45.8°$).

Three poses (P1, P4 and P6) were found within 1 kcal/mol of the minimum, however, Contact Statistics analysis immediately suggested rejecting pose P4 because of the many unfavorable interactions. P6 is the pose with the lowest interaction energy and the lowest RMSD (0.64 Å) from the X-ray solution (6COX), whereas P4 differs only in the sulphonamide orientation which is similar to that shown in the 1CX2 structure.

Contact Statistics analysis for P6 and P4 did not differ from analysis of the X-ray solution; in P6, the sulphonamide is better able to take advantage of its hydrogen bond capabilities.

The analysis obtained generating the poses with GA produced four poses (PGA4, PGA6, PGA12 and PGA17) within 5 kcal/mol of the minimum. Contact Statistics enabled PGA4, PGA6 and PGA12 to be eliminated. PGA17 is the pose with minimum energy, the closest RMSD to the X-ray solution (6COX) and corresponds to the pose P6.

In this case the procedure is able to unambiguously determine the correct pose, thus, reproducing the X-ray data despite the search procedure applied.

Rofecoxib and nimesulide docking

Rofecoxib

Rofecoxib (Vioxx®) also belongs to the vicinal diaryl heterocyclic class of cyclooxygenase inhibitors. In **5**, the central ring is a cyclopentenone with a 4-(methylsulphonyl)phenyl moiety and a phenyl ring attached to it. Rofecoxib is a highly selective COX-2 inhibitor, endowed with three degrees of freedom: the two angles between the three rings (τ_2 and τ_3 in Figure 1) and the angle between the methylsulphonamide group and the phenyl ring (τ_1).

The QMD procedure applied to **5** gave 4 conformers, which may be grouped into two classes of 2 conformers each. The first class is at lower energy than the second and both conformers are superimposable on those of SC-558 obtained by QMD, with slight differences: in the first conformer the methylsulphone group is approximately perpendicular to the phenyl plane ($\tau_1 = 86.6$), the phenyl plane forms an angle of −38.6° (τ_2) with the cyclopentenone ring and finally there is an angle of −35.2° (τ_3) between the central ring and the phenyl ring; the second conformer is the mirror image of the first ($\tau_1 = -86.6°$; $\tau_2 = 38.6°$; $\tau_3 = 35.2°$).

The docking results for rofecoxib give two solutions: the lowest non-bond energy P5 and pose P6 at higher energy (about 0.5 kcal/mol). In P5, the methylsulphone moiety is roughly superposed to the SC558 sulphonamide group and is located in Cavity C close to Phe518, His90 and Asp515. A sulphone oxygen can form hydrogen bonds with the backbone nitrogen of Phe518 and Ile517. The phenyl group is located in Cavity A and occupies a position comparable to that of the SC558 bromophenyl ring with close contact with the backbone atoms of Val523. Finally, the furanone moiety is roughly superposed to the SC558 pyrazole ring with close contact with the Val349 and Leu531 side-chains.

In P6, the ligand is oriented in a completely different manner: the phenyl ring is directed toward the entrance of the hydrophobic channel connecting with the active site of the enzyme, the methylsulphone moiety is located in Cavity A and furanone group is close to the entrance of Cavity C with the exo-cyclic oxygen interacting with Arg120 (3.1 Å) and Tyr355 (3.1 Å). In P6, the methylsulphone moiety is close to Ala527 and Gly526

backbone atoms and Trp385 and Ser530 side chains. The phenyl ring is close to Tyr355, Leu531 and Val116 side-chains.

In P5, Contact Statistics analysis shows at least three region of favorable contact: the phenyl group in Cavity A, as shown with other ligands, the methylsulphone phenyl ring and the major part of the sulphone moiety. One of the two oxygens of the sulphone group is in a hydrophobic preferred contact region.

Contact Statistics analysis revealed that P6 is a reasonably good pose because the phenyl group is again surrounded by a hydrophobic preferred regions at the entrance of the hydrophobic channel that leads to the active site. The exo-cylic oxygen is located in a hydrophilic preferred contact region; unfavorable regions are occupied by the methylsulphone moiety and partially by the furanone group.

Although P6 cannot be rejected, pose P5 is better located and is the one that permits rofecoxib to best exploit its interaction capability. In P5 the methylsulphone moiety is located near Arg513, in accordance with the major role played by residue Arg513, as indicated in the literature [6, 8, 26].

