Partially Unified Multiple Property Recursive Partitioning (PUMP-RP) Analyses of Cyclooxygenase (COX) Inhibitors

Shashidhar N. Rao and Thomas P. Stockfisch*

Accelrys, Inc., 9685 Scranton Road, San Diego, California 92121

Received August 10, 2002

We have carried out partially unified multiple property recursive partitioning (PUMP-RP) analyses on a database of cyclooxygenase (COX) inhibitors, using CART methods implemented in Cerius[2]. Three sets of physicochemical descriptors (ISIS public keys, DAYLIGHT Fingerprints, and Cerius[2]) were computed for the database molecules which were divided into two groups, assigned as training (89%) and test (11%—selected using diversity analyses tools in Cerius[2]) sets. The descriptors which led to the discrimination of active and selective COX-2 inhibitors included ISIS Key #59 (Snot%A%A), Balaban electrotopological index JY, partition coefficient AlogP, and Jurs surface area descriptors (FNSA, FPSA, and PPSA). A strong correlation is obtained between the predicted and experimental COX-2 inhibitory activity and a moderate correlation for selectivity of the COX-2 inhibitors, both in the training and test sets. Application of the RP trees to a validation set of Merck cyclooxygenase inhibitors shows good consistency with the COX-1 and COX-2 activity data, albeit moderate consistency with the selectivity data. Compared to the independent RP models (obtained by considering each activity separately), the PUMP-RP decision trees provide easier identification and interpretation of those descriptors that are common to both COX-1 and COX-2 activities. Similarly, they are easier to distinguish the descriptors that discriminate the two activities. The study represents a preliminary validation of the PUMP-RP method described in the previous article of this issue.

INTRODUCTION

Recursive partitioning methods have been useful in high throughput data analyses spanning a wide variety of fields including structure−activity relationships between drug candidates and their biological receptors. These methods typically disseminate the data into decision trees using a number of classification techniques such as CART (Classification and Regression Trees). However most of these methods deal with the classification of only one dependent variable (Y) at a time, as a function of several independent descriptors (X). The previous article in this issue[1] describes a novel approach based on CART in which multiple Y variables are disseminated into decision trees, as implemented in Cerius[2]. In this manuscript, we present an application of the PUMP-RP to the study of a database of cyclooxygenase (COX) inhibitors.

Cyclooxygenase (COX) is one of the key enzymes in the biochemical pathways leading to the production of proinflammatory prostaglandins and other isotypes of this compound, which in turn play a pivotal role in gastrointestinal protection. Researchers have identified two distinct isoforms of this enzyme in the last two decades. The first one is the constitutive form (COX-1) which is involved in the production of prostaglandins important in gastrointestinal protection.[2] The second isoform (COX-2) is induced under the conditions of immunological insult (such as inflammation).[3] Hence, selective and potent inhibitors of COX-2 are deemed to be potential antiinflammatory agents without the side-effects associated with traditional NSAIDs (Non-Steroidal Antiinflammatory Drugs) such as aspirin, which inhibit both COX-1 and COX-2 in a nonselective manner. Currently marketed selective COX-2 inhibitors, Celebrex and Vioxx, are two examples of widely prescribed antiinflammatory drugs with minimal gastrointestinal side-effects (e.g. ulcer formation).[4]

In light of the availability of extensive structure−activity relationships (SAR) on COX inhibitors, we have chosen a set of these compounds belonging to the Celebrex family,[5−18] for our recursive partitioning analyses. The main reason to adhere to a family of molecules is to have access to a pool of consistently measured biological data (i.e. the same COX-1 and COX-2 inhibition assays for all the molecules). Typically, these molecules consist of a central heterocyclic ring system with two substituted phenyl groups attached on neighboring atoms of the central ring, which in turn carries other substituents. A number of X-ray crystallographic studies on the complexes of COX-2 with its inhibitors have been reported in the literature. However, no structure-based docking study has yet been found to correlate the experimentally observed inhibitory constants to theoretical docking scores. Such a correlation can be particularly challenging in light of the fact that COX-2 is a membrane bound enzyme where significant dynamical aspects of protein domains are believed to play an important role in the phenomenon of binding and inhibition. In that light, the main goal of the present study is to obtain theoretical models (in the form of decision trees) which rationalize the biological activity and selectivity of the therapeutically desirable COX-2 inhibitors as a function of their physicochemical properties. Specifically, the focus is on identifying the descriptors which help the identification of COX-2 active and COX-2 selective (i.e.

* Corresponding author phone: (858)799-5000; fax: (858)799-5100; e-mail: tps@accelrys.com.

10.1021/ci0203803 CCC: $25.00
Published on Web 07/24/2003

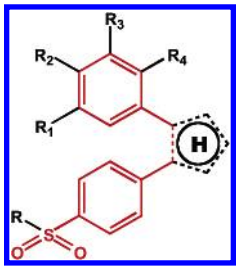

Figure 1. Schematic illustration of a typical COX inhibitor, which forms the template for all the molecules investigated in the present study. Atoms shown in red were used to overlap the 3-D structures of the molecules in the training and test sets. The various substituents (R groups) present in the compounds of the training and test sets TR and TE are listed in Chart 2. It should be noted that the library of compounds is ***not fully combinatorial*** in nature.

simultaneously COX-1 inactive) compounds. The models obtained in our analyses demonstrate a strong correlation between the experimental data and theoretical predictions. Their simplicity makes them amenable for testing an external database for potential new leads as COX-2 inhibitors.

METHODS AND NOMENCLATURE

Preliminary structures of 454 COX inhibitors used in the present investigation[5−18] were built and energy optimized using the CFF Force Field[19] in Cerius[2]. All the compounds contained a central ring in which two neighboring atoms have substituted phenyl rings attached to them (Figure 1). The central ring (designated H) was typically a heteroaromatic system (e.g. pyrrole, pyrazole, thiophene, etc.). However, some of the molecules were based on the H ring being a cyclopentene and benzene. One of the phenyl rings was always substituted with either a sulfonamide or a methyl sulfone moiety. All the compounds were aligned on the common substructure element (shown in red) in Figure 1. The energy minimized and aligned compounds were stored in an SD file format together with their experimental IC_{50} values for inhibition of COX-1 and COX-2 (available at http://www.accelrys.com). A collection of 50 most diverse compounds was obtained via a diversity analysis in the following manner: (1) the molecules were imported into a Cerius[2] study Table. (2) A standard set of 2D descriptors (COMBICHEM default, described earlier in refs 20−23]) was computed followed by a computation of the principal components. Eleven principal components were deemed adequate to explain 93% of variance in the descriptor space, the first three of which accounted for 67% of the variance. (3) A subset of the 50 most diverse molecules was selected in the space of the 11 principal components using the MaxMin diversity function described by Hassan et al.[21] This collection was regarded as the test set (TE) for the purpose of PUMP-RP analyses. (4) The remaining collection of 404 compounds was used as the training set (TR) for developing PUMP-RP trees. **It may be noted that although the test set (TE) is a diverse subset of the total data set of COX inhibitors, they are very much representative of the physicochemical space of the latter. This is because all the compounds are derived from a group of structurally and chemically similar collection of templates. Furthermore, the choice of a diverse subset (versus a random subset) for a test set was to help identify those RP trees that would be robust enough to recover truly active and selective compounds from an environment of chemical diversity. Occurrence of compounds close in the property space in a random subset could artificially bias the results.**

