

Comparative residue interaction analysis (CoRIA): a 3D-QSAR approach to explore the binding contributions of active site residues with ligands

Prasanna A. Datar · Santosh A. Khedkar ·
Alpeshkumar K. Malde · Evans C. Coutinho

Received: 7 February 2006 / Accepted: 27 May 2006 / Published online: 29 September 2006
© Springer Science+Business Media B.V. 2006

Abstract A novel approach termed comparative residue-interaction analysis (CoRIA), emphasizing the trends and principles of QSAR in a ligand–receptor environment has been developed to analyze and predict the binding affinity of enzyme inhibitors. To test this new approach, a training set of 36 COX-2 inhibitors belonging to nine families was selected. The putative binding (bioactive) conformations of inhibitors in the COX-2 active site were searched using the program *DOCK*. The docked configurations were further refined by a combination of Monte Carlo and simulated annealing methods with the *Affinity* program. The non-bonded interaction energies of the inhibitors with the individual amino acid residues in the active site were then computed. These interaction energies, plus specific terms describing the thermodynamics of ligand–enzyme binding, were correlated to the biological activity with G/PLS. The various QSAR models obtained were validated internally by cross validation and boot strapping, and externally using a test set of 13 molecules. The QSAR models developed on the CoRIA formalism were robust with good r^2 , q^2 and r_{pred}^2 values. The major highlights of the method are: adaptation of the QSAR formalism in a receptor setting to answer both the type (qualitative) and the extent (quantitative) of ligand–receptor binding, and use of descriptors that account for the complete thermodynamics of the ligand–receptor binding. The

CoRIA approach can be used to identify crucial interactions of inhibitors with the enzyme at the residue level, which can be gainfully exploited in optimizing the inhibitory activity of ligands. Furthermore, it can be used with advantage to guide point mutation studies. As regards the COX-2 dataset, the CoRIA approach shows that improving Coulombic interaction with Pro528 and reducing van der Waals interaction with Tyr385 will improve the binding affinity of inhibitors.

Keywords CoRIA · COX-2 · Docking · G/PLS statistics · Simulations · QSAR

Introduction

The quantitative structure activity relationship (QSAR) approach has been one of the most successful tools in drug design. Since its first inception by Hansch et al. [1–3], the QSAR methodology has undergone revolutionary changes in every aspect of the approach, and the limitations inherent in the Hansch approach have spurred the development of the current 3D-QSAR methods.

QSAR analysis, whether 2D or 3D, has invariably been applied to data sets for which the geometry of the receptor is unknown. In cases where the geometry of the receptor is known, either by X-ray crystallography, NMR, or comparative modeling, docking and structure based drug design are preferred over QSAR analysis. Current 3D-QSAR techniques attempt to characterize ligand–receptor interactions indirectly using probes positioned at intersections of a lattice (grid or box) spanning a three dimensional region.

P. A. Datar · S. A. Khedkar ·
A. K. Malde · E. C. Coutinho (✉)
Department of Pharmaceutical Chemistry, Bombay College
of Pharmacy, Kalina, Santacruz (E), Mumbai 400 098, India
e-mail: evans@bcpindia.org

Usually, steric and electrostatic interactions between the probe and the ligand are calculated at the grid points. Comparative molecular field analysis (CoMFA) [4] was the first 3D-QSAR method developed based on this approach. Since then, several other methods like molecular shape analysis (MSA) [5], Molecular Similarity Matrices (for e.g., CoMSIA [6]), Distance Geometry [7], the Hypothetical Active Site Lattice method (HASL) [8], Genetically Evolved Receptor Models (GERM) [9], CoPASS [10], Catalyst [11], etc. have been developed around the CoMFA theme. Each method has its strengths and limitations which has now been realized.

One of these limitations is the wealth of knowledge available in ligand–receptor complexes is ignored. There is now a growing interest in developing receptor-dependent 3D QSAR models in which the receptor geometry is used in computing QSAR-independent variables. The first approach along this line was that of Hopfinger and colleagues [12] who modeled anticancer anthracyclines intercalating into DNA. More recently, COmparative BINDing Energy (COMBINE) [13] has been developed, in which the ligands are divided into fragments, which are chosen according to their spatial location in the protein–binding site, rather than by chemical identity. The intermolecular interaction energies between each fragment of the ligand and the amino acid residue are calculated. These energy terms are then correlated to experimental binding free energies (biological activities) by partial least squares (PLS) method. A reverse variant of CoMFA is Adaptation of fields for molecular comparison (AF-MoC) developed by Gohlke and Klebe [14] and is based on adaptive knowledge-based pair-potentials tailored for the protein while simultaneously considering ligand-based information in a CoMFA-type approach.

A conceptual flaw in most 3D-QSAR approaches is that contribution from the solvent and entropy factors, which many times influence and dictate the binding process are ignored. However, Vedani et al. have developed QSAR methods beyond the third dimension by accounting for the effect of different conformations as the fourth dimension [15], the induced fit mechanism as the fifth dimension [16] and evaluation of different solvation models as the sixth dimension [17], additionally incorporating entropy of ligand binding into the analysis. Much progress has been made in recent years in the theoretical foundations and practical computations of ligand–receptor binding thermodynamics [18]. Hence, an approach that includes such terms has the potential to enhance the quality and predictive power of a 3D-QSAR model.

We describe here an approach termed comparative residue–interaction analysis (CoRIA) which tries to take into account ligand–protein interactions as well the thermodynamics of binding to explain the variation in activity of 36 cyclooxygenase-2 (COX-2) inhibitors belonging to different structural classes. The COX-2 enzyme has been well studied at the molecular level and many structurally diverse inhibitors have been synthesized and tested with a good amount of success. Also, the X-ray structures of the COX-2 enzyme in complex with several inhibitors are available. These factors make the COX-2 dataset a good test case for the CoRIA hypothesis.

Ligand–receptor binding

The complete thermodynamics of ligand–receptor binding leading to a biological response, involves many events like interaction, solvation and entropy changes [19], all of which can profitably be taken into account in a 3D-QSAR analysis.

The binding free energy difference (ΔG_{bind}) between free and bound states of receptor and ligand ($\Delta G_{\text{complex}} - \Delta G_{\text{uncplexed}}$) is related to the binding constant (K_d) of ligand to the receptor and can be written as an additive interaction of different events using the classical binding free energy equation [15],

$$\Delta G_{\text{bind}} = \Delta G_{\text{solv}} + \Delta G_{\text{conf}} + \Delta G_{\text{inter}} + \Delta G_{\text{motion}} \quad (1)$$

i.e., the total free energy of binding (ΔG_{bind}) is an additive interaction of different events such as solvation of ligand (ΔG_{solv}) which is the difference between unbound (e.g. cellular) and bound state, conformational changes that occur in the receptor and/or ligand (ΔG_{conf}), specific interactions between the ligand and receptor as a consequence of their proximity (ΔG_{inter}), and the motion in the receptor and ligand once they are close to each other (ΔG_{motion}).

Every term in Eq. 1 contributes more or less to the ligand–receptor binding process. The binding free energy contribution by solvation at physiological conditions is the hydration free energy. Both the protein and the ligand are solvated before binding, the solvent molecules reorganize, as the interactions with water compete with protein–ligand interactions. Because of the relatively high surface area of the receptor compared to the ligand, the solvation binding free energy due to the receptor is negligible as compared to that of ligand. For the ligand, the electrostatic contribution to the hydration free energy can be calculated using the

finite difference approximation to the linear form of the Poisson–Boltzmann equation [20].

The next part of the binding free energy is the contribution from changes in the conformation of both the protein and the ligand due to their association (ΔG_{conf}). In many complexes, the conformation of the receptor on binding does not change much from the native structure, but in some (e.g., HIV protease [21]), the protein undergoes a large change in its conformation on inhibitor binding. The conformational change associated with the ligand upon binding to the receptor is more significant compared to that for the receptor and can be estimated by the strain energy upon binding. The ligand conformational energy can be calculated using a molecular mechanics potential function [22] as the energy associated with changes in bond lengths, angles, torsions and non-bonded interactions,

$$E = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}} + E_{\text{vdW}} + E_{\text{coul}} \quad (2)$$

The third term in Eq. 1, the interaction energy (ΔG_{inter}), is the dominant factor in the overall binding process in many cases, and is the main theme in the CoRIA formalism. Since it is purely an enthalpic contribution, ΔG can be replaced by ΔE , and is the total energy of the complex minus the energy of the free protein and free ligand. The electrostatic and van der Waals interactions are the major contributors to the interaction energy [23].

