RESEARCH PAPER

Accepted 18 July 2004
Revised 2 April 2004
Received 8 October 2003
Indian J. Pharm. Sci., 2004, 66(5): 613-620

3D-QSAR CoMFA CoMSIA Studies on Indomethacin Derivatives as Selective Cyclooxygenase-2 Inhibitors

M. AROCKIA BABU, M. ELIZABETH SOBHIA[1], MEENAKSHI SHARMA[1], SMRITI KHANNA[1], P. V. BHARATAM[1] AND S. G. KASKHEDIKAR*
Department of Pharmacy, S. G. S. I. T. S, Indore-452003.
[1]Medicinal Chemistry Division, NIPER., S. A. S Nagar (Mohali)-160062.

A set of sixty-four compounds of indomethacin derivatives was subjected to three-dimensional quantitative structure activity relationship analysis using comparative molecular field analysis and comparative molecular similarity indices methods. The comparative molecular field analysis model gave cross-validated and conventional r^2 values of 0.742 and 0.992, respectively, for the forty-five compounds of training set with optimum number of components as 6. The comparative molecular similarity indices model has cross-validated and conventional r^2 values of 0. 594 and 0.935, respectively, for the forty-five compounds of training set with optimum number of components as 4. Nineteen compounds were used to validate the comparative molecular field analysis and comparative molecular similarity indices models, which were not included in the model generation. The comparative molecular field analysis and the comparative molecular similarity indices results have been compared. The comparative molecular field analysis model was found to be highly predictive and it can be used to design potent cyclooxygenase-2 inhibitors prior to their synthesis.

The quest for selective cyclooxygenase-2 (COX-2) inhibitors began when this enzyme was first described in 1990[1-6]. It is now well established that cyclooxygenase exists in at least two isoforms, COX-1 and COX-2. COX-1 is a regulatory enzyme that is present almost throughout the body and produces prostaglandins that help in proper functioning of kidney and stomach. COX-2 is an inducible enzyme that catalyzes the production of prostaglandins leading to inflammation[7,8]. Most of the existing non-steroidal antiinflammatory drugs (NSAIDs) suffer from non-selective binding to both COX-1 and COX-2, thus producing undesirable gastrointestinal and renal side effects. It is believed that inhibitors that bind with a greater or lesser extent of selectivity towards COX-2 are different from the conventional NSAIDs. Further more selective COX-2 inhibitors are believed to play a vital role in ovulation and labour, as well as in the treatment of colon cancer and Alzheimer's disease[9-11].

*For correspondence
E-mail: babuphd2001@yahoo.co.in

During the last decade, a number of diaryl heterocyclic and other compounds have been extensively studied as selective COX-2 inhibitors[12]. Among them, two compounds, celecoxib and rofecoxib have been marketed for the treatment of acute pain, osteoarthritis and rheumatoid arthritis. However the structures of only few compounds such as aspirin, indomethacin, meclofenamic acid, zomepirac and flurbiprofen have been modified to have selective COX-2 inhibitory activity[13-17]. These were found to have similar efficacy to that of selective COX-2 inhibitors and greater gastro intestinal safety. However, due to lack of structural requirements for selective inhibition of COX-2 enzyme, these attempts were not so fruitful. In the recent past, number of research groups have independently tried to modify the structure of indomethacin, this included glycerol ester and amide, thiazole analogues of indomethacin, indolalkonic acid derivatives and extending the acetic acid side chain along with

replacement of the N-benzoyl moiety of indomethaicn[18, 19]. Thus, these structural modifications of indomethacin have led to the clinical trial stage of L-761, 000, as potent and selective COX-2 inhibitory activity[20]. Based on these results, it was thought worthwhile to design COX-2 inhibitors from classical NSAIDs, by focusing attention on indomethacin derivatives. Kalgutkar *et al.* has recently reported ester and amide derivatives of indomethacin, which have exhibited highly selective COX-2 inhibitory activity[21]. To further explore the structural requirements of amide and ester derivatives of indomethacin for selective COX-2 inhibitory activity, two methods of three-dimensional quantitative structure activity relationship namely, Comparative molecular field analysis (CoMFA) and Comparative molecular similarity index analysis (CoMSIA) were performed. No QSAR studies have been reported so far on amide and ester derivatives of indomethacin. The validated models gave important structural insight in terms of contour maps to aid design of COX-2 inhibitors prior to synthesis.