Chart 1

JX	Hbond acceptor	Jurs-FPSA-2	Shadow-Zlength
JY	Hbond donor	Jurs-FNSA-2	Area
Kappa-1	RadOfGyration	Jurs-FPSA-3	Vm
Kappa-2	Jurs-SASA	Jurs-FNSA-3	Density
Kappa-3	Jurs-PPSA-1	Jurs-WPSA-1	PMI-mag
Kappa-1-AM	Jurs-PNSA-1	Jurs-WNSA-1	PMI-X
Kappa-2-AM	Jurs-DPSA-1	Jurs-WPSA-2	PMI-Y
Kappa-3-AM	Jurs-PPSA-2	Jurs-WNSA-2	PMI-Z
PHI	Jurs-PNSA-2	Jurs-WPSA-3	Apol
SC-0	Jurs-DPSA-2	Jurs-WNSA-3	Dipole-mag
SC-1	Jurs-PPSA-3	Jurs-RPCG	Dipole-X
SC-2	Jurs-PNSA-3	Jurs-RNCG	Dipole-Y
SC-3_P	Jurs-DPSA-3	Jurs-RPCS	Dipole-Z
SC-3_C	Jurs-FPSA-1	Jurs-RNCS	DYFP-1024
SC-3_CH	Jurs-FNSA-1	Jurs-TPSA	ISIS_public_key
CHI-0	Rotlbonds	Jurs-TASA	Wiener
CHI-1	CHI-V-3_CH	Jurs-RPSA	log Z
CHI-2	Shadow-Ylength	Jurs-RASA	Zagreb
CHI-3_P		Shadow-XY	AlogP
CHI-3_C		Shadow-XZ	MR
CHI-3_CH		Shadow-YZ	AlogP98
CHI-V-0		Shadow-Xyfrac	LogP
CHI-V-1		Shadow-Xzfrac	Fh2o
CHI-V-2		Shadow-Yzfrac	Foct
CHI-V-3_P		Shadow-nu	MolRef
CHI-V-3_C		Shadow-Xlength	MW

Chart 2

R	R_1	R_2	R_3	R_4	H
CH_3	H	F	Cl	F	PYRROLE
NH_2	F	OCH_3	H	Cl	PYRAZOLE
	Cl	Cl	CH_3	H	IMIDAZOLE
	Br	CH_3	F	CH_3	PHENYL
	CH_3	OCF_3	OCH_3	OCH_3	THIOPHENE
	CF_3	CF_3			OXAZOLE
	OCH_3	H			CYCLOPENTENE
	SCH_3	$NHCH_3$			
	CH_2OCH_3	$N(CH_3)_2$			
	$N(CH_3)_2$	SCH_3			
	$NHCH_3$	$SOCH_3$			
	NH_2	$COCH_3$			
	NO_2				

The SD file corresponding to the training set molecules (prealigned as stated above) was loaded into a Cerius[2] Study Table. Both 2-D and 3-D descriptors (Chart 1), used as independent variables (X) in the RP analyses, were computed. The biological activity of the compounds, measured in terms of their IC_{50} values for inhibition of COX-1 and COX-2, were entered into two separate columns. The activities were classified into two categories, 0 (inactive) and 1 (active). In the case of COX-1 inhibition, molecules with class 0 corresponded to molecules which had IC_{50} values of greater than 5 μM, while in the case of COX-2 inhibition, the corresponding molecules had IC_{50} values of greater than 0.5 μM. The rest of the molecules were assigned into class 1 in both of the assays. The basis for different cutoff values of active and inactive compounds lay in the therapeutic

significance of the two enzymes. Generally, in cyclooxygenase based antiinflammatory programs, it is desirable to inhibit COX-2 preferentially over COX-1, for reasons stated earlier. Hence, it is essential to identify molecular properties associated with lower values of IC_{50} for COX-2 inhibition versus COX-1 inhibition. The classified columns of COX-1 and COX-2 inhibition data were assigned variable names Y1 and Y2, respectively, for the purposes of statistical analyses. The study tables for the training set (TR) and test set (TE) were saved into binary data files and were used in all the RP runs.

The PUMP-RP runs were carried out on the TR using the following parameters: (a) pruning factor (α) values were varied between 3 and 6; (b) four values were considered for the minimum number of samples in any node − 4, 6, 8, and 10; and (c) default values were used for the maximum tree depth (10 layers), maximum number of generic splits (30), and the number of knots per variable (20). The activity classes were weighted equally, and the splits were scored using the Gini Impurity scoring function. The RP trees were stored in .dep files, which were subsequently used to compute the predictions of COX-1 and COX-2 activities for the compounds in TE.

In the course of the analyses, two criteria were employed to define the selectivity of a compound for COX-2 inhibition, over COX-1. A compound was deemed as COX-2 selective if it showed activity against COX-2 (class 1) and was inactive against COX-1 (class 0). Further, any compound with a ratio of IC_{50}(COX-1) to IC_{50}(COX-2) greater than 100 was deemed to be COX-2 selective, irrespective of its activity class. Once the trees were built, the predicted selectivity values were computed from the predicted activity values. That is, a compound was predicted to be COX-2 selective if its predicted activity against COX-2 belonged to class 1, and the predicted activity against COX-1 belonged to class 0.

In addition to the test set described above, two validation sets were employed to test the application of the RP trees obtained in this study. The first one consisted of a collection of 25 Merck cyclooxygenase inhibitors and represents a different class of chemistry than that covered by the training and test sets, TR and TE, respectively. The second set is a group of eight well-known NSAIDs, which are nonselective and potent inhibitors of human cyclooxygenase enzymes. Since these compounds were not used in training the RP trees, the expectation for high levels of consistency with experimental observations would be relatively modest compared to TR and TE.

In addition to PUMP-RP analyses, we have carried out single-Y analyses by choosing the generality factor (β) that gave rise to a pure specific tree which by definition has zero generic nodes and one Y_k split. The single-Y RP analyses were carried out for pruning factors of 3, 4, and 5, with the minimum sample size of 4. The rest of the parameters were set to default values as described earlier (*vide infra*).

It is customary to report enrichment factors in studies of this nature. However, we have not reported these factors in this study since our training set contains a much higher percentage of active and selective compounds (as discussed below) than in a typical large database (e.g. HTS with less than 1% actives). In such a situation, enrichment factors are not high enough to be statistically significant.