The last term in Eq. 1, ΔG_{motion} , incorporates the effects of molecular motion and accounts for the loss of torsional, vibrational, rotational and translational free energies upon binding. When two molecules bind, there is a loss of three rotational and three translational degrees of freedom. There is both an enthalpic and entropic contribution to this free-energy loss. The enthalpic contribution is about $3RT$ [24], and is usually assumed to be a constant value, independent of the size of the ligand. The entropic contribution is approximately 7–11 kcal/mol [19]. This contribution would cancel out when comparing different ligands binding to the same receptor.

We first describe the preparation of the test system, before we explain how each one of the above events in ligand–receptor binding has been accounted for in the CoRIA approach.

Materials and methods

The building of molecular models of COX-2 inhibitors and the enzyme-inhibitor complexes were carried out

with the *InsightII* (v 2000.3 L; Accelrys Inc., USA) [25] molecular modeling package. Energy minimization and molecular dynamics simulations were carried out with the *Discover* module (v 2000.3 L; Accelrys Inc., USA) and the CVFF force field [26]. The CVFF charges for atoms were used throughout this study. *DOCK* (v. 4.0) [27] and *Affinity* [28] (Accelrys Inc., USA) suite of programs were used to search the optimal binding poses of inhibitors. The statistical analysis was carried out using G/PLS as implemented in *Cerius2* (v 4.8; Accelrys Inc., USA) [29].

Biological data

The test system was a set of COX-2 inhibitors comprising pyrrole [30], imidazole [31, 32], cyclopentene [33, 34], benzene [35], pyrazole [36], spiroheptene [37], isoxazole [38], pyridine [39], thiazolone [40], thiadiazole and oxadiazole [41] derivatives. The inhibitory activities of all these molecules were measured on human recombinant enzyme by the same protocol, and are reported as IC_{50} values [30–41]. The enzyme inhibitory activities were converted to the negative logarithmic scale (pIC_{50}) and as seen in Table 1 spans 5 log orders.

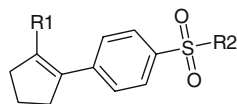
Docking of COX-2 inhibitors

The crystal structure of murine COX-2 [42], entry 1CX2 in PDB (www.rcsb.org/pdb) [43] complexed with SC-558 was used for the modeling studies. Hydrogens were added corresponding to pH 8.0 using the *InsightII Biopolymer* program, which resulted in a +1 charge for arginines and lysines, and –1 charge for the aspartates and glutamates. The histidines were unionized at this pH and were considered as such in this study. The enzyme along with the inhibitor was energy minimized using steepest descents and conjugate gradients, with the enzyme backbone atoms tethered with a force constant of 100 kcal/mol/Å². The electrostatic energy was computed with a distance dependent dielectric constant (r). During the minimization procedure, the ligand (SC-558) was free to move, and minimization was carried out until a gradient of 0.01 kcal/mol/Å was reached for the system.

For ligand docking, residues within a 15 Å radius from the centroid of SC-558 defined the active site. The Connolly surface for these residues was composed with the *MS* program [44] using a probe of radius 1.4 Å. Spheres were generated to occupy the binding pocket with the program *SPHGEN* in *DOCK*. The spheres were then enclosed in a grid and steric and electrostatic

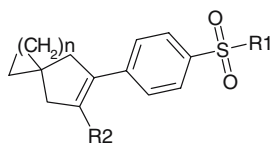
Table 1 Training and test set (T) molecules used in the present QSAR studies of COX-2 inhibitors

Cyclopentene



ID	pIC ₅₀	R1	R2
1	7.67		CH ₃
2	8.69		NH ₂
T1	6.82		NH ₂
T2	8.69		NH ₂

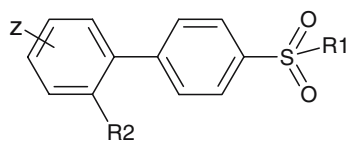
Spiroheptene



ID	PI _{C50}	R1	R2	<i>n</i>
3	8.39	CH ₃		2
4	6.86	CH ₃		1
5	8.00	CH ₃		1
6	8.09	CH ₃		1
T3	7.20	CH ₃		3
T4	9.0	NH ₂		1

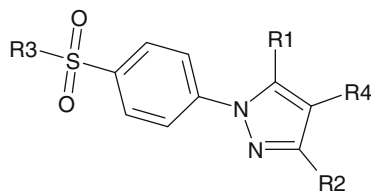
Table 1 continued

Benzenes



ID	pIC ₅₀		R1	R2
7	7.08		CH ₃	
8	6.46		CH ₃	
9	8.09		CH ₃	
T5	8.69		NH ₂	

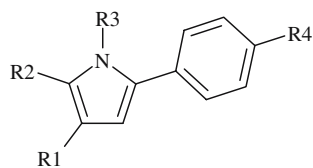
Pyrazoles



ID	pIC ₅₀	R1	R2	R3	R4
10	5.44		CF ₃	NH ₂	OH
11	7.00		CF ₃	CH ₃	H
12	7.38		CF ₃	NH ₂	H
13	5.48		CHF ₂	NH ₂	H
14	7.50		CF ₃	NH ₂	H
T6	6.46		CN	NH ₂	H

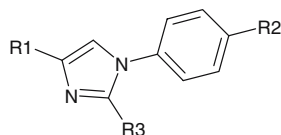
Table 1 continued

Pyrroles



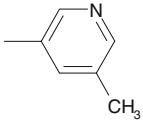
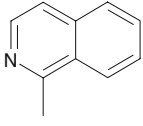
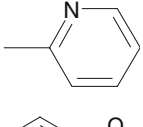
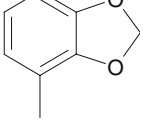
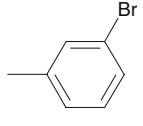
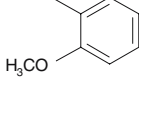
ID	pIC ₅₀	R1	R2	R3	R4
15	7.22	H	CH ₃		SO ₂ CH ₃
16	5.54	H	CH ₃		SO ₂ CH ₃
17	4.00		CH ₃		SO ₂ CH ₃
18	5.79	COCH ₃	CH ₃		SO ₂ CH ₃
T7	6.28	H	CH ₃		SO ₂ CH ₃
T8	6.29	H	H		F

Imidazoles

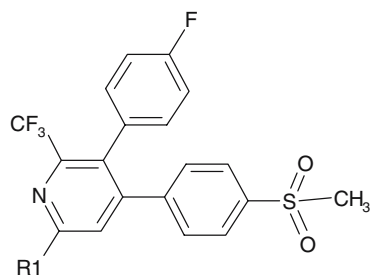


ID	pIC ₅₀	R1	R2	R3
19	5.24	CF ₃	SO ₂ CH ₃	
20	6.04	CF ₃	SO ₂ CH ₃	
21	7.09	CF ₃	SO ₂ CH ₃	
22	4.91	CF ₃	SO ₂ CH ₃	

Table 1 continued

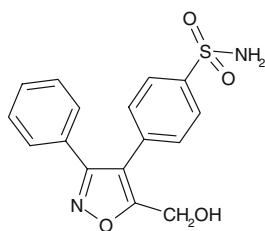
23	6.29	CF ₃	SO ₂ NH ₂	
24	5.76	CF ₃	SO ₂ CH ₃	
25	5.77	CF ₃	SO ₂ CH ₃	
26	6.32	CF ₃	SO ₂ CH ₃	
T9	8.15	CF ₃	SO ₂ NH ₂	
T10	4.00	CF ₃	SO ₂ CH ₃	

Pyridines



ID	pIC ₅₀	R1
27	6.22	Br
28	6.69	OCH ₂ C ≡ CH
29	6.52	OCH ₃
30	6.52	OCH ₂ CH ₃

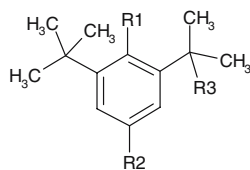
Isoxazoles



ID	PI _C ₅₀
31	6.74

Table 1 continued

Thiazolones, Thiadiazoles, Oxadiazoles



ID	pIC ₅₀	R1	R2	R3
32	6.85	OH		CH ₃
33	7.19	OH		CH ₃
34	5.76	OH		CH ₃
35	6.24	OH		H
36	5.49	OH		CH ₃
T11	5.18	OH		CH ₃
T12	7.32	OH		CH ₃
T13	5.82	OCH ₃		CH ₃

interactions of the enzyme were then calculated at each grid point by *DOCK*. Both ligand orientation and flexibility were considered during the docking process. The best docking mode for each ligand was selected on the basis of the ligand–enzyme interaction energy.

