Electronic-Topological Study of Structurally Diverse Cyclooxygenase-2 Inhibitors

A. Dimoglo*, E. Sim, N. Shvets and V. Ahsen

Gebze Institute of Technology, POB 141, Gebze/Kocaeli, 41400, Turkey

Abstract: A large series of cyclooxygenase-2 (COX-2) inhibitors with diverse skeletons were investigated by means of the Electronic-Topological Method. A system for the COX-2 inhibitor activity prognostication was built with 6 pharmacophores and 6 anti-pharmacophores. The forecasting ability of the system was also tested on different structures, which differ from those that characterize the series studied.

Keywords: COX-2, NSAIDs, Structure-Activity Relationships (SAR).

INTRODUCTION

Non-steroidal anti-inflammatory drugs (NSAIDs) have been used for the treatment of arthritis and pain widely, but significant side effects such as gastric and intestinal toxicity and decreased renal function limit their use. The cyclooxygenase (COX) enzymes catalyze the oxidative conversion of arachidonic acid into thromboxane and prostaglandin H_2, which mediates both beneficial and pathological effects. Two forms of COX that have been identified as COX-1 and COX-2 are known. COX-1 is expressed as a constitutive enzyme, and it is involved in homeostasis of the gastrointestinal tract and other functions. It is responsible for the physiological production of prostaglandins, whereas the expression of the COX-2 isoform is induced in response to inflammatory stimuli such as cytokines [1]. Different research groups were involved into the search for COX-2 as selective inhibitor. The search was carried out as follows:

1) Synthesis of new compounds, which are analogues to NSAIDs and widely used in therapy;

2) Pharmacological studies of known and new synthesized potential COX-2 inhibitors;

3) Theoretical studies, such as QSAR, which are based on diverse experimental data. The theoretical investigations are being carried out to optimize the reactivity search for the most active and selective COX-2 inhibitors. It is supposed that QSAR equations might help in understanding the mechanism and action of the compound on the corresponding receptor [2].

Statistical analyses for a series of 4,5-diarylpyrroles were conducted by Copeland *et al.* [3] to describe and predict activities of these compounds *in vivo* and *in vitro*. Empirical parameters comparison for different QSAR models was carried out for pyrrole derivatives [2] to reveal their best activities and anti-inflammatory reagents. The research on COX-2 inhibitors was focused on inflammation and pain. However, experimental and epidemiological data suggest that COX-2 inhibitors could be used in the treatment or prevention of a broader range of diseases. A review [4] on some unresolved issues related to the discovery of COX-2 inhibitors presents the kinetic and structural basis for their selectivity and possible complications in their development and use. It is an example of putting structural information to use in a COX-2 inhibitor design. Also Zhu *et al.* [5] give some structural models of NSAIDs, which are expected to possess a high level of inhibitory activity and selectivity. The DycoBlock method used in the article is based on the multiple-copy stochastic dynamics simulation in the presence of a receptor molecule. Acantoic acid is reported as a novel COX-2 inhibitor and its SAR is studied by docking [6]. As a result, natural pimarane diterpene and a series of its analogues have been discovered as novel COX-2 inhibitors. Later on a series of 4-aryl/cycloalkyl-5-phenyl oxazole derivatives synthesis and SAR-study was reported [7]. It was shown that they are potent and selective COX-2 inhibitors. Their SAR study resulted in a few rules considered as being useful for the COX-2 inhibitor activity prediction. One more approach, Monte Carlo extended linear response calculations have been used for prediction of binding affinities of celecoxib analogues with the COX-2 enzyme [8]. From all these works it follows that either their authors are dealing with only one structural class, or they compare their results with a definite structural foretype of a NSAID. These limitations, however, can be bypassed by using the electronic-topological method (**ETM**) [9].

MATERIALS AND METHOD

Data Sets

From multiple publications on NSAIDs, four structural classes were selected for the SAR study in the frameworks of the **ETM**. They are benzopyrane derivatives [10], sulfonyl-

*Address correspondence to this author at the Gebze Institute of Technology, PK 141, Gebze/Kocaeli, 41400, Turkey; Tel: +90 262 653 84 97/1428; Fax: +90 262 653 84 90; E-mail: dimoglo@penta.gyte.edu.tr

1389-5575/03 $41.00+.00

containing terphenyls [11], diaryl derivatives of pyrazole [12], and diaryl-substituted pyridines [13] (see Fig. **1**).