RESULTS AND DISCUSSIONS

The training set TR consists of 404 compounds, 246 (61%) of which are active as COX-2 inhibitors and 86 (21%) active as COX-1 inhibitors. In this set, 181 compounds (45%) are COX-2 selective, while 21 are COX-1 selective (i.e. active in COX-1 and inactive in COX-2; ratio of the IC50(COX-2) to IC50(COX-1) > 100). A total of 202 compounds are nonselective, of which 65 compounds show activity against both COX-1 and COX-2 and the remaining 137 are inactive. Of the 50 compounds in TE, 27 (54%) are COX-2 active while only 12 molecules are COX-1 active. The number of COX-2 selective compounds in this set (17; 34%) represents a somewhat smaller percentage when compared to TR. The test set also consists of two COX-1 selective compounds and 31 nonselective compounds.

Table 1 lists the RP-predicted hits for COX-1 inhibitory activity (class 1) and COX-2 inhibitory activity (class 1), in both the TR and TE sets. Also listed is the number of true positives among the predictions (i.e. those compounds, which are active against the COX-1, and COX-2 assays using the criteria defined earlier). It is noted that for a majority of the RP trees, in the TR set, around 60−80% of the predicted COX-2 actives are true hits, while around 78−91% of the predicted COX-1 actives are true hits. In the test set (TE), a majority of the RP trees predict COX-1 inhibitory actives to contain around 60−100% true hits, while the corresponding percentages for COX-2 actives ranges from 50 to 89. In addition, in the case of the training set, about 42−65% of the compounds are predicted to have the experimental COX-1 and COX-2 inhibition classes simultaneously. For the test set, the corresponding range is from 52 to 66. Table 1 also indicates that the highly truncated RP trees corresponding to a pruning factor of 6 show poorer returns of true hits for both COX-1 and COX-2 active compounds, compared to the trees obtained with the pruning factors of 4 and 5. Hence, the RP trees corresponding to PF value of 6 were not considered for further analyses. In the case of trees with PF values of 4 and 5, the value of minimum number of samples in a node does not affect the above percentages significantly, although values of 4 and 6 for this parameter tend to return slightly higher percentages of COX-2 actives. Also, there is no systematic correlation between the variations in these percentages with the generality factor (β) values.

Table 2 lists the numbers of compounds which are predicted to be COX-2 selective for both the TE and TR sets. Also listed are the number of true positives for COX-2 selectivity (i.e. the predicted selective compounds which are experimentally COX-2 selective). Unlike the case of COX-2 inhibitory actives, the percentages of true COX-2 selective compounds are typically less than 50 for most of the trees generated. This is particularly true of trees that have greater generic character (increasing values of generality factor β).

The RP trees obtained at the pruning factor of 3 were generally very complex and difficult to use in arriving at descriptors that were deemed to be important in the determination of the activity and selectivity of COX-2 inhibitors. By contrast, the RP trees obtained with a pruning factor of 6 were oversimplified and contained typically 1 or 2 layers of descriptors before the terminal nodes were seen. Such trees are not particularly useful as their terminal nodes contain significantly higher percentages of false positives. In this

Table 1. PUMP-RP Predicted Hits (Class 1) for COX-1 and COX-2 Inhibition as a Function of the Pruning Factor α, Minimum Sample Size (MS), and Generality Factor β[a]

			COX-1 and COX-2 actives TRG (86 and 246)				COX-1 and COX-2 actives TE (12 and 27)			
α	MS	β	Pred COX-1	TP COX-1	Pred COX-2	TP COX-2	Pred COX-1	TP COX-	Pred COX-2	TP COX-2
4	4	0.280	131	67	214	178	18	6	30	21
		0.642	139	71	195	164	18	8	31	21
		0.915	173	75	173	156	28	11	18	12
		1.107	157	75	167	153	26	10	17	12
		1.593	142	73	166	152	24	10	17	12
		2.543	138	73	175	154	24	10	18	12
		4.859	144	75	158	137	25	10	15	11
4	6	0.466	192	76	231	193	30	11	31	20
		0.911	192	76	218	184	30	11	31	21
		1.523	176	76	212	181	28	10	30	21
		2.543	172	76	221	183	28	10	31	21
		>1000	178	78	204	166	29	10	28	20
4	8	0.127	152	73	211	182	23	10	18	12
		0.970	152	73	187	164	23	10	20	14
		1.236	136	73	181	161	21	9	19	14
		>1000	157	75	191	167	26	10	25	17
4	10	0.091	152	73	211	182	23	10	18	12
		0.578	152	73	210	184	23	10	20	14
		1.212	136	73	204	181	21	9	19	14
		1.614	139	74	214	187	24	10	25	17
		1.990	162	81	214	187	27	10	23	16
		>1000	169	81	214	187	25	9	23	16
5	4	2.020	218	77	256	195	33	12	35	24
		2.354	218	77	242	181	33	12	35	24
		4.391	224	79	218	162	34	12	31	22
5	6	0.814	192	76	227	186	30	11	31	21
		1.162	192	76	202	166	30	11	32	22
		1.621	176	76	196	163	28	10	31	22
		2.170	172	76	205	165	28	10	32	22
		4.390	178	78	181	146	29	10	28	20
5	8	1.058	192	76	214	177	30	11	34	23
		>1000	176	76	208	174	28	10	33	23
5	10	1.290	192	76	214	177	30	11	34	23
		>1000	176	76	208	174	26	10	33	23
6	4	6.925	106	44	100	78	24	8	21	15
6	6	3.113	192	76	186	145	30	11	27	19
6	8	>1000	100	39	193	150	21	7	29	21
6	10	1.402	192	76	214	177	30	11	34	23
		>1000	161	71	208	174	28	10	33	23

[a] TP refers to true positives in the hit list of actives (class 1). "Pred COX-1" and "Pred COX-2" correspond to the total number of compounds predicted to be COX-1 active and COX-2 active, respectively. By definition, these would then consist of the corresponding true positives and false positives. The training set TR consists of 86 COX-1 active and 246 COX-2 active compounds. The test set TE consists of 12 COX-1 active and 27 COX-2 active compounds.

light, we have focused our attention on the trees obtained with pruning factors between 3.5 and 5. To arrive at tree/s that are likely to be most useful in predicting the activity and selectivity of the test set of compounds, we carried out further RP analyses with the variations of the pruning factor between 3.5 and 5 at the interval of 0.1. The minimum sample size was retained at 4 in all these investigations, and the rest of the parameters were set to default values as described earlier. The results of the analyses on the TE and TR sets are presented in Table 3. These indicate no major shift from the patterns of hits returned as noted in Tables 1 and 2 either for the percentages of COX-2 selectivity or COX-1/COX-2 actives. It is further noted that these percentage values also do not differ significantly for pruning factors of 3.5 and 4.1. However, the trees obtained for the latter value of PF are simpler and hence more desirable to use in further analyses.