Refining docked conformations

The initial “*DOCK*” conformations were refined using *Affinity*. The protein structure was first classified into

three distinct regions: (i) residues within 5 Å radius from the centroid of the ligand were kept free; (ii) residues present in the region 5–10 Å from the centroid of the ligand were tethered with a force that allowed a limited mobility and (iii) the rest of the protein was fixed. The best scored structures from *DOCK*, were refined by *Affinity* through a combined Monte Carlo (MC) and simulated annealing (SA) procedures. In the MC procedure, all degrees of freedom for the ligand namely, translational, rotational and torsional,

were considered. In the MC step, the maximum permissible translational movement of the ligand was restricted to 1.0 Å along with 180° as the maximum angle of rotation. Configurations were screened with a quartic term for the Lennard Jones potential while neglecting the Coulombic energy. Selected configurations, based on their energies, were then minimized. The best structures were then screened with a standard Lennard Jones potential and a more accurate estimate of the electrostatic energies by the Cell multipole method [45]. The 10 lowest energy structures were then refined with SA, where a linear decrease in temperature from 500 K to 300 K in steps of 100 K for 100 fs each was adopted. The annealed structures were then minimized to a gradient of 0.001 kcal/mol/Å. The above protocol was followed for all the molecules in the data set. The final conformation selected for the QSAR study was the one with the lowest binding energy. The enzyme–inhibitor complexes thus obtained were used for the computation of the non-bonded interaction energies. The docking protocol was validated to reproduce the crystal structure of COX-2 with SC-558 (1CX2) and the same protocol was used to generate the ligand–receptor complexes for all molecules in the training and test sets. The other descriptors (*vide infra*) were calculated for the ligands extracted from this bound conformation.

Computation of individual terms in the CoRIA approach

Electrostatic contribution to the solvation energy (ΔG_{solv} in Eq. 1)

The electrostatic contribution to the solvation free energy of ligands as explained earlier, was calculated using the module *Delphi* (Accelrys Inc., USA). The program calculates the electrostatic energy of transfer of a molecule from vacuum to water, using the approximated Poisson-Boltzmann equation [46]. The ligand charges for *Delphi* were calculated using the CFF91 [47] force field, as these partial atomic charges have been reported to be more accurate in reproducing experimental solvation free energy [48].

Strain energy (ΔG_{conf} in Eq. 1)

The ligands were extracted from their respective enzyme complexes and minimized to relieve the “strain” due to constraints imposed by the protein environment. This strain energy accounts for the ΔG_{conf} term

in Eq. 1. The minimization routine involved several cycles of steepest descents, conjugate gradients and Newton Raphson (BFGS) methods, and terminated when the gradient was below 0.001 kcal/mol/Å. The strain energy was computed as a difference between the energy of the ligand in complex with the protein and the energy of the ligand after free minimization *in vacuo*. The protein strain energy was found to be insignificant in the present case.

Non-bonded interaction energies (ΔG_{inter} in Eq.1)

The van der Waals and Coulombic interaction energies for the enzyme–ligand complexes were calculated using the CVFF forcefield in *Discover3* program. The functional forms of the two non-bonded interaction energies are expressed as,

$$E_{\text{vdw}} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad (3)$$

$$E_{\text{ele}} = \frac{q_i q_j}{\epsilon r_{ij}} \quad (4)$$

for the van der Waals and Coulombic energies, respectively.

A total of 84 residues in the enzyme enclosed by a 10 Å radius from the inhibitor were identified as the binding region and considered for calculation of nonbonded energies (Eqs. 3 and 4). For each ligand there are thus 84 entries each for the van der Waals, Coulombic and the sum of the two non-bonded interaction energies, respectively, making a total of 252 non-bonded interaction energies in the QSAR study table.

Miscellaneous descriptors

Besides the above three major components in the ligand binding process several other descriptors for the ligand have been calculated which directly or indirectly can be related to one or the other terms in the classical equation. These are free energy for desolvation of ligand in water (*Fh2o*) and octanol (*Foct*) [49]. These linear free energy properties have proven useful as molecular descriptors in structure activity analysis [50]. The computations of *Fh2o* and *Foct* are based solely on the connectivity of the atoms in the molecule and are not dependent critically on the conformation of the molecule. The *Fh2o* and *Foct* were calculated using the *Cerius2* program.

Jurs descriptors [51]

Jurs descriptors combine shape and electronic information that encode features responsible for polar interactions between molecules. The molecular surface is defined by the van der Waals radii of the atoms, as traced by a sphere representing a solvent molecule (water by default). The surface traced out by the center of the solvent sphere is termed the solvent-accessible surface. The molecule is further defined by a specific electron distribution, thus yielding a representation of a charged contact surface where polar intermolecular interactions can take place. It is calculated by mapping the atomic partial charges on the solvent-accessible surface area of individual atoms. The sum of the solvent-accessible surface areas of all positively charged atoms (PPSA-1) and the sum of solvent-accessible surface areas of all negatively charged atoms (PNSA-1) have been calculated using following equations as implemented in *Cerius2*.

$$\text{PPSA} - 1 = \Sigma(+\text{SA}_i) \quad (5)$$

$$\text{PNSA} - 1 = \Sigma(-\text{SA}_i) \quad (6)$$

where, $+\text{SA}_i$ and $-\text{SA}_i$ are the surface area contributions of the i th positively or negatively charged atom in the molecule respectively.

Molar refractivity

The molar refractivity index of a molecule is a combined measure of its size and polarizability. This fragment constant thermodynamic descriptor relates the effect of substituents on a reaction center from one type of process to another. The basic idea behind using such descriptor is that similar changes in structure are likely to produce similar changes in reactivity, ionization, or binding. It was calculated using the method described by Ghose and Crippen [52] as implemented in the *Cerius2* program.

Molecular volume

Molecular volume is a 3D spatial descriptor that defines the molecular volume inside the contact surface. This property is a function of conformation and is related to binding and transport. It was calculated using the *Cerius2* program.

Lipophilicity

The octanol/water partition coefficient ($\log P$) is related to the hydrophobic character of the molecule. It was also calculated using Ghose and Crippen's [52] atom based approach as implemented in *Cerius2*, in which each atom of the molecule is assigned to a particular class. The total value of $A \log P$ is an additive contribution of each atom, which is calculated by the following equation.

$$A \log P = \sum n_i a_j \quad (7)$$

where, n_i is number of atoms of type i and a_j is contribution of an atom of type i .

Surface area

This 3D spatial descriptor was computed using the *Cerius2* program. It is a measure of the van der Waals area of a molecule and determines the extent to which the molecule exposes itself to the external environment. The molecular surface area is related to binding, transport, and solubility of the molecule.

Statistical analysis

Genetic function approximation (GFA) [53] in conjunction with PLS [54] referred to as G/PLS was the statistical tool used to develop meaningful QSAR models. The GFA algorithm develops an initial population of individuals; a fitness function which is a measure of least square error—termed 'lack of fit' (LOF) [53], is then applied as an estimate of the quality of each individual. Individuals with the best fitness scores are allowed to mate and propagate their genetic material to offspring through the crossover and/or mutation operation. After repeatedly performing these steps, the average fitness of the individuals in the population increases, as good combination of 'genes' (descriptors in the present case) are discovered and spread through the population. This can be observed as the frequency at which each term (descriptor) appears in all equations in the final population. The entire population of equations can then be searched for information on features, patterns and regions in which different equations predict well.