MATERIALS AND METHODS

The IC_{50} values of 64 compounds were found in the literature[21]. The IC_{50} values had been obtained using the *in vitro* biological method on human COX-2 enzyme and these were converted in to pIC_{50} for our study. To assess the predictive power of the models, a training set of 45 compounds, together with a test set of 19 molecules were selected randomly with respect to their activity. (Tables 1 and 2). Molecular modeling studies were carried out using SYBYL 6.8[22], implemented on SGI octane-2 workstation. CoMFA[23-30], the most commonly used 3D-QSAR methods, is part of the QSAR module in SYBYL. The three-dimensional structures of all compounds were built from standard library available in the module and the molecules were subjected to energy minimization using Tripos force field. The minimized template molecule was used for building the rest of the molecules in the series. All the molecules were optimized with GAUSSION 98 using the Austin model-1 (AM-1) semi-empirical method. The key words OPT, PRECISE, were used in the AM_1 optimization procedure. Electrostatic term was calculated using Gasteiger-Hückel atomic charges[31,32].

One of the most important parameters in CoMFA is the relative alignment of all the compounds to one another so that they have a comparable conformation and a similar orientation in space. In this series, the most active molecule 11 was used as the reference molecule. Nine features were selected for aligning the compounds as numbered atom 1 to 9 in fig. 1.

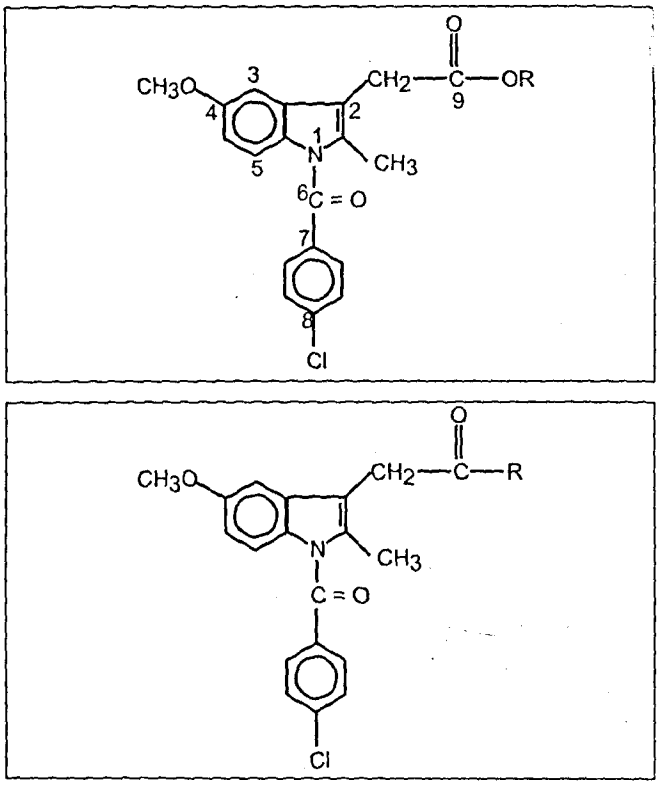
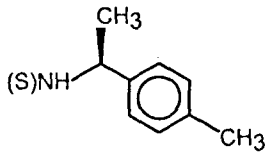


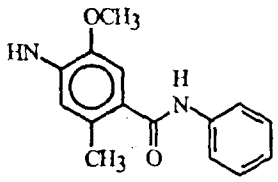
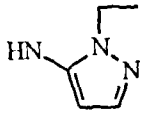
Fig. 1: General structures of ester and amide derivatives of indomethacin used for 3D-QSAR analysis.