In light of these observations, and in light of the fact that COX-2 inhibitory activity and selectivity against COX-1 are the key end-points, we have chosen to use the RP trees corresponding to PF value of 4 and minimum sample size of 4, for further discussions. Figures 2 and 3 show the RP trees obtained β values of 0.3 and 1.1, respectively. In Figure

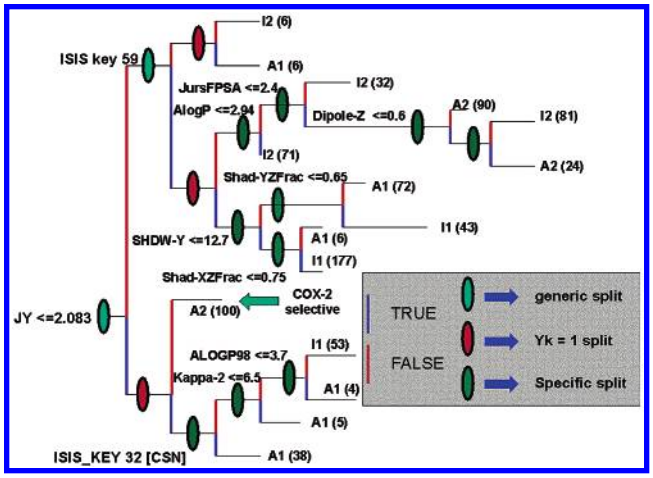

Figure 2. RP tree generated with pruning factor (α) = 4, minimum sample size 4, and generality factor (β) = 0.3. The terminal nodes marked A1, A2, I1, and I2 correspond to COX-1 active, COX-2 active, COX-1 inactive, and COX-2 inactive groups of compounds.

Table 2. COX-2 Selectivity Prediction by the PUMP-RP Trees Generated as a Function of Pruning Factor α, Minimum Sample Size (MS), and Generality Factor β[a]

			COX-2 selectivity (TRG-181)				COX-2 selectivity (TE-17)			
α	MS	β	TP	FP	FN	NS	TP	FP	FN	NS
4	4	0.280	92	41	89	182	10	9	7	24
		0.642	58	18	123	205	8	8	9	25
		0.915	35	7	146	216	2	4	15	29
		1.107	39	7	142	216	3	4	14	29
		1.593	50	11	131	212	5	4	12	29
		2.543	59	18	122	205	5	5	12	28
		4.859	37	17	144	206	3	3	14	30
4	6	0.466	50	10	131	213	2	4	15	29
		0.911	41	5	140	218	2	2	15	31
		1.523	45	5	136	218	3	2	14	31
		2.543	54	12	127	211	3	3	14	30
		>1000	32	11	149	212	1	1	16	32
4	8	0.127	73	23	108	200	4	5	13	28
		0.970	55	16	126	207	5	4	12	29
		1.236	59	16	122	207	6	4	11	29
		>1000	51	14	130	209	5	5	12	28
4	10	0.091	73	23	108	200	4	5	13	28
		0.578	74	20	107	203	5	4	12	29
		1.212	78	20	103	203	6	4	11	29
		1.614	86	20	95	203	7	5	10	28
		1.990	86	20	95	203	6	4	11	29
		>1000	78	21	103	202	6	6	11	27
5	4	2.020	39	13	142	210	3	2	14	31
		2.354	24	6	157	217	3	2	14	31
		4.391	0	0	181	223	0	0	17	33
5	6	0.814	43	13	138	210	3	3	14	30
		1.162	24	6	157	217	3	2	14	31
		1.621	28	6	153	217	4	2	13	31
		2.170	37	13	144	210	4	3	13	30
		4.390	13	7	168	216	1	1	16	32
5	8	1.058	34	8	147	215	4	3	13	30
		>1000	38	8	143	215	5	3	12	30
5	10	1.290	34	8	147	215	4	3	13	30
		>1000	38	8	143	215	5	3	12	30
6	4	6.925	0	0	181	223	0	0	17	33
6	6	3.113	0	0	181	223	0	0	17	33
6	8	>1000	59	34	122	189	4	4	13	29
6	10	1.402	34	8	147	215	4	3	13	30
		>1000	43	10	138	213	5	3	12	30

[a] TP, FP, FN, and NS correspond to numbers of true positive, false positive, false negative, and nonselective compounds, respectively.

2, the terminal leaf node #12 corresponds to a bin containing COX-2 selective compounds (blue arrow). A detailed analysis of this bin indicates that all the 53 compounds in COX-1 inactive leaf node #13 also occur in #12. Further, 47 of these compounds corresponds to true positives (that is they are selective in the experimental assays), while six others are false positives. The rest of the COX-2 selective compounds are located in other COX-2 active bins.

In Figure 3, on the other hand, we observe two different bins with COX-2 selective molecules (blue arrows). One of them occurs after a specific split (Shadow-YZFrac $< = 0.7$), while the other one occurs after a Yk split following a generic split (Jurs-PPSA-3 $< = 43$) in the left half (bottom half in the figure) of the tree. Interestingly, below this generic split, if Jurs-PPSA-3 > 43, then bins corresponding to COX-1 selective compounds are obtained; on the other hand, Jurs-PPSA-3 $< = 43$ leads to COX-2 selective bins. Thus, we are presented with an example of a situation where a certain type of inequality in the generic part of the tree could lead to selective bins. This is also indicative of the significance of the descriptor, which in this case represents the positive polar surface area of the molecules. COX-1 selective

Table 3. Percentages of True COX-2 Active [COX-2(A)], True COX-1 Active [COX-1(A)], and True COX-2 Selective [COX-2_Sel] Compounds Retrieved from the Training and Test Sets (TR and TE) by RP Trees Generated with α Values of 3.5, 3.7, 4.1, 4.3, and 4.8[a]

		training set (TR)			test set (TE)		
α	β	COX-2 (A)	COX-1 (A)	COX-2_sel	COX-2 (A)	COX-1 (A)	COX-2_sel
3.5	−0.21	72	78	51	78	50	59
	0.26	71	83	38	81	67	53
	0.76	68	87	25	48	83	18
	0.87	67	87	30	48	75	24
	1.02	72	93	34	56	75	24
	1.23	72	91	40	56	75	35
	1.39	69	92	37	52	75	35
3.7	0.01	72	78	51	78	50	59
	0.33	71	83	38	81	67	53
	0.84	68	87	25	48	83	18
	1.02	67	87	30	48	75	24
	1.25	66	85	36	48	75	35
	1.8	68	84	38	63	75	53
	2.72	69	84	40	63	75	53
4.1	0.35	72	78	51	78	50	59
	0.65	67	83	32	78	67	47
	0.93	63	87	19	48	92	12
	1.12	62	87	22	48	83	18
	1.59	62	85	28	48	83	29
	2.49	63	85	33	48	83	29
	4.80	56	87	20	41	83	18
4.3	0.62	67	83	32	78	67	47
	0.97	64	87	19	56	92	12
	1.56	63	87	22	56	83	18
	2.41	63	87	27	56	83	18
	4.69	57	90	14	52	83	6
4.8	2.09	79	90	22	89	100	18
	2.43	74	90	13	89	100	18
	4.44	66	92	0	81	100	0

[a] The percentages are relative to the total number of COX-2 active, COX-1 active, and COX-2 selective compounds in TR and TE.