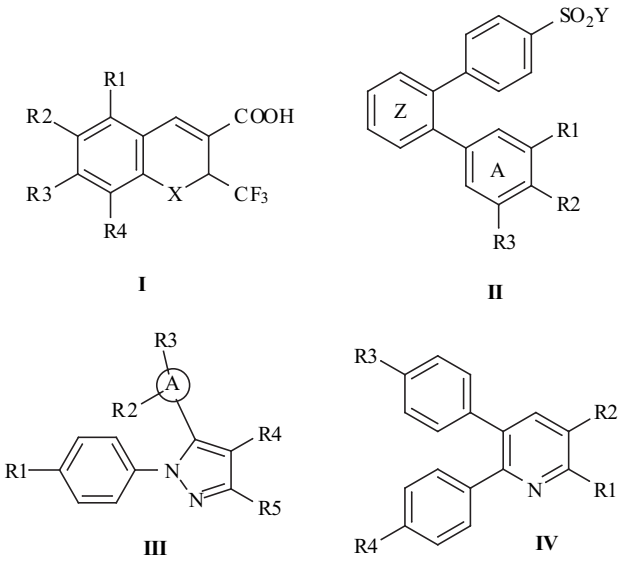


Fig. (1). Common molecular skeletons for the compounds of the training set.

There are two reasons for this selection. First, these compounds belong to different structural classes; next, related experimental data are obtained by the same research group [11-16]. The latter circumstance is considered to be very important for the correctness of biological parameters' comparison. At the same time, the search for common pharmacophores in the structures of active compounds is more informative when they possess diverse skeletons rather than similar ones. The total number of compounds investigated at the training phase is 192. All they are divided into classes of active compounds (118 molecules with $IC_{50} \leq 0.1$ μM) and inactive compounds (74 molecules with $IC_{50} \geq 2.0$ μM, see Tables **1-4**) relative to the IC_{50} index of activity (that is, substance concentration sufficient for inhibiting 50% of the enzyme).

Compounds from series **I** are characterized by the presence of condensed benzopyrane cycles. Most of the compounds have no sulfonyl group. All they are carbonic acids because there is a carboxyl group attached in position 3. The distinguishing feature of the compounds from series **II** is that the latter have no heterocycles (excluding compounds **81, 85, 87**). Phenyl rings with various substituents form the basis of their skeletons. All compounds in this series contain a sulfonyl group. Compounds from series **III** are characterized by the presence of sulfonamide group and pyrazole cycle with trifluoromethyl group attached to it. The cycle A in this series is phenyl, as a rule. Series **IV** is formed as pyridine derivatives. It contains three-cyclic structures with phenyl rings vicinal attachment to the central heterocycle, similar to the previous three series. Normally, there is a methyl-sulfonyl group in *para*-position in one of the phenyl rings. From the point of view of pharmacology, compounds from series **III** are of interest because some preparations that are used in clinical practice are in the list. In particular, compound **118** is well known and well documented NSAID and named as celecoxib [13]. In series **II** there are compounds possessing high activity and selectivity. Compounds of series **I** and series **IV** have approximately the same level of activity of celecoxib. Their selectivity, however, is a few times higher than the selectivity of this known drug.

Table 1. Compounds of Series I and their IC_{50} Indices

N comp	R1	R2	R3	R4	X	IC_{50}, μM	Activity**	
							exper.	calcul.
1	H	CH_3S	H	H	O	<0.1	+	+
2	H	Cl	CH_3	H	O	<0.1	+	+
3	H	Cl	*t*-Butyl	H	O	<0.1	+	+
4	H	3-Cl-4-CH_3O-PhO	H	H	O	<0.1	+	–
5	H	CF_3O	H	H	O	<0.1	+	+
6	H	Cl	H	Et	O	<0.1	+	+
7	H	Cl	Ph	H	O	<0.1	+	+
8	H	Cl	Cl	H	O	<0.1	+	+
9	H	CH_3O	H	Cl	O	<0.1	+	+
10	H	Cl	H	H	S	<0.1	+	+
11	H	CF_3O	H	H	O	<0.1	+	+
12	H	HCF_2	H	H	O	<0.1	+	+
13	H	CF_3	H	H	O	<0.1	+	+

(Table 1). contd.....