The GA protocol produced two poses, but Contact Statistics indicated that only PGA2, the pose with the lowest non-bond energy, is acceptable. PGA2 corresponds to pose P5. In the model obtained with MOE the methylsulphone group and Arg513 are closer than in the results obtained with the Accelrys package. Thus a hydrogen bond is possible between the methylsulphone group and the Arg513 side chain, showing the importance of Arg513 in rofecoxib/COX-2 binding.

In conclusion, the results obtained with the two procedures suggest that rofecoxib binds in the COX-2 site in pose P5 (or PGA2) (Figure 5).

No X-ray structure of the complex rofecoxib/COX-2 is currently available in the literature so far, but Llorens et al. [8] docked rofecoxib in the active site of COX-2. Site mutagenesis studies have highlighted the role of Arg513 in the interaction: the authors manually docked the ligand, placing the sulphone in the vicinity of Arg513 and the carbonyl oxygen close to the side chain of Ser530. These interaction keys are present in P5 (PGA2), confirming the consistency of our findings.

Nimesulide

Nimesulide belongs to the class of diaryl- and aryl-heteroaryl ethers [1] and is a selective COX-2 inhibitor (**6** in Figure 1). Its pharmacology, together with clinical results in rheumatic disease, osteoarthritis and acute inflammation, shows it to possess novel anti-inflammatory qualities; the mechanism of action was clarified by the discovery of the COX-2 isozymes.

Nimesulide has four degrees of freedom, as shown in Figure 1 (τ_1–τ_4). The QMD procedure applied to **6** gave 8 conformers and, as discussed above, only two of them were submitted to docking runs. The two most stable conformers are equivalent in terms of energy and are mirror images of each other. In the first conformer, the four dihedral angles take the following values: $\tau_1 = -92.8°$, $\tau_2 = 13.2°$, $\tau_3 = 58.2°$ and $\tau_4 = 22.5°$. In the second conformer the dihedral angles have the opposite values ($\tau_1 = 92.8°$, $\tau_2 = -13.2°$, $\tau_3 = -58.2°$ and $\tau_4 = -22.5°$).

Because of its presence at the pH of interest ($pK_{a\,=}$ 6.56 [40]), anionic nimesulide was also submitted to conformational analysis; the results were very similar to those of the neutral version.

The results of the docking procedure gave two poses: the low non-bonded energy pose P9 and pose P1, with a higher energy of about 0.5 kcal/mol.

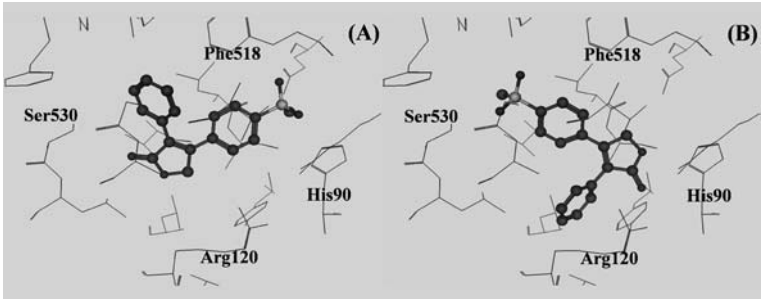

Figure 5. The results of the docking procedure applied to the rofecoxib/COX-2 complex are reported. (A) Pose P5; (B) pose P6.

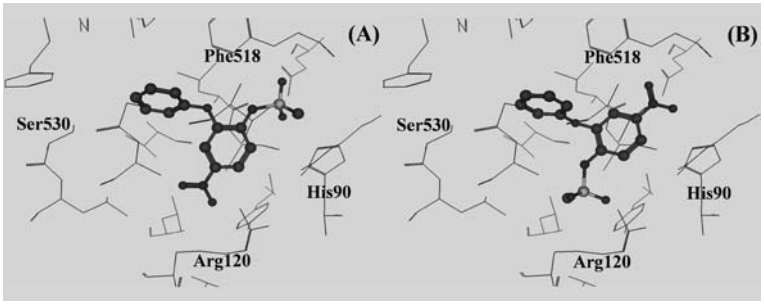

Figure 6. The results of the docking procedure applied to the nimesulide/COX-2 complex are reported. (A) Pose P1; (B) pose P9.

In P9, the nitro group is located in Cavity C close to the Phe518, Ala516 and His90 residues. The methanesulphonamide group is directed towards the channel where the active site is located. It forms a hydrogen bond with the hydroxyl group of Tyr355, it is located close to Ala527 and Arg120 (about 3.1–3.2 Å).