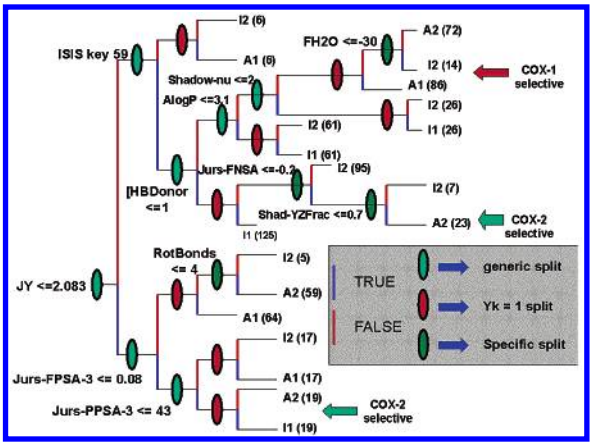

Figure 3. RP tree generated with pruning factor (α) = 4, minimum sample size 4, and generality factor (β) = 1.1. The terminal nodes marked A1, A2, I1, and I2 correspond to COX-1 active, COX-2 active, COX-1 inactive, and COX-2 inactive groups of compounds.

compounds are also seen to occur in the terminal node #5 (in the top right corner of the figure indicated by a red arrow). An analysis of the leaf nodes 12 and 13 (corresponding to [A2(23)] and [I1(125)], respectively) indicates that all the compounds in leaf node 12 occur in leaf node 13. Interestingly, as with Figure 2, all but one of the 23 compounds were true positives for COX-2 selectivity, implying a high degree of confidence for node 12 as being predictive of COX-2 selectivity.

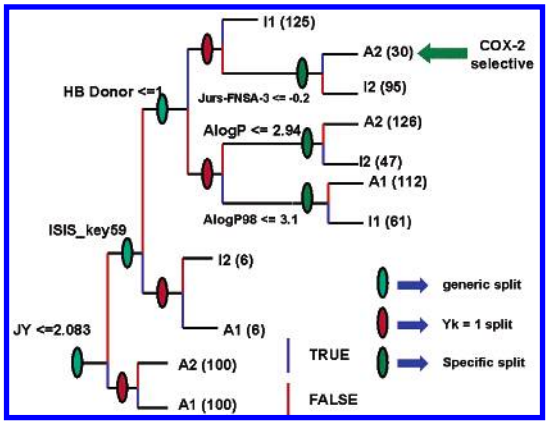

Figure 4. RP tree generated with pruning factor (α) = 5, minimum sample size 4, and generality factor (β) = 2.04. The terminal nodes marked A1, A2, I1, and I2 correspond to COX-1 active, COX-2 active, COX-1 inactive, and COX-2 inactive groups of compounds.

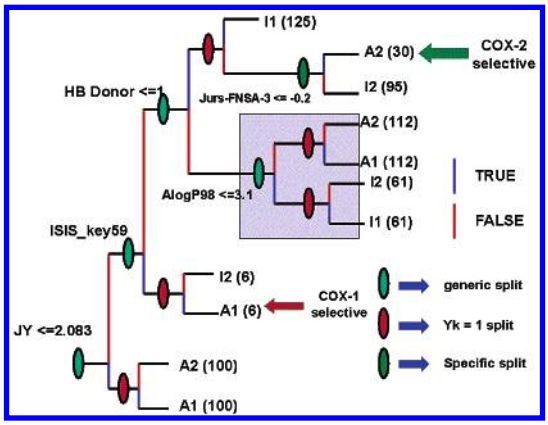

Figure 5. RP tree generated with pruning factor (α) = 5, minimum sample size 4, and generality factor (β) = 2.35 (corresponding to the next higher generic tree to the one Figure 4). The terminal nodes marked A1, A2, I1, and I2 correspond to COX-1 active, COX-2 active, COX-1 inactive, and COX-2 inactive groups of compounds.

We have also used the RP trees corresponding to PF value of 5 to illustrate tree simplification as the degree of its generic character is increased. Figure 4 shows the RP tree corresponding to $\alpha = 5$, minimum sample size of 4, and $\beta = 2.04$. The reason for the choice of this tree is based on the best retrieval percentages of COX-1 and COX-2 inhibitory activity and inactivity from the test set. Also, this tree retrieves the best percentages of the true actives (Table 1). It is noted that in the upper half of the tree, two specific splits (AlogP $< = 2.94$, AlogP98 $< = 3.1$) follow a Y_k split. However, these two correlated descriptors (which were accidentally included together in the descriptor set) have similar physicochemical meaning and their split values are also close within experimental error. Increasing the generality factor β (which increases the generic character of the tree) to 2.35 (Figure 5) leads to an expected coalescence of the two correlated descriptors, resulting in a single generic split with AlogP $< = 3.1$, followed by two Y_k splits which lead to terminal nodes. Thus, the more generic tree is simpler to interpret and has utilized the AlogP data from both the COX-1 and COX-2 inhibitors to generate the node.

The results generated in this study also indicate that despite the ease of finding a common alignment, the molecular field analyses (MFA or CoMFA) of the cyclooxygenase inhibitors would be less appropriate than the recursive partition analyses presented here. The former employ only steric and electrostatic field values as the input descriptors, while the RP studies are amenable to the incorporation of a number of different conformationally dependent and independent descriptors. The time-dependent nature of the COX inhibition by molecules of the classes represented in our study[24] is suggestive of the role of protein dynamic effects (commonly associated with membrane-associated proteins such as COX). These effects are not readily captured in only the steric and electrostatic field terms, which are more representative of the active site. Although such effects are not necessarily represented in the two-dimensional descriptors of the molecules, the latter may play a role in modulating the overall binding phenomenon. This may imply that terms other than those relating to interactions with the enzyme are important in the binding and inhibition. The occurrence of a number of two-dimensional descriptors in the RP trees (e.g. the electrotopological JY descriptor) lends strength to this observation. Furthermore, a model based on molecular field descriptors alone would be harder to use for predictions in situations, where the alignment of the test set of molecules with the training set molecules is ambiguous. In such circumstances, RP trees with alignment independent descriptors would be useful in making activity predictions.

Table 4. Predictions of COX-1 and COX-2 Inhibitory Activities (Class 1) from Single-Y RP Trees Obtained with α Values of 3, 4, and 5 in the Training and Test Sets[a]

α	COX-2 Pred	COX-2 TP	COX-2 %	COX-1 Pred	COX-1 TP	COX-1 %
Training Set (86 COX-1 Active, 246 COX-2 Active)						
3	136	123	50	163	76	88
4	173	139	57	155	73	85
5	173	139	57	155	73	85
Test Set (12 COX-1 Active, 27 COX-2 Active)						
3	19	16	59	23	10	83
4	23	19	70	22	10	83
5	23	19	70	22	10	83

[a] The columns "COX-2 Pred" (#2) and "COX-1 Pred" (#5) refer to the total number of compounds which are predicted to be COX-2 active and COX-1 active, respectively. The columns COX-2 TP and COX-1 TP refer to the number of compounds which are true positives for COX-2 and COX-1 inhibition, respectively. The data in columns #4 (COX-2%) and #7 (COX-1%) correspond to the percentage of true positives retrieved by the RP trees relative to the total number of COX-2 and COX-1 actives in the total data set.