N comp	R1	R2	R3	R4	X	IC_{50}, μM	Activity**	
							exper.	calcul.
14	Cl	Cl	Cl	H	O	<0.1	+	+
15	H	CF_3CF_2	H	H	O	<0.1	+	+
16	H	SCF_3	H	H	O	<0.1	+	+
17	H	4-CF_3-PhO	H	H	O	<0.1	+	+
18	H	4-Cl-PhO	H	H	O	<0.1	+	+
19	H	Cl	H	Ethenyl	O	<0.1	+	+
20	H	Cl	H	Ethynyl	O	<0.1	+	+
21	H	Cl	H	2-Thienyl	O	<0.1	+	+
22	H	Cl	H	Ph-Ethynyl	O	<0.1	+	+
23	H	Cl	H	4-Cl-Ph	O	<0.1	+	+
24	H	Cl	CH_3	H	S	<0.1	+	+
25	H	SCF_3	H	H	S	<0.1	+	+
26	H	Cl	H	Cl	S	<0.1	+	+
27	H	4-CH_3O-PhO	H	H	O	0.1	+	+
28	H	Cl	Cl	Cl	O	0.1	+	+
29	H	4-OH-Benzoyl	H	H	O	2.0	–	–
30	H	NH_2	H	H	O	2.2	–	–
31	H	H	H	*t*-Butyl	O	7.0	–	–
32	H	*t*-Butyl	H	*t*-Butyl	O	7.0	–	–
33	H	4-Morpholino-SO_2	H	H	O	8.0	–	–
34	H	H	H	Ph	O	11.0	–	–
35	Cl	H	Cl	H	O	12.0	–	–
36	H	CH_3SO_2	H	H	O	12.0	–	–
37	H	[(*t*-Butyl)-NH]SO_2	H	H	O	19.0	–	–
38	H	[(CH_3)-NH]SO_2	H	H	O	20.0	–	–
39	H	[($PhCH_2$)-NH]SO_2	H	Cl	O	21.0	–	–
40	H	[($PhCH_2$)-NH]SO_2	H	H	O	24.0	–	–
41	H	Cl	H	*i*-Pr	O	36.0	–	+
42	H	Ph-acetyl	H	H	O	43.0	–	–
43	H	(*i*-Pr)-$NHSO_2$	H	H	O	46.0	–	–
44	H	Cl	Et	H	O	67.0	–	–
45	H	$(CH_3)_2NSO_2$	H	H	O	79.0	–	–
46*	H	Cl	H	H	O	82.0	–	–
47	H	H	H	Cl	O	>100	–	–
48	H	H	CH_3	CH_3	O	>100	–	–
49	H	H	*i*-Pr	H	O	>100	–	–
50	H	H	Ph	H	O	>100	–	–

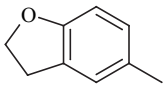
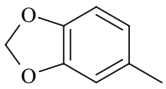
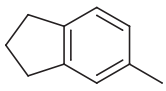
*(S)-isomer

**- "+" for active compounds, "–" for inactive ones.

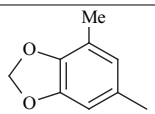
Table 2. Compounds of Series II and their IC_{50} Indices

N comp	Z	Y	A	R1	R2	R3	IC_{50}, μM	Activity	
								exper.	calcul.
51	4,5-F_2-Ph	NH_2	Ph	H	F	Cl	0.002	+	+
52	4,5-F_2-Ph	NH_2	Ph	H	F	CH_3	0.002	+	+
53	4,5-F_2-Ph	NH_2	Ph	H	CH_3	Cl	0.003	+	+
54	4,5-F_2-Ph	NH_2	Ph	H	Cl	CH_3	0.003	+	+
55	4,5-F_2-Ph	NH_2	Ph	H	F	H	0.004	+	+
56	4,5-F_2-Ph	NH_2	Ph	H	CH_3	H	0.004	+	+
57	4,5-F_2-Ph	NH_2	Ph	H	-OCH_2O-		0.004	+	+
58	4,5-F_2-Ph	CH_3	Ph	H	F	CH_3	0.005	+	+
59	4,5-F_2-Ph	NH_2	Ph	H	CH_3O	CH_3	0.005	+	+
60	4,5-F_2-Ph	NH_2	Ph	H	CH_3	CH_3	0.005	+	+
61	Ph	NH_2	Ph	H	Cl	H	0.006	+	+
62	4,5-F_2-Ph	CH_3	Ph	H	Cl	CH_3	0.006	+	+
63	4,5-F_2-Ph	NH_2	Ph	H	$(CH_3)_2N$	Cl	0.006	+	+
64	4,5-F_2-Ph	CH_3	Ph	H	CH_3	H	0.007	+	+
65	4,5-F_2-Ph	CH_3	Ph	H	$(CH_3)_2N$	Cl	0.008	+	+
66	4,5-F_2-Ph	CH_3	Ph	H	F	Cl	0.010	+	+
67	4,5-F_2-Ph	CH_3	Ph	H	-OCH_2O-		0.012	+	+
68	4,5-F_2-Ph	NH_2	Ph	H	CH_3O	F	0.013	+	+
69	4,5-F_2-Ph	NH_2	Ph	H	CH_3O	Cl	0.013	+	+
70	4,5-F_2-Ph	CH_3	Ph	H	CH_3O	CH_3	0.013	+	+
71	4,5-F_2-Ph	CH_3	Ph	H	CH_3	Cl	0.013	+	+
72	4,5-F_2-Ph	CH_3	Ph	H	F	H	0.014	+	+
73	Ph	NH_2	Ph	Cl	F	H	0.017	+	+
74	Ph	NH_2	Ph	Cl	CH_3O	H	0.019	+	+
75	4,5-F_2-Ph	CH_3	Ph	H	CH_3O	Cl	0.019	+	+
76	4,5-F_2-Ph	CH_3	Ph	H	CH_3O	F	0.021	+	+
77	4,5-F_2-Ph	NH_2	Ph	Cl	CH_3O	Cl	0.021	+	+
78	4,5-F_2-Ph	CH_3	Ph	H	CH_3	CH_3	0.023	+	+
79	4,5-F_2-Ph	NH_2	Ph	H	-$O(CH_2)_2O$-		0.032	+	+
80	Ph	NH_2	Ph	F	CH_3O	H	0.033	+	+
81	4,5-F_2-Ph	NH_2	3-Pyridinyl	-	CH_3	H	0.051	+	+
82	Ph	CH_3	Ph	F	CH_3O	H	>100	–	–
83	1,2,3,4-F_4-Ph	CH_3	Ph	H	F	H	>100	–	–
84	Naphthalene	CH_3	Ph	H	F	H	>100	–	–
85	2,3-Pyrazine	CH_3	Ph	H	F	H	>100	–	–
86	4,5-F_2-Ph	CH_3	Ph	Cl	CH_3O	Cl	>100	–	–
87	4,5-F_2-Ph	CH_3	2-Pyridinyl	H	CH_3	H	52.3	–	–
88	Ph	CH_3	Ph	Cl	CH_3O	H	11.4	–	–