CoMFA interaction fields were calculated at each lattice intersection of a regularly spaced grid of 2.0Å. The CoMFA fields[33], depicting the steric and electrostatic interaction with an sp^3 carbon atom with +1.0 charge as the probe were calculated using Tripos force field. The steric and electrostatic fields were truncated at ±30.0 kcal/mol and the electrostatic fields were ignored at the points with maximal steric interactions. The regression analysis of CoMFA field energies was performed using the partial least squares (PLS) algorithm and the leave-one-out (LOO) method was adopted for cross-validation. The cross validation was performed to obtain the optimum number of components which were then used in deriving final CoMFA model without cross validation. The column filtering value (σ) was set to 2.0 for cross-validated runs. Equal weights were assigned to steric and electrostatic fields. Final analysis was carried out to calculate the conventional r^2 value using the optimum number of components.

The recently reported CoMSIA method is based on molecular similarity indices; this method overcomes some of the drawbacks arising from the functional form of the Lennard-Jones and Coulomb potentials[34] used in CoMFA.

TABLE 1: THE TRAINING SET WITH EXPERIMENTAL (PIC_{50}) AND CALCULATED VALUES

NO	R	Act.[a]	Cal.[b]	Cal.[c]
1	CH_3	6.60	6.63	6.76
2	C_2H_5	7.00	6.79	6.81
3	C_3H_7	7.00	7.00	6.96
4	i -C_3H_7	6.60	6.64	6.77
5	C_6H_{13}	7.22	7.33	7.35
6	C_7H_{15}	7.40	7.39	7.40
7	*Trans*- $CH_2CH{=}CH\ (CH_2)_2CH_3$	7.30	7.30	7.22
8	$CH_2C^{\circ}C\ (CH_2)_3CH_3$	6.60	6.60	6.93
9	$(CH_2)_2NHCOOC(CH_2)_3$	7.35	7.34	7.43
10	α -$C_{10}H_7$	5.30	5.32	5.29
11	$(CH_2)_2C_6H_5$	7.40	7.42	7.44
12	C_6H_4 (4-OCH_3)	7.40	7.39	7.47
13	C_6H_4 (4-$NHCOCH_3$)	7.30	7.21	7.27
14	C_6H_4 (4-F)	7.12	7.21	7.33
15	3-Pyridyl	7.30	7.25	7.22
16	$NHCH_3$	6.15	6.24	6.48
17	$N(CH_3)_2$	4.74	4.72	5.06
18	$N(C_2H_5)_2$	4.60	4.55	4.59
19	NHC_8H_{17}	7.40	7.43	7.34
20	NHC_9H_{19}	7.40	7.46	7.38
21	$NHCH_2CH_2CH_2Cl$	7.30	7.24	6.87
22	$NH\ CH_2CH_2OH$	6.60	6.62	6.86
23	$NHCH_2COOCH_3$	5.40	5.39	5.20
24	(D)-$NHCH\ (CH_3)\ COOCH_3$	6.40	6.41	6.58
25	(L)-$NHCH\ (CH_3)\ COOCH_3$	6.72	6.73	6.84
26	$NH\ (CH_2)_2C_6H_5$	7.22	7.14	7.01
27	NH_2	6.15	6.18	6.11
28	$NHCH_2C_6H_4$ (2-CH_3)	6.82	6.77	6.88
29		6.70	6.73	6.75