How different are the results of the recursive partitioning analyses when the single-Y method is used, that is, when trees are obtained separately for COX-1 and COX-2 activity? To address this, we have represented the pure specific RP trees (with number of generic nodes = 0 and number of Yk splits = 1) for the pruning factor values of 3, 4, and 5, in Figure 6a−c. Interestingly, the overall success rates of correct predictions for the COX-1 and COX-2 inhibitory activities are similar to those observed in the PUMP-RP trees (Table 4). However, it is clear that when the activities against the two enzymes are analyzed separately, there is generally no common set of descriptors in the trees for COX-1 and COX-2 inhibition. This is counterintuitive since a number of the compounds in our training and test set are nonspecific and potent inhibitors of these enzymes, as are a number of NSAIDS (not included in this study). Furthermore, the crystal structures of the two enzymes share a lot of sequence and structural similarities, although they may differ in their dynamic behavior, a possible factor governing specificity.

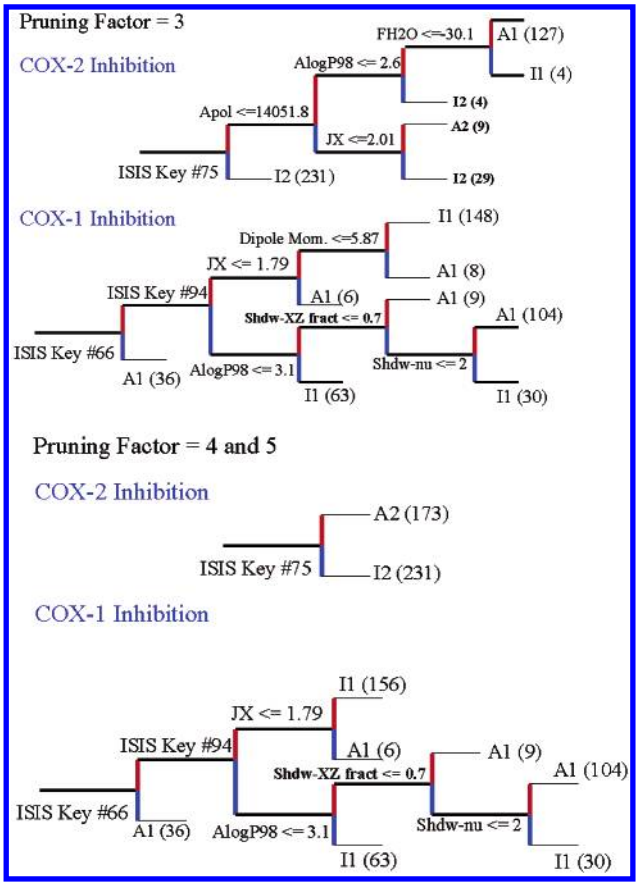

Figure 6. a: Independent RP trees obtained by the single-Y recursive partition of the COX-1 and COX-2 inhibition data, at pruning factor (α) value of 3. b: Independent RP trees obtained by the single-Y recursive partition of the COX-1 and COX-2 inhibition data, at pruning factor (α) values of 4 and 5.

Interestingly, at the PF values of 4 and 5 as seen in Figure 6b, the independent COX-2 decision tree is highly truncated and is governed by the value of only one variable, namely, the ISIS public key #75 (A!NSA). The absence of this key corresponds to COX-2 inhibitory activity, while the presence of this key corresponds to a lack of COX-2 inhibition.

As stated earlier, the training and test sets used in this study are quite rich in COX-2 active and COX-2 selective compounds (which are of therapeutic interest) and atypically high random hit rates. In that light, it is meaningless to discuss enrichment factors for the prediction of COX-2 activity and selectivity by various RP trees generated by this study. However, at the suggestion of one of the referees, we have evaluated the statistical significance of our models by computing Cohen's kappa values following the methods described in ref 25. Typically, kappa values range from 0 (random corresponding to 50% prediction in a two-class system) to 1 (perfect model). We note that in general, our models are reasonably predictive and are characterized by kappa values of between 0.3 and 0.55. In general, the RP models demonstrate higher kappa values for the training set when compared to the test set. Also, the kappa values are higher for prediction of COX-2 and COX-1 actives when compared to those for the prediction of COX-2 selectivity.

We have applied the RP trees obtained with the pruning factor of 4 and minimum sample size of 4 to a validation set of 25 COX inhibitors from Merck,[26−29] which are outside the training and test sets used in this study. Vioxx, the currently marketed Merck COX-2 selective antiinflammatory agent, is one of the members of the validation set. The structures of Merck COX inhibitors follow the general theme of diaryl heterocycles depicted in Figure 1, with the exception of a few flurbiprofen analogues.[29] All the molecules in this test set were oriented on the common template used for the training set molecules. The flurbiprofen analogues were oriented with the template in a manner consistent with their X-ray crystallographically known orientations relative to a diaryl heterocycle template. Experimentally, 13 of these compounds are COX-2 selective while the remaining 12 are nonselective. Of the latter group, four compounds are inactive against both COX-1 and COX-2.

Table 5. Analysis of 25 Merck Compounds against the RP Tree Obtained Using a Pruning Factor (α) of 4.0 and a Minimum Sample Size of 4[a]

	COX-2 Sel		NS-A		COX-1 Sel		COX-2		COX-1	
β	Pred	TP	Pred	TP	Pred	TP	Pred	TP	Pred	TP
0.280	7	4	18	7	0	0	25	21	18	7
0.642	7	4	18	7	0	0	25	21	18	7
0.915	2	2	16	7	7	0	18	16	23	8
1.107	2	2	16	7	7	0	18	16	23	8
1.593	5	3	14	7	6	0	19	16	20	7
2.543	5	3	14	7	6	0	19	16	20	7
4.859	5	3	14	7	6	0	19	16	20	7

[a] TP refers to true positives. NS refers to nonselective compounds that are active against both COX-1 and COX-2 enzymes. Experimentally, 13 compounds are COX-2 active and selective against COX-1, while 8 compounds are both COX-1 and COX-2 active (hence nonselective) and the remaining 4 are COX-2 inactive.

The predictions from RP trees are listed in Table 5 in terms of the number of COX-2 selective, COX-1 selective, nonselective, COX-2 active, and COX-1 active compounds. Interestingly, all the trees identify Vioxx to be COX-2 active inhibitor. However, none of the trees identify it as a COX-2 selective compound. Interestingly, the tree represented in Figure 2 identifies Vioxx as belonging to leaf #12, which is one of the COX-2 selective leaves in this tree. The results in Table 5 indicate that higher numbers of false negatives are obtained for COX-2 selectivity than for COX-2 activity. No more than four (∼31%) of the truly COX-2 selective compounds are identified by all the trees. On the other hand, the RP trees do capture a significant portion of the COX-2 and COX-1 actives.