Both linear (x) terms and linear in combination with spline $\langle x - a \rangle$ terms were used to develop the QSAR equations. A total of 8 models, four with only linear terms and four with a combination of linear and

spline terms, were developed using the above described parameters. To preserve the information contained in the descriptor dataset, pretreatment based on correlation matrix was avoided [55]. The ‘unit scaling’ method was used for all the descriptors, where all the values in a given column are divided by its standard deviation to obtain a mean of 0 and a variance of 1. The unit scaling places all the columns, to be considered for statistical analysis, on the same platform and all the columns have equal weight (importance) while deriving the equations. The optimal number of components was selected as six, for which the cross-validated r^2 (q^2) was found to be the highest; the length of the equations was set to six terms with a smoothness value of 1.0 (the smoothness function controls the bias in the scoring factor between equations with different number of terms), and number of generations were limited to 10,000. Crossover and mutation probabilities of 50% (default settings) were used.

Validation

Cross validation using both the leave-one-out (LOO) and leave-five-out procedures with 50 trials was followed to calculate the q^2 for every model. The activity of the test set molecules was predicted using the CoRIA models and predictive r^2 (r_{pred}^2) was calculated. The features and validation results of the CoRIA models are reported in Table 2.

Results and discussion

Physico-chemical properties derived from ligand–protein interaction studies can be easily incorporated into the QSAR formalism. Two such attempts to include physico-chemical properties from the ligand–protein complex to formulate 3D-QSAR equations have appeared [11, 56]. Overall, all 3D-QSARs reported till date have only probed and extracted information about a specific interaction involving the ligand, which is mostly restricted to its binding site while neglecting critical elements of the ligand binding process. The CoRIA formalism was developed to fully account for the thermodynamics of ligand–protein binding within the QSAR framework. In order to perform a CoRIA analysis, one needs (i) a 3D structure of the protein (X-ray, or NMR or Homology model) preferably with a bound ligand; (ii) a docking algorithm; (iii) a method to compute intermolecular interactions (van der Waals and Coulombic energies), (iv) program(s) to calculate binding descriptors that the medicinal chemist thinks are important for describing the events during the

ligand–protein binding process, and (v) a robust statistical method, such as G/PLS or a neural network, for deriving the QSAR models.

To prove the capability of the CoRIA approach, we have selected the COX-2 enzyme as a test case since there is a vast group of inhibitors with diverse chemical structures and a wide spectrum of biological activity. Forty-nine molecules with COX-2 inhibitory activity, that included cyclopentenones, spiroheptenes, benzenes, pyrazoles, pyrroles, imidazoles, pyridines, isoxazoles, thiazoles, thiadiazoles, and oxadiazoles, were docked into the enzyme active site to determine their binding conformation as there is no experimental evidence of this property. Since, the non-bonded interaction energies are sensitive to small changes in the conformation, the docked conformations were further refined using Monte Carlo and SA procedures. Finally, the ligand–receptor complexes were carefully examined for any bad clashes between ligand and receptor atoms. The overall goal is to secure ligand–receptor complexes as accurate as possible since the whole idea of calculation of various QSAR descriptors is critically dependent on the “goodness” of these complexes. This is the reason why great care was taken at each stage of calculation. There are no reports in the literature about water molecules being crucial in the binding site of COX-2, hence they were not considered in the calculations.

To account for the major determinants of the binding process, we have computed descriptors that appropriately define these events. These include the non-bonded interaction energy terms (van der Waals and Coulombic, ΔG_{inter}), strain energy of the ligand in the bound conformation (ΔG_{conf} for ligand), electrostatic contribution to the solvation energy (ΔG_{solv} of ligand, desolvation free energy of ligand in water (*Fh2o*) and octanol (*Foct*), Jur’s partial positive surface area (PPSA-1) and partial negative surface area (PNSA-1), surface area of the ligands (related to ΔG_{solv} for ligand), lipophilicity ($A \log P$, also related to ΔG_{solv}), molecular volume and molecular refractivity. An inclusion of these binding descriptors in a QSAR model can rationalize the full process of binding from the transfer of free ligand in solution to its complexation with receptor.

From the dataset of 49 molecules, 36 were grouped into a training set to construct the CoRIA models, based on chemical (Daylight fingerprints) and biological (pIC_{50}) diversity. This was achieved with the Tanimoto coefficient using the “select diverse” utility in *Cerius2*. The remaining 13 molecules were reserved for testing the predictivity of the models. Four different models were developed using various combinations of descriptors. In Model 1, the Coulombic (C) and van

Table 2 CoRIA equations for various combinations of the descriptors

Model	CoRIA equations	r^2	q^2 by LOO	q^2 by LGO	r^2_{pred}	P^2
1L	$\text{pIC}_{50} = 11.89 + 0.74 \text{ "V_Tyr385"} - 7.25 \text{ "C_Pro528"} + 51.09 \text{ "V_Tyr115"} + 60.00 \text{ "V_Arg433"} - 11.71 \text{ "C_Met113"}$	0.94	0.94	0.83	0.57	0.66
2L	$\text{pIC}_{50} = 11.85 + 81.16 \text{ "V_Tyr115"} + 0.75 \text{ "V_Tyr385"} - 7.40 \text{ "C_Pro528"} + 58.52 \text{ "V_Arg433"} - 11.37 \text{ "Met113"}$	0.94	0.93	0.85	0.60	0.70
3L	$\text{pIC}_{50} = 11.89 + 60.00 \text{ "V_Arg433"} - 11.71 \text{ "C_Met113"} + 0.74 \text{ "V_Tyr385"} - 7.25 \text{ "C_Pro528"} + 51.09 \text{ "V_Tyr115"}$	0.94	0.94	0.72	0.57	0.70
4L	$\text{pIC}_{50} = 11.85 + 81.16 \text{ "V_Tyr115"} + 0.75 \text{ "V_Tyr385"} - 7.40 \text{ "C_Pro528"} + 58.52 \text{ "V_Arg433"} - 11.37 \text{ "Met113"}$	0.94	0.94	0.78	0.60	0.72
1LS	$\text{pIC}_{50} = 9.35 - 60.76 < -0.04 - \text{ "V_Arg433"} > - 7.36 \text{ "C_Pro528"} + 2.52 \text{ "V_Val116"} + 0.74 \text{ "V_Tyr385"} - 11.48 \text{ "C_Met113"}$	0.94	0.90	0.77	0.41	0.76
2LS	$\text{pIC}_{50} = 9.12 + 0.23 \text{ "Arg120"} + 0.78 \text{ "V_Tyr385"} - 270.28 < -0.01 - \text{ "V_Val89"} > - 7.13 \text{ "Pro528"} + 10.47 \text{ "V_Gln350"}$	0.93	0.89	0.70	0.24	0.69
3LS	$\text{pIC}_{50} = 7.91 + 19.69 < 0.019 - \text{ "C_Phe470"} > + 0.78 \text{ "V_Tyr385"} + 7.62 < 0.214 - \text{ "C_Pro528"} > - 180.55 < -0.01 - \text{ "V_Val89"} > - 54.41 < -0.04 - \text{ "V_Arg433"} >$	0.92	0.93	0.84	0.43	0.76
4LS	$\text{pIC}_{50} = 9.74 + 61.36 \text{ "V_Arg433"} + 0.74 \text{ "V_Tyr385"} + 1.85 \text{ "V_Val116"} - 10.54 \text{ "C_Met113"} + 7.38 < 0.231 - \text{ "C_Pro528"} >$	0.93	0.89	0.77	0.54	0.69

Number of molecules in the training set is 36; r^2 : correlation coefficient; q^2 by LOO and LGO: cross-validation correlation coefficient by LOO and Leave-Group-Out (group of 5) respectively; r^2_{pred} : predictive correlation coefficient; P^2 : predictive correlation coefficient calculated as per Vedani et al. [15–17]; V_Tyr385: van der Waals interaction with residue Tyr385; C_Pro528: Coulombic interaction with Pro528; Met113: van der Waals plus Coulombic interaction with Met113

der Waals (V) interaction energies between the ligand and the residues in the receptor active site were considered for construction of the CoRIA equations (i.e., C + V). Model 2 included the total non bonded interaction energy (NB), in addition to C and V terms (i.e., C + V + NB). To account for the different processes leading to binding, Model 3 was developed using the ten descriptors (D) i.e. strain energy (ΔG_{conf}), Jurs PPSA-1, Jurs PNSA-1, molecular volume, surface area, $A \log P$, $Fh2o$, $Foct$, molar refractivity and electrostatic contribution to the solvation energy (ΔG_{solv}) in addition to C & V (i.e., C + V + D), and finally Model 4 was developed with all descriptors (C + V + NB + D) described above. Each of the models described above were further categorized into those that included only linear terms (Models 1L to 4L) and those that considered both linear and spline terms in the construction of the models (Models 1LS to 4LS).