30	$NHCH_2C_6H_4$ (4-$COCH_3$)	7.10	7.09	7.10
31	NHC_6H_4 (4-F)	7.22	7.09	7.10
32	NHC_6H_4 (4-Cl)	7.26	7.15	6.90
33	NHC_6H_4 (4-SCH_3)	6.92	6.94	6.95
34	NHC_6H_4 (3-SCH_3)	6.66	6.71	6.84
35	NHC_6H_4 (3-OC_2H_5)	6.19	6.13	6.28
36	NHC_6H_4 (4-$NHCOCH_3$)	6.92	6.96	6.79
37	NHC_6H_4 (4-CH_2COOCH_3)	7.24	7.25	7.08
38	NHC_6H_4 (4-$CONH_2$)	6.49	6.52	6.48
39		6.22	6.24	6.09
40	NHC_6H_4 (4-C_6H_5)	6.30	6.38	6.52
41	NH (3-Pyridyl)	7.28	7.33	7.12
42	NH (5-Chloro-3-pyridyl)	7.32	7.23	7.18
43		6.15	6.22	5.78
44	$NHOCH_2C_6H_5$	7.30	7.31	7.23
45	$NHOCH_2C_6H_4$ (4-NO_2)	7.22	7.20	7.27

The structure of training set compounds 1 to 15 and 16 to 45 are ester and amide derivatives of indomethacin, respectively (Fig. 1), Act.[a]- Observed activity, Cal.[b]-Calculated activity using CoMFA model and Cal.[c]-Calculated activity using CoMSIA model.

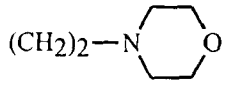
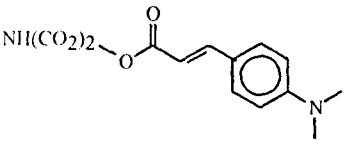
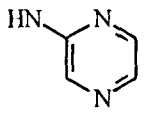
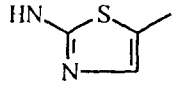
Molecular similarity is expressed in terms of five different properties viz., steric, electrostatic, hydrophobic, hydrogen-bond donor and acceptor properties. Using a common probe atom, similarity indices are calculated for a data set of pre-aligned molecules at regularly spaced grid points[35]. Similarity indices A_F between the compounds of interest and a probe atom are calculated according to (e.g. at grid point q for molecule j of the data set) the equation, $A_F{}^q(j) = -\sum w_{probe,k}\, w_{ik} \exp(-\alpha r^2_{iq})$

Where i is the summation index over all atoms of the molecule j, $w_{i,k}$ is actual value of the physicochemical property k of atom I, $w_{probe,k}$ is the probe atom with charge +1, radius 1Å, hydrophobicity +1, H-bond donor acceptor property +1, α is the attenuation factor and r_{iq} is the distance between probe atom at grid point q and atom i of the test molecule.

The value of the so-called attenuation factor a was set

to 0.3. Similarity indices can be calculated at all grid points inside as well as outside the molecules and are subsequently evaluated in a PLS analysis following the usual CoMFA protocol.

TABLE 2: TEST SET

No	R	Act.[a]	Cal.[b]
1	C_4H_9	7.30	7.16
2	C_5H_{11}	7.30	7.25
3	C_6H_{11}	6.03	6.77
4	CH_2CH_2C6H11	6.00	6.65
5	$(CH_2)_2O(CH_2)_3CH_3$	6.78	6.76
6	$CH(CH_3)CH_2CCCH_2CH_3$	6.92	7.29
7	C_8H_{17}	7.04	7.41
8		6.76	7.35
9	C_6H_5	6.39	7.24
10	C_6H_4 (4-SCH_3)	6.52	7.41
11	C_6H_4 (2-SCH_3)	7.22	6.56
12		6.73	6.10
13	$NHCH_2C_6H_4$ (4-CH_3)	7.22	6.83
14	(R) NHCH (CH_3) C_6H_4 (4-CH_3)	7.22	6.90
15	NHC_6H_4 (4-OCH_3)	7.25	6.84
16	NH (2-Chloro 3-pyridyl)	7.30	6.10
17		5.39	6.12
18		5.39	6.20
19	$NHNHCH_2C_6H_5$	5.60	6.14

The structure of test set compounds 1 to 11 and 12 to 19 are ester and amide derivatives of indomethacin, respectively (Fig. 1) and Cal.[b]-Calculated activity using CoMFA model.