In pose P1, the methanesulphonamide is located in Cavity C and interacts with the backbone nitrogens of the residues Phe518 and Ile517 and with the chain nitrogen of Gln192. The nitro group is directed towards the active site hydrophobic channel.

In both poses, the phenoxy ring occupies a position comparable to that of the bromophenyl ring of SC558.

Contact Statistics analysis revealed that the phenoxy ring was located in a hydrophobic preferred contact region in both poses. In P1, the nitro group was surrounded by an unfavorable hydrophobic preferred contact grid and it was distant from Arg120, whereas the methanesulphonamide group was nicely located, suggesting the presence of favorable interactions. In pose P9, the nitro group was located in a region with potential favorable and unfavorable interactions, but the central phenyl ring appeared very well located with the nitrogen of the methanesulphonamide group in a region without particular interactions. Finally, in this pose the methanesulphonamide group was well located with the potential to interact favorably with Arg120 and Tyr355.

All poses produce both favorable and unfavorable interactions. In particular, in P9 the methanesulphonamide can interact nicely with Arg120 and Tyr355 and the nitro group is better located than in P1. In conclusion, P9 is the most favorable pose for nimesulide (Figure 6).

The GA protocol produced five poses but only two are retained on the indications of Contact Statistics: the lowest non-bond energy pose PGA2, corresponding to pose P9, and pose PGA1, corresponding to pose P1.

Other molecular modeling experiments [26] with nimesulide have led to different proposals for its interaction with COX-2. Garcia-Nieto et al. [9] used the full range of possible conformers of nimesulide for their docking and molecular dynamics simulations. Two alternative binding modes were proposed: in the first the nitro group interacts with Arg120 and the methylsulphone group points towards Arg513 (corresponding to P1 or PGA1), in the second the methylsulphone group is placed close to Arg120 and the nitro group is close to Arg513 (corresponding to P9 or PGA2). The latter is consistent with pose P9 (see above), with published docking studies [26], with unpublished X-ray data [41] and mutational analysis [42].

Conclusions

This study furnishes a number of noteworthy results.

The study on rofecoxib confirms that COX-2 selective tricyclic inhibitors occupy all three cavities. In particular, Contact Statistics analysis shows that, in the area occupied by the sulphone moiety, there are both regions of preferred hydrophilic interactions and regions of preferred hydrophobic interactions. It is therefore reasonable to suppose that substitution of the sulphonamide moiety with the methylsulphone group does not reduce the activity of the ligand (e.g. celecoxib and SC-58125) [43].

Moreover, some relevant indications emerge concerning the probable binding orientation of rofecoxib [8] and nimesulide [9], for which X-ray data are not yet available.

To sum up, Contact Statistics analysis coupled with different docking search strategies can provide detailed information about interaction of the ligand with the binding site, and determine the most reasonable orientation among a number of low energy poses.