Models 1L to 4L have a correlation coefficient (r^2) greater than of 0.90 as seen in Table 2. The cross validated correlation coefficient (q^2), using LOO and leave-five-out, is greater than 0.7 for all the models, indicating a good predictive power of the models for the molecules in the training set. The q^2 (LOO) values are greater than 0.9, however q^2 values derived by leave-group-out (group of five) are smaller and around 0.6. The q^2 (LOO) reported using GFA are high, and there are several literature reports with such high q^2 values (0.85–0.91) [57–61]. The predictive r^2 (r^2_{pred}) of Models 1L to 4L range from 0.54 to 0.60, which indicates a good prediction of the activity of molecules outside the training set. The p^2 of all models is above 0.6 indicating good internal predictive power of the test set. An examination of all the eight CoRIA models (Table 2) reveals that Model 1L contains most of the terms present in the other models but has the best statistics. This model was therefore used to predict the activity of the molecules in the test set. A plot of the predicted versus experimental activity of the training set using Model 1L is shown in Fig. 1. The 10 best equations associated with Model 1L were analyzed for the frequency with which a descriptor occurs in the population of equations (Fig. 2). An examination of the population of equations shows that the terms describing van der Waals interactions with residues Tyr385 and Arg433 and Coulombic interaction with Pro528 dominate in Models 1L to 4L. Coulombic interaction with Met113 and van der Waals interaction with Tyr115 are seen variably in some of the ‘best’ equations.

In order to gauge the influence of spline factors on the structure or behavior of the models, four other

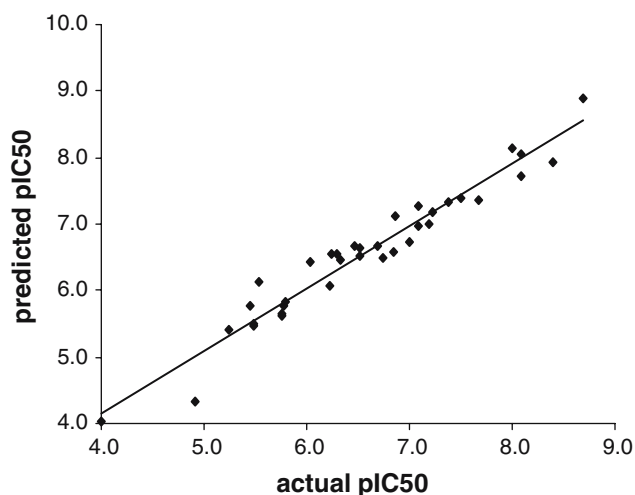
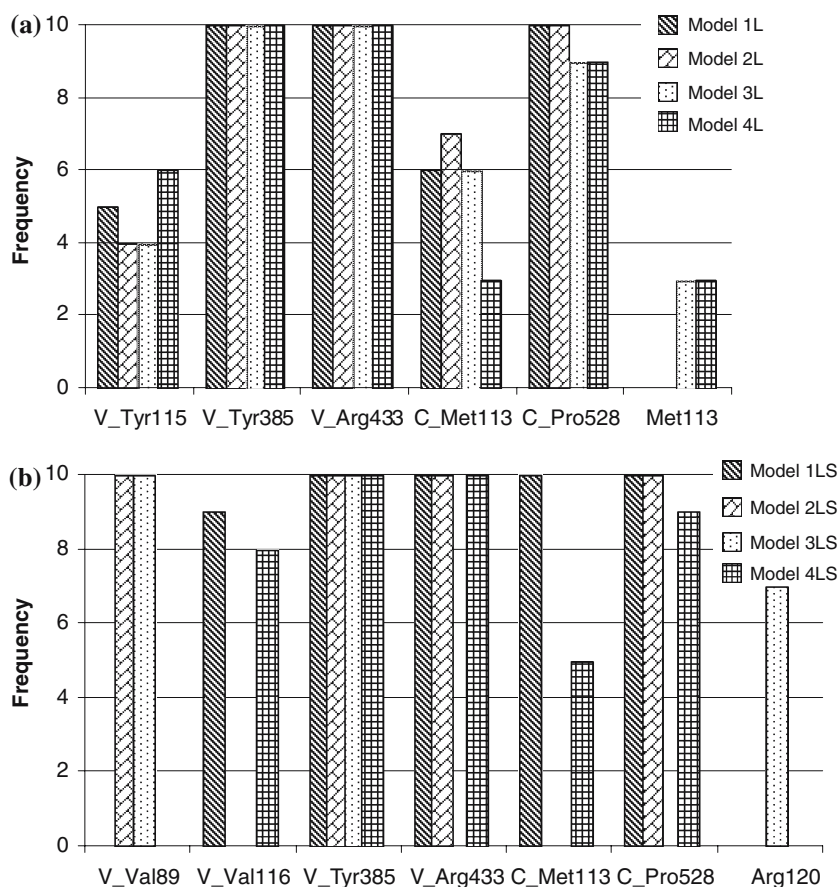


Fig. 1 The plot of predicted versus experimental pIC_{50} values of the training set molecules using CoRIA Model 1 L

models (Models 1LS to 4LS) were developed using a combination of linear and spline $\langle x-a \rangle$ terms. In all these models (Model 1LS to 4LS), terms describing van der Waals interaction (V) with Tyr385 and Coulombic interaction (C) with Pro528 dominate as previously (in Models 1L to 4L); however, Coulombic interaction with Met113 is absent in Models 2LS and 3LS. The van der Waals interaction with Arg433 also appears as a spline term in Models 1LS and 3LS, as a linear term in Model 4LS, but is not found in Model 2LS. Except for Model 2LS, the other models, namely Models 1LS, 3LS and 4LS, have a good predictive power.

An analysis of all CoRIA models reveals that ligand interaction with Tyr385 plays a dominant role in the strength of its COX-2 inhibition. This is in congruence with the reported point mutation data [57–59]. The residue Tyr385 is present in the catalytic center at the apex of the channel near the end of helix H8. Variation in the strength of the interaction between Tyr385 and the 5-aryl ring of pyrazole, for example, may be one of the factors that modulate the activity of this class of molecules. The other residues revealed by CoRIA to be important in modulating the activity are Pro528, Arg433, Met113 and Tyr115. There is no point mutation data yet reported for these residues, showing their importance in the COX-2 activity, but the CoRIA analysis highlights the interactions of inhibitors with these residues as being crucial for molecular recognition. A close appraisal of the ligand–receptor complexes in light of the CoRIA equations reveals residues predicted to be important in controlling the activity lie within a 5 Å radius of the inhibitors. This is noteworthy since non-bonded interactions are significant within a 5 Å radius of the inhibitor and decay gradually at

Fig. 2 Frequency of descriptors appearing in the 10 best equations of CoRIA Models 1L to 4L (a) and Models 1LS to 4LS (b) Coulombic interactions with Pro528 and Met113 and van der Waals interactions with Tyr115, Tyr385 and Arg433 appear more frequently in the models