RESULTS AND DISCUSSION

The major objective of CoMFA and CoMSIA analysis is to find the best predictive model for the series. Partial Least Square (PLS) was used to derive the 3D QSAR models, on a series of forty five compounds of indomethacin derivatives as training set. The experimental and calculated pIC_{50} values for the training set are shown in Table 1. Predictive 3D QSAR models were established using SYBYL fit atom molecular alignment rule, which had conventional r^2 and cross-validated coefficient respectively as 0.992 and 0.742 for the CoMFA and 0.935 and 0. 594 for CoMSIA. A cross-validated r^2 obtained as result of this analysis, served as a quantitative measure of predictive ability of the final QSAR models. The CoMFA model has better predictive ability than the CoMSIA model. The r^2_{cv} value is a statistical indication of how well a model can predict the activity for members left out of the model formation. In contrast, the conventional r^2 is simply a reflection of how well the fit equation reproduces input values. The relative contributions of steric and electrostatic molecular fields were 54% and 46 %, respectively. The boot strapping reflects the accuracy of our models with r^2_{bs} being 0.993 and 0.935 for CoMFA and CoMSIA, respectively.

To validate our model, we attempted to predict the activities of the test set containing 19 compounds. The predicted pIC_{50} values for the test set are given in the Table 2. The results of PLS analysis are summarized in the Table 3.

The CoMFA contour plots show green colored regions where increased steric bulk is associated with enhanced activity and yellow colored regions where increased steric bulk is associated with diminished activity. The region where increased positive potential is favorable for activity is indicated in blue, while regions where increased negative potential is favorable for activity are indicated in red.

In the case of CoMSIA contour, greater value of bioactivity measurements are correlated with more bulk near green, less bulk near yellow, more positive potential near blue, and more negative potential near red. In a similar way, more hydrophobic groups near orange, more hydrophilic groups near white, hydrogen bond donor favour near cyan, hydrogen bond disfavor near purple, hydrogen bond acceptor favour near magenta and hydrogen bond acceptor disfavor near red.

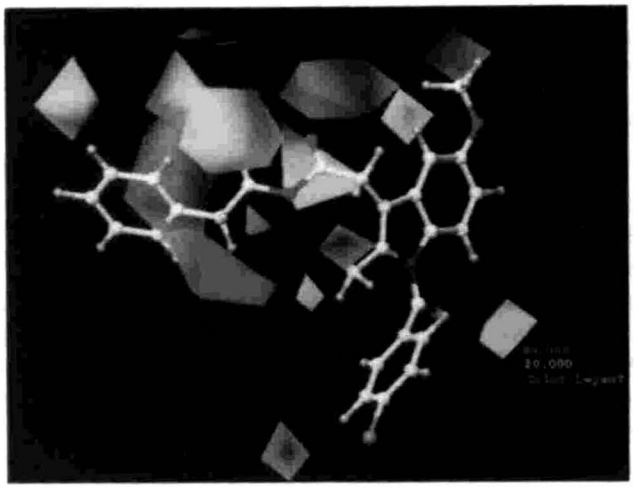


Fig. 2: CoMFA steric contour field plot.

Green contours indicate regions where bulky groups increase activity, whereas yellow contours indicate regions where bulky groups decrease activity. The most active molecule 11 is shown in the background.

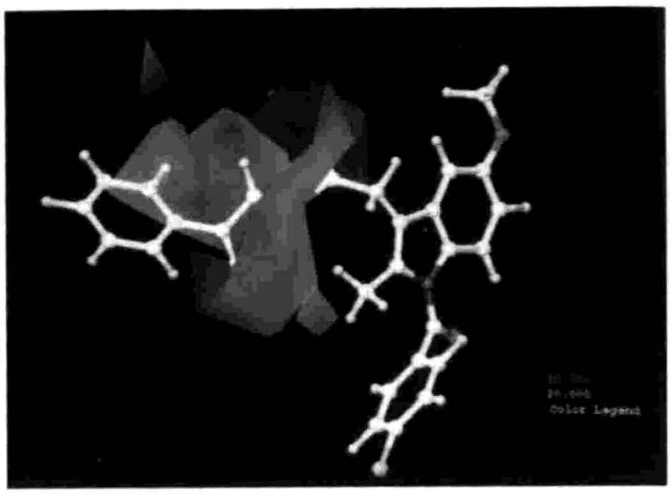


Fig. 3: CoMFA electrostatic contour field plot.