Table 3. Compounds of Series III and their IC_{50} Indices

N comp	A	R1	R2	R3	R4	R5	IC_{50},μM	Activity	
								exper.	calcul.
89	Ph	SO_2NH_2	H	H	F	CF_3	0.0017	+	+
90	Ph	SO_2NH_2	3-CH_3	4-SCH_3	H	CF_3	0.0037	+	+
91	Ph	SO_2NH_2	H	4-$N(CH_3)_2$	H	CF_3	0.0047	+	+
92	Ph	SO_2NH_2	H	4-Cl	Cl	CF_3	0.0053	+	+
93	Ph	SO_2NH_2	3-F	4-$N(CH_3)_2$	H	CF_3	0.0057	+	+
94	Ph	OCH_3	H	4-SO_2NH_2	H	CF_3	0.0067	+	+
95	Ph	SO_2NH_2	H	4-OCH_3	H	CF_3	0.008	+	+
96	Ph	SO_2NH_2	H	4-SCH_3	H	CF_3	0.009	+	+
97	Ph	SO_2NH_2	3-CH_3	4-OCH_3	H	CF_3	0.0093	+	+
98	Ph	SO_2NH_2	H	4-Cl	H	CF_3	0.01	+	+
99	Ph	SO_2NH_2	H	4-Cl	H	CHF_2	0.01	+	+
100	Ph	SO_2NH_2	H	4-Cl	Cl	CN	0.01	+	+
101	Ph	F	H	4-SO_2NH_2	H	CF_3	0.01	+	+
102	Ph	SO_2NH_2	H	4-CH_3	H	CHF_2	0.013	+	+
103	Ph	SO_2NH_2	H	4-OCH_3	H	CHF_2	0.015	+	+
104	Ph	SO_2NH_2	3-Cl	4-Cl	H	CF_3	0.015	+	+
105	Ph	SO_2NH_2	H	4-$NHCH_3$	H	CF_3	0.016	+	+
106	Ph	SO_2NH_2	3,5-Cl_2	4-OCH_3	H	CF_3	0.021	+	+
107		SO_2NH_2	H	H	H	CF_3	0.021	+	+
108	Ph	SO_2NH_2	H	4-Cl	CH_3	CF_3	0.022	+	+
109		SO_2NH_2	H	H	H	CF_3	0.024	+	+
110	5-Cl-2-Thienyl	SO_2NH_2	H	H	H	CF_3	0.026	+	+
111	Ph	SO_2NH_2	3-Cl	4-OCH_3	H	CHF_2	0.027	+	+
112	Ph	SO_2NH_2	3-Cl	4-$NHCH_3$	H	CF_3	0.027	+	+
113	Ph	SO_2NH_2	H	4-Cl	Et	CF_3	0.028	+	+
114	Ph	SO_2NH_2	H	H	Cl	CH_3	0.028	+	+
115	Ph	SO_2NH_2	H	4-Cl	H	CH_2OCH_2Ph	0.029	+	+
116		SO_2NH_2	H	H	H	CF_3	0.031	+	–
117	Ph	SO_2NH_2	H	H	H	CF_3	0.032	+	+
118	Ph	SO_2NH_2	H	4-F	H	CF_3	0.040	+	+
119	Ph	SO_2NH_2	2-Cl	H	H	CF_3	0.041	+	+

(Table 3). contd.....