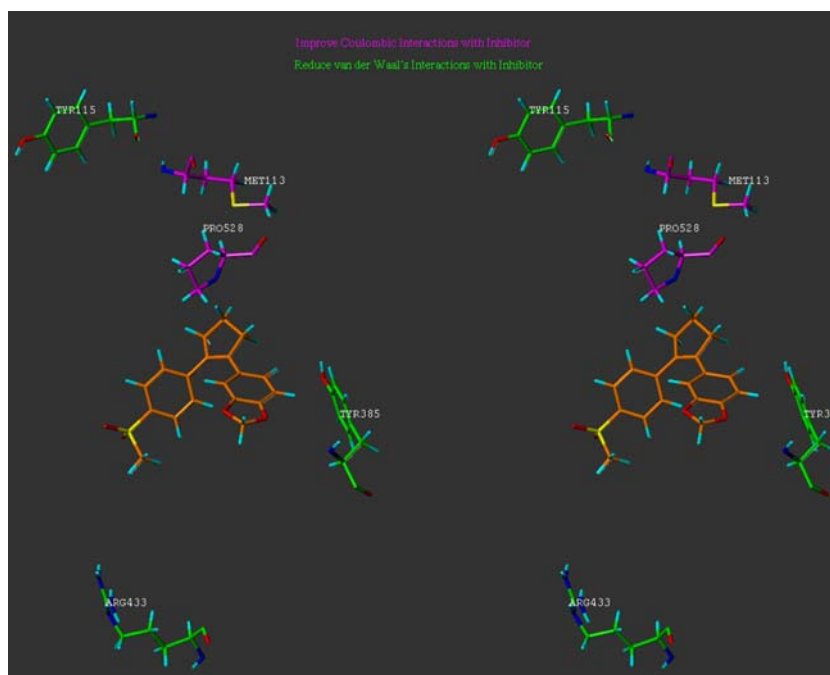


longer distances. There are four residues—Arg120, Tyr355, Arg513 and Glu524—which are involved in a hydrogen bonding network and form a gate for the entry [60] of inhibitors and play an important role in the time-dependent irreversible inhibition of COX-2. The molecules considered in this study are all reversible inhibitors; as a result none of these active site residues appear in the CoRIA equations with the exception of Arg120. Only those inhibitors having a carboxylate group like the aryl propionates are seen to interact with Arg120 [61].

The CoRIA models can easily be utilized in the design of new ligands. Figure 3 shows molecule 6 surrounded by the active site residues reflected in the CoRIA equations. It can be inferred from Model 1L (Table 2) that the Coulombic interactions with residues Pro528 and Met113 need to be maximized and van der Waals interaction with residues Try115, Tyr385 and Arg433 needs to be reduced, to improve the activity. This is evident from the interaction energy data presented in Table 3. In this table, pairs of molecules with different activities, but with a seemingly minor modification have been compared and the variation in their activity explained based on the CoRIA models.

Modification of $-\text{SO}_2\text{CH}_3$ (molecule 1) to $-\text{SO}_2\text{NH}_2$ (molecule 2) causes an increase in activity by one log unit. The variation in the activity as predicted by Model 1L, is a result of the increase in the Coulombic interaction with Pro528 (from -0.100 for molecule 1 to -0.138 for molecule 2) and a concomitant decrease in the van der Waals interaction with Tyr385 (from -2.856 for molecule 1 to -1.125 for molecule 2). In molecules 4 and 5, replacement of the 4-trifluoromethoxyphenyl (in molecule 4) by 3-fluoro-4-methoxyphenyl (molecule 5) causes a one log unit increase in the activity. The factor behind this increase in the activity is a decrease in the van der Waals interaction with Tyr385, which is in close proximity to this group. A similar analyses carried out for the molecule pairs 8 & 9; 10 & 12; 16 & 15 leads to the conclusion that changes in the interaction energies principally with the residues Pro528 (increase in Coulombic interaction) and Tyr385 (decrease in van der Waals interaction) are sufficient to explain the variation in the activity. These two residues are closer to the inhibitor than the remaining three residues, as seen in Fig. 2, and hence have a decisive role in modulating the activity. Although the coefficient (+0.74) for the van der Waals interaction with Tyr385

Fig. 3 A stereoview of the COX-2 active site in which molecule **1** is shown with the important active site residues appearing in the CoRIA equations. The color codes used to represent atoms are: N: blue, O: red, S: yellow, H: cyan, C (ligand): orange, C (Tyr115, Tyr385, Arg433): green, and C (Met113, Pro528): magenta



is relatively small, the decrease in the van der Waals interaction energy upon modification is significant to cause changes in the activity (Table 3). The coefficient for the Coulombic interaction with Pro528 is -7.25 , i.e. a more negative value for this energy term will have a greater effect on the activity and by optimizing the Coulombic interaction with this residue, it is possible to design molecules with enhanced activity.

None of the thermodynamic descriptors like solvation or entropy appear in the final set of CoRIA models. Some of these descriptors appear transiently during the evolution but gradually disappear as the genetic function progresses. Thus it turns out that in the case of COX-2 inhibition, the interaction energies with residues in the active site are sufficient to explain

the variations in the biological activity. It is important to mention that while deriving the CoRIA models, no bias was placed on any of the descriptors and all had an equal weightage during development of the equations. However, these thermodynamic descriptors do appear in the equations when the number of terms in the equation is increased to 12 or more. The equation with 12 terms for a training set of 36 molecules breaks the principles of QSAR according to which the number of terms in a QSAR equation should not be greater than the number of molecules in the training set divided by five, and disregarding this principle leads to an overfitting of the equation. Also, smaller the number of terms in the QSAR equation, the easier it is to use in optimizing molecules. Further, when the analysis is run

Table 3 The interaction energies of the training set molecules (discussed in text) with the residues reflected in CoRIA models 1L to 4L

Molecule	pIC ₅₀	C_Met113	C_Pro528	V_Tyr115	V_Tyr385	V_Arg433
1	7.67	0.008	-0.100	-0.009	-2.856	-0.043
2	8.69	0.004	-0.138	-0.010	-1.255	-0.042
4	6.86	0.006	-0.102	-0.014	-2.654	-0.046
5	8.00	0.010	-0.135	-0.014	-1.702	-0.044
8	6.46	0.001	0.033	-0.014	-2.339	-0.042
9	8.09	-0.007	0.052	-0.014	-0.361	-0.048
10	5.44	0.040	0.115	-0.014	-1.829	-0.046
12	7.38	0.027	-0.062	-0.013	-1.787	-0.045
16	5.54	-0.009	0.024	-0.012	-2.903	-0.049
15	7.22	0.004	0.113	-0.012	-1.114	-0.040
SC-558	8.09	0.017	-0.157	-0.038	-0.827	-0.035
Celecoxib	7.40	0.030	-0.107	-0.035	-1.208	-0.037

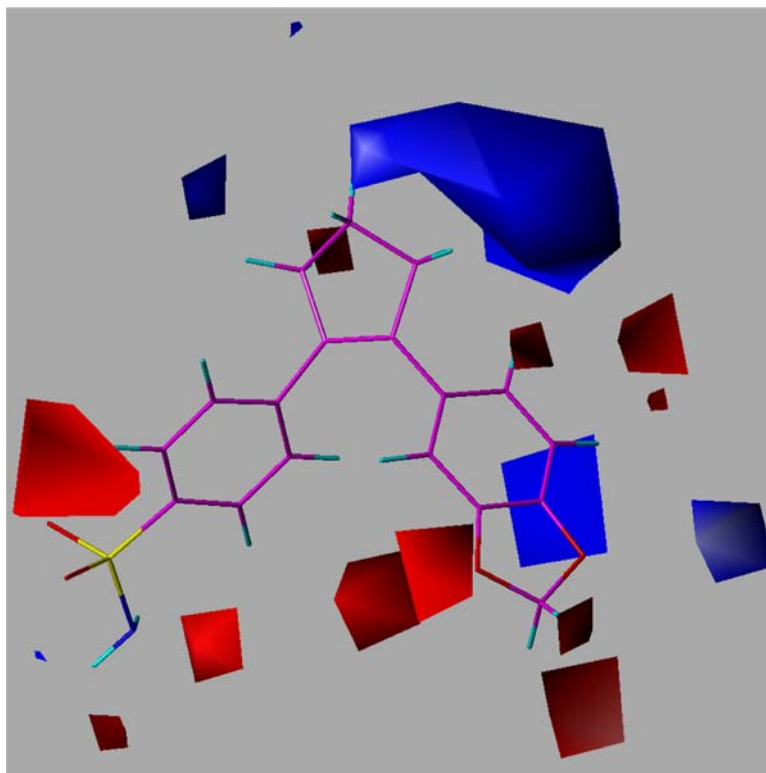
using only the molecular descriptors (D), the r^2 obtained is ~ 0.5 and the above mentioned terms are associated with very small coefficients.