Blue contours indicate regions where positive electrostatic groups increase activity, whereas red contours indicate regions where negative electrostatic groups increase activity.

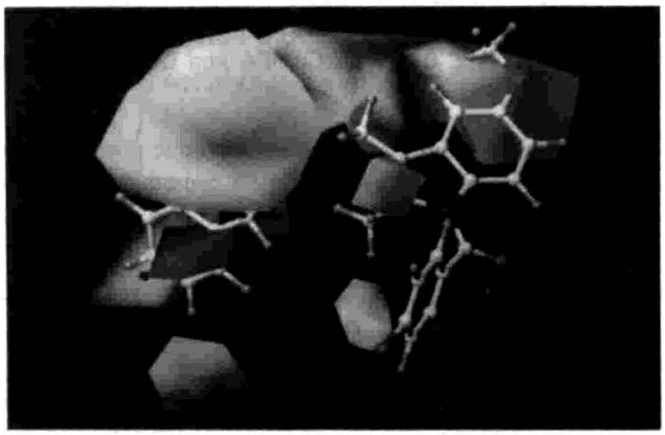


Fig. 4: CoMSIA steric and electrostatic contour field plots.

Green contours indicate regions where bulky groups increase activity, whereas yellow contours indicate regions where bulky groups decrease activity. Blue contours indicate regions where positive groups increase activity, whereas red contours indicate regions where negative groups increase activity.

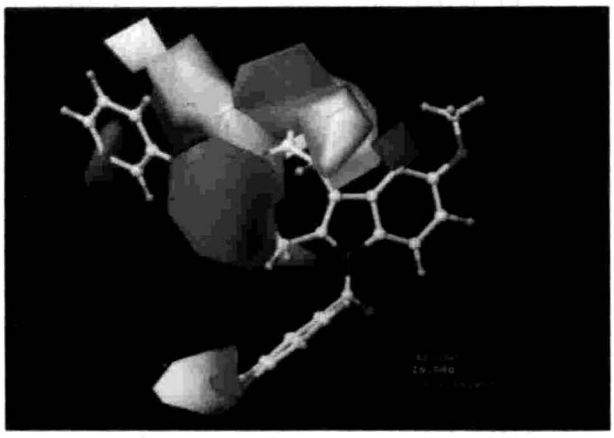


Fig. 5: CoMSIA hydrophobic, hydrogen donor and hydrogen acceptor contour field plots.

Orange contours indicate regions where hydrophobic (lipophillic) groups increase activity, whereas white contours indicate regions where hydrophobic groups decrease activity. Cyan contours indicate regions where hydrogen bond donor groups increase activity, whereas purple contours indicate regions where hydrogen bond donor groups decrease activity. Magenta contours indicate regions where hydrogen bond acceptor groups increase activity, whereas red contours indicate regions where hydrogen acceptors groups decrease activity.

TABLE 3: PLS STATISTICS OF CoMFA AND CoMSIA 3D QSAR

PLS statistics	CoMFA	CoMSIA
Optimum Components	6	4
r^2 cross validated	0.742	0.594
r^2 conventional	0.992	0.935
Std error	0.067	0.185
F value	780.55	144.38
r^2 boot strap[a]	0.993	0.929

[a] Results from 10 run of boot strapped analysis.