								Activity	
N comp	**A**	**R1**	**R2**	**R3**	**R4**	**R5**	**IC_{50},μM**	**exper.**	**calcul.**
120	Ph	SO_2NH_2	H	H	Cl	H	0.049	+	–
121	Ph	SO_2NH_2	3-F	4-OCH_3	H	CHF_2	0.05	+	+
122		SO_2NH_2	H	H	H	CF_3	0.052	+	+
123	5-Cl-2-Thienyl	SO_2NH_2	H	4-Cl	H	CN	0.052	+	+
124	Ph	SO_2NH_2	2-Cl	H	H	CF_3	0.056	+	+
125	Ph	SO_2NH_2	H	4-F	H	CONH(3-ClPh)	0.056	+	+
126	Ph	SO_2NH_2	2-Cl	4-Cl	H	CF_3	0.056	+	+
127	Ph	SO_2NH_2	2-F	H	H	CF_3	0.058	+	+
128	Ph	SO_2NH_2	3-Cl	4-OCH_3-5-CH_3	H	CF_3	0.066	+	+
129	Ph	SO_2NH_2	2-CH_3	H	H	CF_3	0.069	+	+
130	Ph	SO_2NH_2	H	4-CH_3	CN	H	0.076	+	–
131	Ph	SO_2NH_2	H	H	OCH_3	CF_3	0.08	+	+
132	1-Cyclohexyl	SO_2NH_2	H	H	H	CF_3	0.084	+	+
133	Ph	SO_2NH_2	H	4-Cl	H	4-CH_3OPh	0.10	+	+
134	Ph	SO_2CH_3	H	4-F	H	CF_3	0.10	+	+
135	Ph	SO_2NH_2	H	4- SO_2CH_3	H	CHF_2	>100	–	–
136	Ph	SO_2NH_2	H	4-$CONH_2$	H	CHF_2	>100	–	–
137	Ph	SO_2NH_2	H	4-OH	H	CF_3	>100	–	–
138	Ph	SO_2NH_2	2,5-Cl_2	H	H	CF_3	>100	–	–
139	Ph	SO_2NH_2	2,5-$(CH_3)_2$	H	H	CHF_2	>100	–	–
140	Ph	SO_2NH_2	H	4-F	H	H	>100	–	–
141	Ph	SO_2NH_2	H	4-F	H	CO_2H	>100	–	–
142	Ph	SO_2NH_2	H	4-F	H	CO_2CH_3	>100	–	–
143	Ph	SO_2NH_2	H	4-F	H	$CONH_2$	>100	–	–
144	Ph	SO_2NH_2	H	H	H	OCH_3	>100	–	–
145	Ph	SO_2NH_2	H	4-F	*n*-Pr	CF_3	>100	–	–
146	Ph	$NHSO_2CH_3$	H	4-F	H	CF_3	>100	–	–
147	Ph	NO_2	H	4-F	H	CF_3	>100	–	–
148	Ph	$COCF_3$	H	4-F	H	CF_3	>100	–	–
149	Ph	H	H	H	H	CF_3	>100	–	+
150	Ph	SO_2NHCH_3	H	4-F	H	CF_3	>100	–	–
151	Ph	$SO_2N(CH_3)_2$	H	4-F	H	CF_3	<100	–	–
152	Ph	SO_2NH_2	H	4-CH_2OH	H	CF_3	93.3	–	–
153	Ph	OCH_3	H	4-Cl	H	CF_3	74.9	–	–
154	Ph	SO_2NH_2	H	H	NO_2	H	70.0	–	–

(Table 3). contd.....

N comp	A	R1	R2	R3	R4	R5	IC_{50},μM	Activity	
								exper.	calcul.
155	4-Pyridyl	SO_2NH_2	H	H	H	CF_3	64.7	–	–
156	Ph	SO_2NH_2	H	H	H	CH_3	62.8	–	–
157	Ph	SO_2NH_2	H	H	CH_3	H	47.1	–	–
158	Ph	SO_2NH_2	H	4-CO_2H	H	CHF_2	46.8	–	–
159	2-Pyridyl	SO_2NH_2	H	H	H	CF_3	45.6	–	–
160	3-Pyridyl	SO_2NH_2	H	H	H	CF_3	45.0	–	–
161	Ph	SO_2NH_2	H	4-CN	H	CHF_2	29.7	–	–
162	Ph	SO_2NH_2	H	H	NH_2	H	29.7	–	–
163	Ph	SO_2NH_2	H	Cl	SO_2CH_3	H	19.8	–	–
164	Ph	SO_2NH_2	2-$N(CH_3)_2$	H	H	CF_3	14.3	–	–
165	Ph	SO_2NH_2	H	4-CO_2H	H	CF_3	11.2	–	–
166	Ph	SO_2NH_2	H	4-CF_3	H	CF_3	8.23	–	–
167	Ph	SO_2NH_2	3-F	H	H	CF_3	7.73	–	–
168	Ph	Cl	H	4-Cl	H	CF_3	4.79	–	–
169	Ph	SO_2NH_2	H	H	F	H	4.66	–	+
170	Ph	SO_2NH_2	H	H	OH	CF_3	3.58	–	–
171	1-CH_3-2-Furyl	SO_2NH_2	H	H	H	CHF_2	3.29	–	–
172	Ph	SO_2NH_2	H	4-NO_2	H	CF_3	2.63	–	–

As the testing set, 29 molecules were taken (Table 6). It was interesting to test the predictive ability of the ETM on compounds possessing other common skeletons than those taken for the examining set.