We have also studied the prediction made by the above RP tree on the activities of a group of eight NSAIDs (aspirin, ketoprofen, naproxene, desmethylnaproxene, ibuprofen, indomethacin, phenytoin, and diclofenac).[30] As with the case of the Merck flurbiprofen analogues, these compounds were also oriented to the diaryl heterocycle template of the training set consistent with X-ray crystallographic data. All of these compounds are predicted to be nonselective, while five of them (ketoprofen, naproxene, ibuprofen, indomethacin, and diclofenac) are predicted to be active as inhibitors of COX-1 and COX-2 enzymes, consistent with known experimental observations on them. Although aspirin is a widely used antiinflammatory agent, it inhibits both these enzymes weakly (IC_{50} ∼ 150−300 nM). Our studies predict this compound to be inactive against both the enzymes, based on the criteria defined above.

We have further assessed the significance of our models by evaluating them against a random collection of com-

Table 6. Prediction of COX-1 and COX-2 Activity and Selectivity of a Random Collection of WDI Compounds, by the PUMP-RP Trees Generated with an α Value of 4 and a Minimum Sample Size of 4[a]

α	MS	β	COX-1 inhibition	COX-2 inhibition	COX-1 selective	COX-2 selective	NS active
4	4	0.280	467	348	227	108	240
		0.642	465	348	225	108	240
		0.915	469	163	410	104	59
		1.107	469	163	410	104	59
		1.593	438	207	366	135	72
		2.543	438	207	366	135	72
		4.859	441	209	366	134	75

[a] Please note that no experimental data are available for the COX-1 and COX-2 inhibition by the WDI Compounds. Hence all the data refers to predictions by the RP trees.

pounds from the World Drug Index (WDI).[31] Our validation set consisted of 586 compounds with molecular weights less than 500. The results of the analyses carried out using the RP trees generated with a PF value of 4 and minimum sample size of 4 are described in Table 6. This table lists the predicted values for the number of potential COX-1 and COX-2 inhibitors, selectivity for COX-1 and COX-2, and the number of nonspecific compounds. The latter category is divided into two classes, one where compounds are predicted to be active inhibitors of both the isoforms of the cyclooxygenase, while the other corresponds to compounds, which are inactive as inhibitors of both isoforms. Although varying numbers of compounds are predicted to be COX-2 inhibitors, less than 25% of the compounds are predicted to be COX-2 selective. By contrast, a larger percentage of compounds (35−60) are predicted to be COX-1 selective. Further analysis indicates that only 28 compounds are commonly predicted to be selective by all seven trees, leading to the implication that these are most likely to possess both COX-2 inhibitory activity and selectivity against COX-1.

We have also examined the leaf locations for the predicted COX-2 selective compounds in two of the trees (rows 1 and 3 in Table 6) schematically illustrated in Figures 2 and 3. In the case of the tree in Figure 2, all but one of the COX-2 selective compounds fall in leaf #12. Interestingly, the corresponding COX-1 inactivity is due to their location in leaf #13. It may be noted that 78% of the 100 compounds in leaf #12 are true COX-2 active while the rest are COX-2 inactive, whereas nearly 91% of the 53 compounds in leaf #13 are COX-1 inactive. However, the 53 selective compounds represented in these two leaves constitute about 30% of all the COX-2 selective inhibitors in the entire training set TR. Thus, the WDI compounds whose COX-2 activity and COX-1 inactivity are due to their locations in leaves #12 and #13, have about a 22% ($0.78 \times 0.91 \times 0.3$) chance of being COX-2 selective. In the case of the tree in Figure 3, the predicted COX-2 selective compounds all fall in leaves 19 and 20. Interestingly, as mentioned earlier, these bins have the origin of their selectivity in the generic part of the tree. Also, the number of selective compounds from the TR in these leaves is only around 10% of the true COX-2 selective compounds. Hence, the confidence level in assigning COX-2 selectivity of WDI compounds, from this tree is markedly lower than in the tree in Figure 2.

It must be pointed out that a majority of the descriptors that characterize the RP trees obtained in this investigation are independent of the orientation of the molecules relative to the diaryl heterocyclic template alluded to earlier. As indicated in Figures 2−6, the only orientation sensitive descriptor found in our RP trees is the magnitude of the Z-component of total dipole moment (Figure 2). In this light, the models in this study would be applicable to any test set where the molecules are oriented relative to the training set described earlier. Thus in this study, the RP tree Figure 2 would be less applicable to the case of the test set with WDI compounds, which were not preoriented to the training set template. However, the other RP trees represented in Figures 3−5 would be potentially universally applicable for understanding COX-2 activity and selectivity.

CONCLUSIONS

We have employed PUMP-RP methods to analyze the structure−activity relationships in cyclooxygenase inhibitors. The decision trees are highly predictive of the COX-1 and COX-2 activity and moderately predictive of the therapeutically desirable COX-2 selectivity. The enrichment factors are typically less than 3 for the data sets employed as they contain a significant portion of actives, atypical of high throughput screen databases. The decision trees contain useful information that leads to criteria useful in predicting COX-2 selectivity as well as COX-1 selectivity.

ACKNOWLEDGMENT

We are grateful to Dr. John Talley (Monsanto Company) for useful discussions, which were helpful in the choice of the training set used in this study. We thank Dr. Marvin Waldman (Accelrys) for many useful comments and suggestions during the preparation of the manuscript.

REFERENCES AND NOTES

(1) Stockfisch, T. P. *J. Chem. Inf. Comput. Sci*. **2003**, *43*, 1608−1613.
(2) (a) Vane, J. R. *Nature (New Biol.)* **1971**, *231*, 232−235. (b) Smith, J. B.; Willis, A. L. *Nature (New Biol.)* **1971**, *231*, 235−237.
(3) (a) Needleman, P.; Isakson, P. C. *J. Rheumatol*. **1997**, *49*, 6−8. (b) Masferrer, J.; Zweifel, B.; Seibert, K.; Needleman, P. *Proc. Natl. Acad. U.S.A.* **1992**, *89*, 3917−3921.
(4) (a) FitzGerald, G. A.; Patrono, C. *New England J. Med*. **2001**, *345*(6): 433−442. (b) Habeeb, A.; Praveen Rao, P. N.; Knaus, E. E. *J. Med. Chem*. **2001**, *44*, 3039−3042.
(5) Reitz, D. B.; Huang, H.-C.; Li, J. J.; Garland, D. J.; Manning, R. E.; Anderson, G.; Gregory, S.; Koboldt, C. M.; Perkins, W. E.; Seibert, K.; Isakson, P. C. *Bioorg. Med. Chem. Lett*. **1995**, *5*, 867−872.
(6) Reitz, D. B.; Li, J. J.; Norton, M.; Reinhard, E. J.; Collins, J. T.; Anderson, G.; Gregory, S.; Koboldt, C. M.; Perkins, W. E.; Seibert, K.; Isakson, P. C. *J. Med. Chem*. **1994**, *37*, 3878−3881.
(7) Huang, H.-C.; Chamberlain, T. S.; Seibert, K.; Koboldt, C. M.; Isakson, P. C.; Reitz, D. B. *Bioorg. Med. Chem. Lett*. **1995**, *5*, 2377−2380.
(8) Bertenshaw, S. R.; Talley, J. J.; Rogier, D. J.; Graneto, M. J.; Rogers, R. S.; Kramer, S. W.; Penning, T.; Koboldt, C. M.; Veenhuizen, A. W.; Zhang, Y. Y.; Perkins, W. E. *Bioorg. Med. Chem. Lett*. **1995**, *5*, 2919−22.
(9) Li, J. J.; Anderson, G.; Burton, E.; Cogburn, J. N.; Collins, J. T.; Garland, D.; Gregory, S.; Huang, H.-C.; Isakson, P. C.; Koboldt, C. M.; Logusch, E. W.; Norton, M. B.; Perkins, W.; Reinhard, E. J.; Seibert, K.; Veenhuizen, A. W.; Zhang, Y. Y.; Reitz, D. B. *J. Med. Chem*. **1995**, *38*, 4570−4578.
(10) Khanna, I. K.; Weier, R. M.; Yu, Y.; Xu, X.; Koszyk, F. J.; Collins, P. W.; Koboldt, C. M.; Veenhuizen, A. W.; Perkins, W. E.; Casler, J. J.; Masferrer, J. L.; Zhang, Y.; Gregory, S.; Seibert, K.; Isakson, P. *J. Med. Chem*. **1997**, *40*, 1634−1647.
(11) Li, J. J.; Norton, M. B.; Reinhard, E. J.; Anderson, G.; Gregory, S.; Isakson, P. C.; Koboldt, C. M.; Masferrer, J. L.; Perkins, W.; Seibert, K.; Zhang, Y. Y.; Zweifel, B. S.; Reitz, D. B. *J. Med. Chem*. **1996**, *39*, 1846−1856.