References

1. Dannhardt, G. and Kiefer, W., Eur. J. Med. Chem., 36 (2001) 109.
2. Kurumbail, R.G., Kiefer, J.R. and Marnett, L.J., Curr. Opin. Struct. Biol., 11 (2001) 752.
3. Marnett, L.J. and Kalgutkar, A.S., Trends Pharmacol. Sci., 20 (1999) 465.
4. Hinz, B. and Brune, K., J. Pharmacol. Exp. Ther., 300 (2002) 367.
5. Loll, P.J., Garavito, R.M., Carrell, C.J. and Carrell, H.L., Acta Crystallogr. Sect. C – Cryst. Struct. Commun., 52 (1996) 455.
6. Kurumbail, R.G., Stevens, A.M., Gierse, J.K., McDonald, J.J., Stegeman, R.A., Pak, J.Y., Gildehaus, D., Miyashiro, J.M., Penning, T.D., Seibert, K., Isakson, P.C. and Stallings, W.C., Nature, 384 (1996) 644.
7. Flower, R.J., Nat. Rev. Drug Discov., 2 (2003) 179.
8. Llorens, O., Perez, J.J., Palomer, A. and Mauleon, D., J. Mol. Graph. Model., 20 (2002) 359.
9. Garcia-Nieto, R., Perez, C. and Gago, F., J. Comput.-Aided Mol. Des., 14 (2000) 147.
10. Taylor, R.D., Jewsbury, P.J. and Essex, J.W., J. Comput.-Aided Mol. Des., 16 (2002) 151.
11. Schulz-Gasch, T. and Stahl, M., J. Mol. Mod., 9 (2003) 47.
12. Wang, R.X., Lu, Y.P. and Wang, S.M., J. Med. Chem., 46 (2003) 2287.
13. Ferrara, P., Gohlke, H., Price, D.J., Klebe, G. and Brooks, C.L., J. Med. Chem., 47 (2004) 3032.
14. Muegge, I. and Martin, Y.C., J. Med. Chem., 42 (1999) 791.
15. Bruno, I.J., Cole, J.C., Lommerse, J.P.M., Rowland, R.S., Taylor, R. and Verdonk, M.L., J. Comput.-Aided Mol. Des., 11 (1997) 525.
16. Ermondi, G., Lorenti, M. and Caron, G., J. Med. Chem., 47 (2004) 3949.
17. Gohlke, H., Hendlich, M. and Klebe, G., J. Mol. Biol., 295 (2000) 337.
18. Christensen, I.T. and Jorgensen, F.S., J. Comput.-Aided Mol. Des., 11 (1997) 385.
19. Auffinger, P. and Wipff, G., J. Comput. Chem., 11 (1990) 19.
20. Holtje, H.D., Sippl, W. and Rognan, D., Molecular Modeling: Basic Principles and Applications, Wiley-CH, Weinheim, 2003, p. 35.
21. Caron, G., Gaillard, P., Carrupt, P.A. and Testa, B., Helv. Chim. Acta, 80 (1997) 449.
22. Altomare, C., Carrupt, P.A., Gaillard, P., El Tayar, N., Testa, B. and Carotti, A., Chem. Res. Toxicol., 5 (1992) 366.
23. Gaillard, P., Carrupt, P.A. and Testa, B., Bioorg. Med. Chem. Lett., 4 (1994) 737.
24. Ermondi, G., Caron, G., Bouchard, G., van Balen, G.P., Pagliara, A., Grandi, T., Carrupt, P.A., Fruttero, R. and Testa, B., Helv. Chim. Acta, 84 (2001) 360.
25. Accelrys, Extensible Systematic Force Field, San Diego, CA, 1994.
26. Trummlitz, G. and van Ryn, J., Curr. Opin. Drug Discov. Dev., 5 (2002) 550.
27. Filipponi, E., Cecchetti, V., Tabarrini, O., Bonelli, D. and Fravolini, A., J. Comput.-Aided Mol. Des., 14 (2000) 277.
28. Christensen, I.T. and Jorgensen, F.S., J. Biomol. Struct. Dyn., 15 (1997) 473.
29. Plount Price, M.L. and Jorgensen, W.L., J. Am. Chem. Soc., 122 (2000) 9455.
30. Holtje, H.D., Sippl, W. and Rognan, D., Molecular Modeling: Basic Principles and Applications, Wiley-CH, Weinheim, 2003, p. 129.
31. Sadowski, J., Gasteiger, J. and Klebe, G., J. Chem. Inf. Comput. Sci., 34 (1994) 1000.
32. Halgren, T.A., J. Comput. Chem., 17 (1996) 490.
33. Chemical Computing Group Inc., Version 2003.02 (2003).
34. Labute, P., J. Chem. Comput. Group, http://www.chem-comp.com/Journal_of_CCG/Features/cstat.htm (2003).
35. Dixon, J.S., Proteins Struct. Funct. Genet., S1 (1997) 198.
36. Rowlinson, S.W., Kiefer, J.R., Prusakiewicz, J.J., Pawlitz, J.L., Kozak, K.R., Kalgutkar, A.S., Stallings, W.C., Kurumbail, R.G. and Marnett, L.J., J. Biol. Chem., 278 (2003) 45763.
37. Rafols, C., Roses, M. and Bosch, E., Anal. Chim. Acta, 338 (1997) 127.
38. Ippolito, J.A. and Christianson, D.W., Int. J. Biol. Macromol., 14 (1992) 193.
39. Talley, J.J., Prog. Med. Chem., 36 (1999) 201.
40. Fallavena, P.R.B. and Schapoval, E.E.S., Int. J. Pharm., 158 (1997) 109.
41. Marnett, L.J., Rowlinson, S.W., Goodwin, D.C., Kalgutkar, A.S. and Lanzo, C.A., J. Biol. Chem., 274 (1999) 22903.
42. Greig, G.M., Francis, D.A., Falgueyret, J.P., Ouellet, M., Percival, M.D., Roy, P., Bayly, C., Mancini, J.A. and ONeill, G.P., Mol. Pharmacol., 52 (1997) 829.
43. Chavatte, P., Yous, S., Marot, C., Baurin, N. and Lesieur, D., J. Med. Chem., 44 (2001) 3223.