In the general case, the presence in a compound of one or more **P**s from the 6 **P**s calculated allows for declaring the compound as active one, under condition that the compound does not contain **AP**s at all (the similar case are **AP**s for inactive compounds). At the same time, when both **P**s and **AP**s are simultaneously found in a compound, the latter can be qualified as belonging to an intermediate (buffer) class, where compounds can possess a spectrum of activities (including low-active and middle-active compounds).

METHOD

The **ETM** [9] was proposed in 1985 and since that time had been under permanent improvement [17-20]. Since details of the **ETM** can be found in literature [21, 22], we give here only those properties of the **ETM** that differentiate it from other methods used in the structure-activity relationships (SAR) studies. The most important property of the **ETM** is its specific language for chemical structures description. Each compound is described in the form of the so-called Electronic-Topological Matrix of Contiguity (**ETMC**). Any such matrix is being formed of values taken from the results of the corresponding structure optimisation and quantum-chemical calculations for the structure. When the data mentioned are obtained already, main steps of the **ETM**-study are as follows:

1. From the calculated data select and fix one desirable electron characteristic for atoms (e.g. atomic charge) and one more for bonds (here, the Wiberg's indices). These two characteristics determine the meaning of *diagonal* and some *off-diagonal* (i.e. of those representing chemically bonded pairs of atoms) elements, correspondingly.

2. Form an **ETMC** for every molecule by choosing corresponding values of the characteristics. If no bond, use the corresponding three-dimensional (3D) distance.

3. Set a threshold value for the activity in view (to partition all compounds into classes of active and inactive molecules); set a desirable level of the activity prediction, and some precision values for comparing an **ETMC** formed for a compound from the series under study with the fixed **ETMC** ("a template"); as a rule, it is one of the most active compounds.

4. From the comparison of all structures with the template, find structural fragments $\mathbf{S_i}$ ($i \in I$) that are

Table 4. Compounds of Series IV and their IC_{50} Indices

N comp	R1	R2	R3	R4	IC_{50},μM	Activity	
						exper.	calcul.
173	Ph	CN	SO_2CH_3	F	<0.1	+	+
174	SCH_3	H	SO_2CH_3	SCH_3	<0.1	+	+
175	O-*n*-Butyl	CN	SO_2CH_3	F	<0.1	+	+
176	OCH_3Ph	CN	SO_2CH_3	F	<0.1	+	+
177	OCH_2CHCH_2	CN	SO_2CH_3	F	<0.1	+	+
178	O-*n*-Pentyl	CN	SO_2CH_3	F	<0.1	+	+
179	O-*n*-Butyl	CN	SO_2NH_2	F	<0.1	+	+
180	O-*n*-Butyl	CH_2OH	SO_2CH_3	F	<0.1	+	+
181	O-*n*-Butyl	CHF_2	SO_2CH_3	F	<0.1	+	+
182	Cl	CN	SO_2CH_3	F	0.1	+	+
183	$OCH_2CH=C(CH_3)_2$	CN	SO_2CH_3	F	0.1	+	–
184	O-*n*-Butyl	CH_2OCOCH_3	SO_2CH_3	F	0.1	+	+
185	O-*n*-Butyl	$CH(CH_3)OH$	SO_2CH_3	F	0.1	+	+
186	H	$CONH_2$	SO_2CH_3	F	105	–	–
187	O-*n*-Butyl	CH_2OCO-*i*-Pr	SO_2CH_3	F	>100	–	–
188	Cl	H	SO_2CH_3	F	29	–	–
189	Cl	$CONH_2$	SO_2CH_3	F	26.2	–	–
190	O-*n*-Butyl	CH_2NHCH_3	SO_2CH_3	F	3.0	–	–
191	$NHCH_3$	CN	SO_2CH_3	F	2.3	–	–
192	N_3	H	SO_2CH_3	F	2.0	–	–

common for all active compounds but not present in inactive ones.

5. Estimate the found fragments ("activity features") in accordance to a probabilistic criterion (P_a). Choose those that correspond to the desired level of prediction set before calculations. If the fragments are not informative enough, change some initial settings (or all of them, according to the step 3) and repeat steps 4-5. A common scheme of the **ETM** is shown in Fig. (**2**).

So, the aim of the core procedure of the ETM is to find those fragments that are common for all active compounds and satisfy some initial conditions (precision values and probability of their occurrence in active compounds). A criterion that is commonly used in structural methods to evaluate the probability of a fixed structural feature $\mathbf{S_k}$ occurrence in active compounds from a given series is being calculated using the following equation:

$$P_a = (n_1+1)/(n_1+n_2+2).$$

Here n_1 is the number of molecules possessing the feature of activity $\mathbf{S_k}$ in the class of active compounds, and n_2 has the same meaning but in the class of inactive compounds (both obtained from the **ETM** calculations).