The 3D contour plots produced by CoMFA are shown in fig. 2. There is a sterically favorable green contour present at 2nd and 5th position of phenyl ring, and sterically unfavorable yellow contour present in the vicinity of the phenyl ring, suggesting that there is a definite requirement of a substructure with appropriate shape to exhibit high activity. In addition to this, quite a few yellow contours are also spread over the structure indicating that the more steric bulk is unfavorable for COX-2 inhibitory activity. The electrostatic contour (fig. 3) near ester oxygen group, position at which the side chain is attached suggests that positively charged substituents might increase the biological activity. The red region appears around the side chain of the methylene spacer between the phenyl ring and the ester oxygen indicating that high electron density might favour the activity. The most active molecule 11 is shown in the background.

In CoMSIA model, green steric contour present on the side chain and 5th position of the phenyl ring suggests sterically favouable group substituents might increase the biological activity (fig. 4). A big yellow contour spreads on the structure from phenyl to indole, indicating that substitution of bulky group in these regions would decrease activity. The electrostatic fields of CoMSIA such as blue and red contours are similar to that of CoMFA model.

The orange contours around the side chain of the methylene spacer between the phenyl ring and the ester oxygen indicates that hydrophobic groups may be preferred in this region (fig. 5). The red contours show that hydrogen acceptor groups are disfavoured in this region. Magenta coloured contour of CoMSIA represents hydrogen bond donor at meta position of side chain bearing phenyl ring and groups with proton acceptor in these areas might increase activity. Considering all statistical parameter, the CoMFA model performed better than the CoMSIA model. Hence, CoMFA model can be used for prediction of new COX-2 inhibitors with higher biological efficacy.

ACKNOWLEDGEMENTS

The authors thank Director, NIPER, for providing permission to access supercomputer facilities. MAB thanks the Council of scientific and Industrial Research (CSIR), NewDelhi, for a Senior Research Fellowship.

REFERENCES

1. Hia, T. and Nielson, K., **Proc. Natl. Acad. Sci., USA,** 1992, 82, 7384.
2. Jones, D.A., Carlton, D.P., McIntyre, T.M., Zimmerman, G.A. and Prescott, S.M., **J. Biol. Chem.,** 1993, 268, 9049.
3. Kennedy, B., Chan, C.C., Culp, S. and Cromlish, W., **Biochem. Biophys. Res. Commun.**, 1993, 197, 494.
4. Kujubu, D.A., Fletcher, B.S., Varmaum, B.C., Lim, R.M. and Herschman, H., **J. Biol. Chem.**, 1991, 266,12866.
5. Chipmen, J.G. and Erikson, R.L., **Proc. Natl. Acad. Sci., USA,** 1997, 88, 2892.
6. Needleman, P. and Isakson, P.C., **Rheumatology,** 1997, 24, 6.
7. Vane, J.R. and Botting, R.M. **Inflamm. Res.,** 1998, 47, S78.
8. Katori, M., Majima, M. and Harada, Y. **Inflamm. Res.**, 1998, 47, S107.
9. Sawdy, R., Slater, D., Fisk, N., Edmonds, D.K. and Bennet, P., **Lancet**, 1997, 350, 265.
10. Kutchera, W., Jones, D.A., Matsunami, N., Gorden, J., McIntyre, T.M., Zimmerman, G.A., White, R.L. and Prescott, S.M., **Proc. Natl. Acad. Sci., USA,** 1996, 93, 4816.
11. Stewart, W.F., Kawas, C., Corrada, M. and Metter, E.J., **Neurology**, 1997, 48, 626.
12. Dannhardt, G. and Laufer, S., **Current Med. Chem.,** 2000, 7, 1101.
13. Loong, C., Miller, A., Barnett, J., Chow, J., Ramesha, C. and Browner, M.F., **Nat. Struct. Biol.**, 1996, 3, 927.
14. Kalgutkar, A.S., Crews, B.C., Rowlinson, S.W., Garner, C., Seibert, K. and Marnett, L.J., **Science**, 1998, 280, 1268.
15. Kalgutkar, A.S., Kozak, K.R., Crews, B.C., Hochgesang, J.G.P. and Marnett, L.J., **J. Med. Chem.,** 1998, 41, 4800.
16. Bayly, C.I., Black, W.D., Leger, S., Quimet, N., Quellet, M. and Percival, M.D., **Bioorg. Med. Chem. Lett.**, 1999, 9,307.
17. Kalgutkar, A.S., Rowlinson, S.W., Crews, B.C. and Marnett, L.J., **Bioorg. Med. Chem. Lett.**, 2002, 12, 521.
18. Black, W.C., Bayly, C., Belly, M., Chan, C.C., Charleson, S., Denis, D., Gauthier, J.Y., Gordon, R., Guay, D., Kargman, S., Lau, C.K., Leblane, Y., Mancini, J., Quellet, M., Percival, D., Roy, P., Skorey, K., Taggari, P., Vickers, P., Wong, E., Xu, L. and Prasit, P., **Bioorg. Med. Chem. Lett.,** 1996, 6,725.
19. Wood, K.W., Mccroskey, R.W., Michaelider, M.R., Wada, C.K., Hulkower, K.I. and Beil, R.L., **Bioorg. Med. Chem.,** 2002, 11, 1325.
20. Leblanc, Y., Black, W.C., Charleson, C.S., Delrome, D., Denis, D., Gauthier, E.L., Grimm, E.L., Gordon, R., Guay, D., Hamel,