There are several QSAR models reported for COX-2 inhibitors in the literature [62–66]. The CoMFA model developed by Desiraju et al. [62], where a limited set of 1,2-diarylimidazole derivatives were minimized with constraints and aligned to the X-ray conformation of SC-558, has a q^2 of 0.56 and a r^2_{pred} of 0.79 for the test set. Similarly, starting from the crystal structure of SC-558 as the template, Chavatte et al. [63] have reported CoMFA models for various classes of inhibitors with q^2 of about 0.70 and r^2_{pred} ranging from 0.65 to 0.74. Liu et al. [64], have developed a CoMFA model using docking strategies for alignment of 1,5-diarylpyrazoles, with a q^2 of 0.63 and a r^2_{pred} of 0.80. In comparison to these reported QSAR models, the CoRIA models have comparable q^2 and r^2_{pred} values. The r^2_{pred} is for large diverse test set in our study, which is limited to a structural class in each of three studies mentioned above. We had previously reported a CoMFA study of COX-2 inhibitors, using two strategies for generating conformations and alignment of the molecules; one based on ‘template-forcing’ on to SC-558, and other based on ‘flexible docking’ [65]. The CoMFA model based on ‘template forcing’ has a q^2 of 0.66 and an r^2_{pred} of 0.63, while the CoMFA model

developed with ‘flexible docking’ has a q^2 of 0.73 and an r^2_{pred} of 0.76. The CoMFA models of our previous study reveal the presence of sterically unfavorable contours around most of the molecular space except at the *para* position of the *p*-bromophenyl ring of SC-558. This result parallels the observation in the present CoRIA models, which suggests that it is important to decrease the van der Waals interactions particularly with residues Tyr385, Tyr115 and Arg433 to improve the activity. The training set used in the previous study [65] and the one in the present CoRIA study are not exactly the same. For comparison purposes, we have developed a CoMFA model using the present training set of 36 molecules. This CoMFA model exhibits an r^2 of 0.99, a q^2 of 0.4 and an r^2_{pred} of 0.5. The CoMFA contours of this model have been shown in Fig. 4. The statistics of the CoRIA model is superior to the corresponding CoMFA model.

Kim et al. [66] have reported CoMFA, CoMSIA and COMBINE models for a dataset comprising three structural classes of COX-2 inhibitors with q^2 of 0.84, 0.79 and 0.64 respectively. We have included nine structurally different chemical classes of selective COX-2 inhibitors in our study. The COMBINE equations contain 14 terms (for the triaryl ring dataset of 58 molecules), 8 terms (for the diaryl cycloalkanopyrazole dataset of 18 molecules), and 5 terms (diphenyl

Fig. 4 The electrostatic contours of the CoMFA model generated for the same training set used for the CoRIA study, are shown around the highest active molecule **2**. The red and blue contours show how negative and positive electrostatic fields can be used to improve the activity of the molecules



hydrazide dataset of 12 molecules). In the above mentioned study, the number of terms far exceeds the minimum number of molecules in the training set used to derive the model and as explained earlier this has led to an over-fitting in the COMBINE equations. In the CoRIA models, the number of terms in every equation is within the statistical requirements. The active site residues appearing in our CoRIA equations are also present in the COMBINE equations. The output of the CoRIA equations is more meaningful. However, the CoMFA and CoMSIA results discussed in Kim's work are more meaningful than the result of the COMBINE methodology, but not above the CoRIA approach.

Validation of CoRIA Approach

The various CoRIA models were used as a guideline to suitably modify the inhibitors so as to improve their activity. We present here one example. After correcting for any errors in the crystal structure and then subjecting it to a thorough minimization procedure, the activity of SC-558 was predicted based on CoRIA Model 1 L and was calculated to be 8.18 (the experimental activity is 8.09). The SC-558 structure was then modified by replacing –Br with –CH₃ to yield celecoxib. The new complex was minimized as discussed previously. The activity of celecoxib was predicted using Model 1L and found to be 7.41. The reported pIC₅₀ of celecoxib is 7.40. The interaction energies for both SC-558 and celecoxib with various residues predicted by Model 1L to be important in modulating the activity are listed in Table 3. It is clear from this table that once again, van der Waals interaction with Tyr385 and Coulombic interaction with Pro528 are crucial for COX-2 inhibition, as ascribed earlier.

Conclusions

A novel approach termed CoRIA that incorporates the thermodynamics of binding into the QSAR formalism was developed in order to capture those events that are important in the ligand–receptor recognition process. The approach was tested on a diverse set of COX-2 inhibitors. The CoRIA models have a good correlation coefficient (r^2) and predictive power (q^2) compared to previously reported [62–66] 3D-QSAR models for COX-2 inhibitors.

The CoRIA methodology has been successful in extracting key residues in the COX-2 enzyme that have been reported by X-ray crystallography and

site-directed mutagenesis studies to be important in ligand binding. The various equations derived by the CoRIA approach have also brought into focus some other residues that have hitherto not yet been reported and may have a hidden role in creating strong forces that drive the ligand towards specific binding. It is also pertinent to emphasize here that the approach of combining docking studies with the QSAR paradigm allows a more meaningful correlation between ligand–receptor binding and biological activity.

The CoRIA approach can advantageously be used to extract the residues that are important for binding and also as a guide for carrying out mutation studies directed towards understanding ligand binding. The results can be easily interpreted by the medicinal chemist and can provide clear directions for modifications of lead molecules during the lead optimization phase. Thus, the CoRIA methodology, by the inclusion of binding thermodynamics in the QSAR formalism, is a powerful approach to aid the drug design process.

Acknowledgements This work was made possible by a grant [01(1986)/05/EMR-II] from the Council of Scientific and Industrial Research (CSIR), New Delhi to E. C. Coutinho. S. A. Khedkar and A. K. Malde thank CSIR for financial support. We thank Dr. Krishna Iyer for useful suggestions during the course of this work.

References

1. Hansch C, Maloney PP, Fujita T, Muir RM (1962) *Nature* 194:178
2. Hansch C, Muir RM, Fujita T, Maloney PP, Geiger CF, Streich M (1963) *J Am Chem Soc* 85:2817
3. Hansch C, Fujita T (1964) *J Am Chem Soc* 86:5175
4. Cramer RD III, Patterson DE, Bunce JD (1988) *J Am Chem Soc* 110:5959
5. Tokarski JS, Hopfinger AJ (1994) *J Med Chem* 37:3639
6. Klebe G, Abraham U, Mietzner T (1994) *J Med Chem* 37:4130
7. Havel TF, Kuntz ID, Crippen GM (1983) *Bull Math Biol* 45:665
8. Doweiko AM (1988) *J Med Chem* 31:1396
9. Walters DE, Hinds RM (1994) *J Med Chem* 37:2527
10. Jain AN, Koile K, Chapman D (1994) *J Med Chem* 37:2315
11. Müller K (ed) (1995) *Perspectives in drug discovery and design*, vol 3. ESCOM Science Publishers, BV, Leiden, pp 1–20
12. Nelson LM, Malhotra D, Hopfinger AJ (1981) *Comput Chem* 5:19
13. Ortiz AR, Pisabarro MT, Gago F, Wade RC (1995) *J Med Chem* 38:2681
14. Gohlke H, Klebe G (2002) *J Med Chem* 45:4153
15. Vedani A, Briem H, Dobler M, Dollinger K, McMasters DR (2000) *J Med Chem* 43:4416
16. Vedani A, Dobler M (2002) *J Med Chem* 45:2139
17. Vedani A, Dobler M, Lill L (2005) *J Med Chem* 48:3700
18. Tokarski JS, Hopfinger AJ (1997) *J Chem Info Comput Sci* 37:792