When successfully found for a training selection only, the activity features $\mathbf{S_i}$, $i \in I$, can already be used to predict the activity of interest in a new series of compounds with probability P_a estimated for each structural feature at the last step of the **ETM** session.

When subsequently proved for an examining set, they allow for estimating the stability of the prognostication system.

RESULTS AND DISCUSSION

Calculating Pharmacophores and Anti-Pharmacophores

Conformational analysis for all compounds was done by means of a molecular mechanics method (MMX) [23]. Their electronic structures were calculated by semi-empirical methods AM1 [24] and CNDO/2 [25]. A set of pharmacophores was calculated (relative to the series under study) by the **ETM.** Besides the pharmacophores, those molecular fragments that can be found in inactive

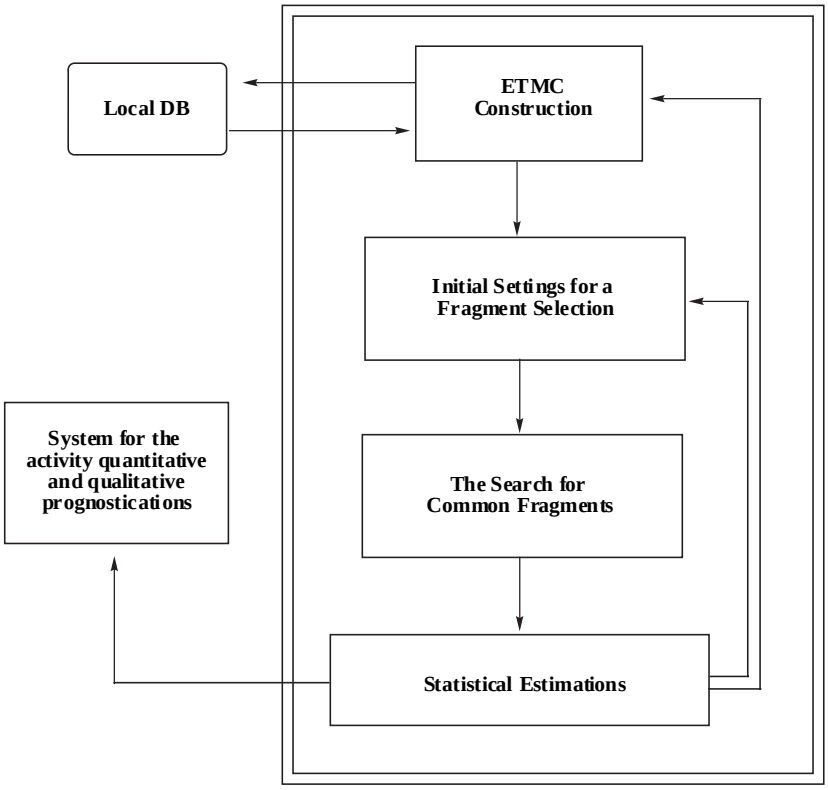


Fig. (2). The framework of the **ETM.**

compounds only and, consequently, cause the activity loss (anti-pharmacophores) were revealed as well. Both pharmacophores and anti-pharmacophores formed the basis of a system for the COX-2 inhibitor activity prediction. In Table **5**, statistical characteristics of six pharmacophores (**P**) and six anti-pharmacophores (**AP**) that form the system are given.

Table 5. Statistical Characteristics of Pharmacophores (P) and Anti-Pharmacophores (AP)