(12) Khanna, I. K.; Weier, R. M.; Yu, Y.; Collins, P. W.; Miyashiro, J. M.; Koboldt, C. M.; Veenhuizen, A. W.; Currie, J. L.; Seibert, K.; Isakson, P. *J. Med. Chem.* **1997**, *40*, 1619−1633.
(13) Huang, H.-C.; Li, J. J.; Garland, D. J.; Chamberlain, T. S.; Reinhard, E. J.; Manning, R. E.; Seibert, K.; Koboldt, C. M.; Gregory, S. A.; Anderson, G. D.; Veenhuizen, A. W.; Zhang, Y. Y.; Perkins, W. E.; Burton, E. G.; Cogburn, J. N.; Isakson, P. C.; Reitz, D. B. *J. Med. Chem.* **1996**, *39*, 253−266.
(14) Penning, T. D.; Talley, J. J.; Bertenshaw, S. R.; Carter, J. S.; Collins, P. W.; Docter, S.; Graneto, M. J.; Lee, L. F.; Malecha, J. W.; Miyashiro, J.; Rogers, R.; Rogier, D. J.; Yu, S. S.; Anderson, G. D.; Burton, E.; Cogburn, J. N.; Gregory, S.; Koboldt, C. M.; Perkins, W.; Seibert, K.; Veenhuizen, A. W.; Zhang, Y. Y.; Isakson, P. C. *J. Med. Chem.* **1997**, *40*, 1347−1365.
(15) Talley, J. J.; Bertenshaw, S. R.; Brown, D. L.; Carter J.; Graneto M. J.; Koboldt, C. M.; Masferrer, J. L.; Norman, B. H.; Rogier, D. J., Jr.; Zweifel, B. S.; Seibert, K. *Med. Res. Rev.* **1999**, *19*(3), 199−208.
(16) Khanna, I. K.; Yu, Y.; Huff, R.; Weier, R. M.; Xu, X.; Koszyk, F. J.; Collins, P. W.; Cogburn, J. N.; Isakson, P.; Koboldt, C. M.; Masferrer, J.; Perkins, W.; Seibert, K.; Veenhuizen, A. W.; Yuan, J.; Yang, D.-C.; Zhang, Y. Y. *J. Med. Chem.* **2000**, *43*, 3168−3185.
(17) Carter, J. S.; Kramer, S.; Talley, J. J.; Penning, T.; Collins, P.; Graneto, M. J.; Seibert, K.; Koboldt, C. M.; Masferrer, J.; Zweifel, B. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1171−1174.
(18) Talley, J. J. *Prog. Med. Chem.* **1999**, *36*, 201−234.
(19) Cerius^2 (version 4.6), Accelrys, Inc., San Diego, CA 92121, 2001.
(20) Waldman, M.; Li, H.; Hassan, M. *J. Mol. Graphics Model.* **2000**, *18*, 412−426.
(21) Hassan, M.; Bielawski, J. P.; Hempel, J. C.; Waldman, M. *Mol. Diversity* **1996**, *2*, 64−74.
(22) C2.*Descriptor*+ version 4.6, Accelrys, Inc.; San Diego, CA 92121, 2001.
(23) Brown, R. D.; Hassan, M.; Waldman, M. *J. Mol. Graphics Model.* **2000**, *18*, 427−437.
(24) Ouellet, M.; Percival, M. D. *Biochem. J.* **1995**, *306*, 247−251.
(25) Cohen, J. *Educ. Psych. Meas.* **1960**, *20*, 37−46.
(26) Riendeau, D.; Percival, M. D.; Brideau, C.; Charleson, S.; Dube, D.; Ethier, D.; Falgueyret, J. P.; Friesen, R. W.; Gordon, R.; Greig, G.; Guay, J.; Mancini, J.; Ouellet, M.; Wong, E.; Xu, L.; Boyce, S.; Visco, D.; Girard, Y.; Prasit, P.; Zamboni, R.; Rodger, I. W.; Gresser, M.; Ford-Hutchinson, A. W.; Young, R. N.; Chan, C. C. *J. Pharmacol. Exp. Ther.* **2001**, *296*(2), 558−566.
(27) Chauret, N.; Yergey, J. A.; Brideau, C.; Friesen, R. W.; Mancini, J.; Riendeau, D.; Silva, J.; Styhler, A.; Trimble, L. A.; Nicoll-Griffith, D. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*(8) Apr 23, 1059−62.
(28) Black, W. C.; Brideau, C.; Chan, C. C.; Charleson, S.; Chauret, N.; Claveau, D.; Ethier, D.; Gordon, R.; Greig, G.; Guay, J.; Hughes, G.; Jolicoeur, P.; Leblanc, Y.; Nicoll-Griffith, D.; Ouimet, N.; Riendeau, D.; Visco, D.; Wang, Z.; Xu, L.; Prasit, P. 2,3-Diarylcyclopentenones as orally active, highly selective cyclooxygenase-2 inhibitors. *J. Med Chem.* **1999**, *42*(7) Apr 8, 1274−81.
(29) Bayly, C. I.; Black, W. C.; Leger, S.; Ouimet, N.; Ouellet, M.; Percival, M. D. *Bioorg. Med. Chem. Lett.* **1999**, *9*(3), 307−12.
(30) Williams, M.; Kowaluk, E. A.; Arneric, S. P. *J. Med. Chem.* **1999**, *42*(9), 1481−1500.
(31) Derwent World Drug Index (Derwent, WDI), 1997, Derwent Information, 14 Great Queen Street, London WC2B 5DF, United Kingdom.

CI0203803