19. Charifson PS (ed) (1997) Practical application of computer-aided drug design, Marcel Dekker Inc., pp 355–410
20. *Delphi 2.5*, User Guide. Accelrys Inc., San Diego, CA, USA
21. Scott WR, Schiffer CA (2000) Structure Fold Des 8:1259
22. Charifson PS (ed) (1997) Practical application of computer-aided drug design. Marcel Dekker Inc., pp 495–538
23. Knetgel RMA, Grootenhuis PDJ (1998) Perspect Drug Discov Des 9–11:99
24. Williams DH, Cox JPL, Doig AJ, Garner M, Gerhard U, Kaye PT, Lal AR, Nicholls IA, Salter J, Mitchell RC (1991) J Am Chem Soc 113:7020
25. InsightII, version 2000.3 L. Accelrys Inc., San Diego, CA, USA
26. Dauber-Osguthorpe P, Roberts VA, Osguthorpe DJ, Wolff J, Genest M, Hagler AT (1988) Proteins 4:31
27. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982) J Mol Biol 161:269
28. Affinity v.98 Molecular Modeling Program Package. Accelrys Inc. San Diego, CA, 1998
29. Cerius2 version 4.8. Accelrys Inc., San Diego CA, 1999
30. Khanna IK, Weier RM, Yu Y, Collins PW, Miyashiro JM, Koboldt CM, Veenhuizen AW, Currie JL, Seibert K, Isakson PC (1997) J Med Chem 40:1619
31. Khanna IK, Weier RM, Yu Y, Xu XD, Koszyk FJ, Collins PW, Koboldt CM, Veenhuizen AW, Perkins WE, Casler JJ, Masferrer JL, Zhang YY, Gregory SA, Seibert K, Isakson PC (1997) J Med Chem 40:1634
32. Khanna IK, Yu Y, Huff RM, Weier RM, Xu X, Koszyk FJ, Collins PW, Cogburn JN, Isakson PC, Koboldt CM, Masferrer JL, Perkins WE, Seibert K, Veenhuizen AW, Yuan J, Yang DC, Zhang YY (2000) J Med Chem 43:3168
33. Reitz DB, Li JJ, Norton MB, Reinhart EJ, Collins JT, Anderson GD, Gregory SA, Koboldt CM, Perkins WE, Seibert K, Isakson PC (1994) J Med Chem 38:3878
34. Li JJ, Anderson GD, Burton EG, Cogburn JN, Collins JT, Garland DJ, Gregory SA, Huang HC, Isakson PC, Koboldt CM, Logusch EW, Norton MB, Perkins WE, Reinhart EJ, Seibert K, Veenhuizen AW, Zhang Y, Reitz DB (1995) J Med Chem 38:4570
35. Li JJ, Norton MB, Reinhart EJ, Anderson GD, Gregory SA, Isakson PC, Koboldt CM, Masferrer JL, Perkins WE, Seibert K, Zhang Y, Zweifel BS, Reitz DB (1996) J Med Chem 39:1846
36. Penning TD, Talley JJ, Bertenshaw SR, Carter JS, Collins PW, Docter S, Graneto MJ, Lee LF, Malecha JW, Miyashiro JM, Rogers RS, Rogier DJ, Yu SS, Anderson GD, Burton EG, Cogburn JN, Gregory SA, Koboldt CM, Perkins WE, Seibert K, Veenhuizen AW, Zhang YY, Isakson PC (1997) J Med Chem 40:1347
37. Huang HC, Li JJ, Garland DJ, Chamberlain TS, Reinhart EJ, Manning RE, Seibert K, Koboldt CM, Gregory SA, Anderson GD, Veenhuizen AW, Zhang Y, Perkins WE, Burton EG, Cogburn JN, Isakson PC, Reitz DB (1996) J Med Chem 39:253
38. Talley JJ, Brown DL, Carter JS, Graneto MJ, Koboldt CM, Masferrer JL, Perkins WE, Rogers RS, Shaffer AF, Zhang YY, Zweifel BS, Seibert K (2000) J Med Chem 43:775
39. Lee LF (1996) Pat. No. WO 9624585, US
40. Song Y, Connor DT, Doubleday R, Sorenson RJ, Sercel AD, Unangst PC, Roth BD, Gilbertsen RB, Chan K, Schrier DJ, Guglietta A, Bornemeier DA, Dyer RD (1999) J Med Chem 42:1151
41. Song Y, Connor DT, Sercel AD, Sorenson RJ, Doubleday R, Unangst PC, Roth BD, Gilbertsen RB, Chan K, Schrier DJ, Guglietta A, Bornemeier DA, Dyer RD (1999) J Med Chem 42:1161
42. Kurumbail RG, Stevens AM, Gierse JK, McDonald JJ, Stegeman RA, Pak JY, Gildehaus D, Miyashiro JM, Penning TD, Seibert K, Isakson PC, Stallings WC (1996) Nature 384:644
43. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) Nucleic Acids Res 28:235
44. MS, QCPE. Creative Arts Bldg. 181, Indiana University, Bloomington, IN 47405
45. Ding HQ, Karasawa N, Goddard WA III (1992) J Chem Phys 97:4309
46. Nicholls A, Honig B (1991) J Comput Chem 12:435
47. Maple JR, Hwang M-J, Stockfisch TP, Dinur U, Waldman M, Ewig CS, Hagler AT (1994) J Comput Chem 15:162
48. Dixit SB, Bhasin R, Rajasekaran E, Jayaram B (1997) J Chem Soc Faraday Trans 93:1105
49. Waller CL, Marshall GR (1993) J Med Chem 36:2390
50. Kodithala K, Hopfinger AJ, Thompson ED, Robinson MK (2002) Toxicol Sci 66:336
51. Stanton DT, Jurs PC (1990) Anal Chem 62:2323
52. Ghose A, Crippen G (1986) J Comput Chem 7:565
53. Rogers D, Hopfinger AJ (1994) J Chem Inf Comput Sci 34:854
54. Kubinyi H. (ed) (1993) 3D QSAR in drug design: theory, methods and applications, ESCOM, Leiden, pp 523–550
55. Cerius2 4.8, User's Manual. Accelrys Inc., San Diego, CA, USA
56. Vieth M, Cummins DJ (2000) J Med Chem 43:3020
57. Ryn JV, Trummelitz G, Pairet M (2000) Curr Med Chem 7:1145
58. Shimokawa T, Kulmacz RJ, DeWitt DL, Smith WL (1990) J Biol Chem 265:20073
59. Kalgutkar AS, Crews BC, Rowlinson SW, Marnett AB, Kozak KR, Rimmel RP, Marnett LJ (2000) Proc Natl Acad Sci USA 97:925
60. So OY, Scarafia E, Mak AY, Callan OH, Swinney DC (1998) J Biol Chem 273:5801
61. Greig GM, Francis DA, Falguyet JP, Ouellet M, Percival MD, Roy P, Bayly C, Mancini JA, O'Neil GP (1997) Mol Pharmacol 52:829
62. Desiraju GR, Gopalakrishnan B, Jetti RKR, Raveendra D, Sarma JARP, Subramanya HS (2000) Molecules 7:945
63. Chavatte P, Yous S, Lesieur D (2001) J Med Chem 44:3223
64. Liu H, Huang X, Shen J, Luo X, Li M, Xiong H, Chen G, Shen J, Yang Y, Jiang H, Chen K (2002) J Med Chem 45:4816
65. Datar PA, Coutinho EC (2004) J Mol Graph Model 23:239
66. Kim H-J, Chae CH, Yi KY, Park K-L, Yoo S (2004) Bioorg Med Chem 12:1629