P., Ilargman, S., Lau, C.K., Quellet, M., Pekival, D., Roy, P., Skovey, K., Tagari, R., Vickers, P., Wong, E., Xu, L. and Prasit, P., **Bioorg. Med. Chem.,** 1996, 6,731.
21. Kalgutkar, A.S., Marnett, A.B., Crews, B.C., Remmel, R.P., Marnett, L.J., **J. Med. Chem.**, 2000, 43, 2860.
22. SYBYL 6.8, Molecular Modeling Software, Tripos Associates Inc., St. Louis MO 63144, U.S.A.
23. Gokhale, V.M. and Kulkarni, V.M., **J. Med. Chem.**, 1999, 42, 5348.
24. Waller, C.L., Oprea, T.I., Giolitti, A. and Marshall, G.R., **J. Med. Chem.**, 1993, 36, 4152.
25. Tong, W., Collantes, E.R., Chen, Y. and Welsh, W.J., **J. Med. Chem.**, 1996, 39, 380.
26. Desiraju, G.R., Gopalakrishnan, B., Jetti, R.K.R., Raveendra, D., Sarma, J.A.R.P. and Subramanya, H. S., **Molecules**, 2000, 5, 945.
27. Doytchinova, I. A. and Flower, R. F., **J. Med. Chem.**, 2001, 44, 3572.
28. Nair, A.C., Jayatilleke, P., Wang, X., Miertus, S. and Welsh, W.J., **J. Med. Chem.**, 2002, 45, 973.
29. Schaal, W., Karlsson, A., Ahlsen, G., Lindberg, J., Andersson, H.O., Danielson, U.H., Classon, B., Unge, T., Samuelsson, B., Hulten, J., Hallberg, A. and Karlen, A., **J. Med. Chem.**, 2001, 44, 155.
30. Huang, X., Xu, L., Luo, X., Fan, K., Ji, R., Pei, G., Chen, K. and Jiang, H., **J. Med. Chem.**, 2002, 45, 333.
31. Gasteiger, J. and Marsili, M., **Tetrahedron**, 1980, 36, 3219.
32. Gasteiger, J. and Marsili, M., **Organ. Magn. Reson.**, 1981, 15, 353.
33. Clark, M., Cramer III R.D., Jones, D.M., Patterson, D.E., Simeroth, P.E., **Tetrahedron Comput. Method**, 1980, 3, 47.
34. Cramer III, R.D., Patterson, D.E. and Bunce, J.D., **J. Amer. Chem. Soc.**, 1988, 110, 5959.
35. Klebe, G., Abraham, U. and Mietzner, T., **J. Med. Chem.**, 1994, 37, 4130.