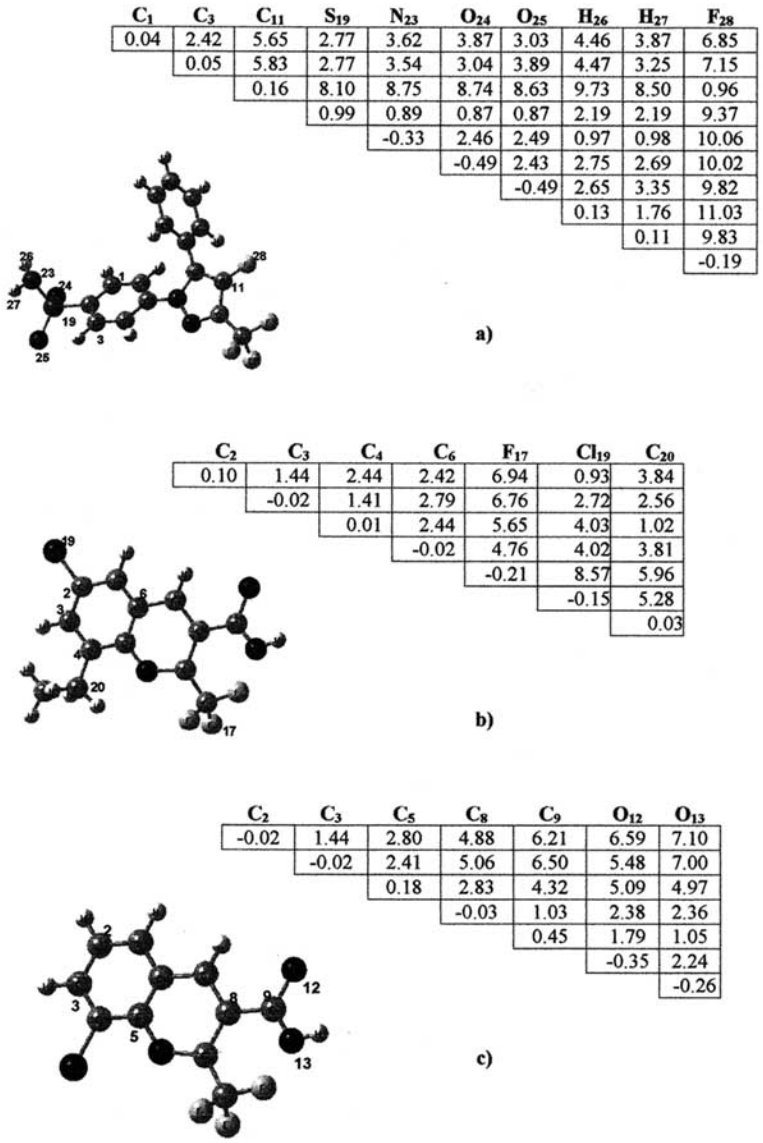
Type of pharmacophore/ antipharmacophore	n_1	n_2	P_a	P_{in}*
P1	31	1	0.94	0.06
P2	31	2	0.94	0.06
P3	29	1	0.94	0.06
P4	34	3	0.90	0.10
P5	26	0	0.96	0.04
P6	25	0	0.96	0.04
AP1	3	16	0.19	0.81
AP2	2	15	0.16	0.84
AP3	2	15	0.16	0.84
AP4	0	11	0.08	0.92
AP5	1	13	0.12	0.88
AP6	0	5	0.14	0.86

*** probability estimates for inactive compounds**

All **P**s and **AP**s taken together define the best partitioning of the compounds taken for the study into classes of active and inactive compounds. Most of the active compounds (~80%) contain simultaneously by four **P**s of the 6 **P**s found; up to two **P**s are realized in 15% of the active compounds; all six **P**s are found in 5% of the compounds. To form the system for the COX-2 inhibitor activity prediction, compounds representing each series and possessing high-level inhibitory activity and selectivity were tried as templates for the **ETMC**s comparison. In Fig. (**3a-3c**), skeletons of the template compounds and *submatrices* of their **ETMC**s (**ETSC**s, for short) are shown as an example; the **ETSC**s given correspond to some of calculated **P**s and **AP**s.

So, pharmacophore **P1** (see Fig. (**3a**)) consists of 10 atoms (compound **89** was taken as the template). All atoms of sulfonylamide group, atom C_3 of the phenyl ring, atom C_{11} of pyrazole and attached to pyrazole F_{28} atom enter the **P1**. This fragment reveals itself in 31 active compounds and in 1 inactive compound. Thus, the probability of the **P1** realization (P_a) in the compounds of the training set is equal to 0.94. The maximal distance for atoms described by the **P1**'s submatrix is observed for the H_{26} atom of amino group and F_{28} (11.03Å). Four atoms of **P1** have high negative charges ($Q \approx -0.19e^- - -0.49e^-$) while the rest of atoms are charged positively. The maximal positive charge ($Q=0.99e^-$) is on the sulfur atom. Dihedral angle between phenyl ring with sulfonylamide group attached and pyrazole ring was found as 54º.

The **P3** pharmacophore found from the template compound **6** is given in Fig. (**3b**). In this substructure, 4 of



C_1	C_3	C_{11}	S_{19}	N_{23}	O_{24}	O_{25}	H_{26}	H_{27}	F_{28}
0.04	2.42	5.65	2.77	3.62	3.87	3.03	4.46	3.87	6.85
	0.05	5.83	2.77	3.54	3.04	3.89	4.47	3.25	7.15
		0.16	8.10	8.75	8.74	8.63	9.73	8.50	0.96
			0.99	0.89	0.87	0.87	2.19	2.19	9.37
				-0.33	2.46	2.49	0.97	0.98	10.06
					-0.49	2.43	2.75	2.69	10.02
						-0.49	2.65	3.35	9.82
							0.13	1.76	11.03
								0.11	9.83
									-0.19

a)

C_2	C_3	C_4	C_6	F_{17}	Cl_{19}	C_{20}
0.10	1.44	2.44	2.42	6.94	0.93	3.84
	-0.02	1.41	2.79	6.76	2.72	2.56
		0.01	2.44	5.65	4.03	1.02
			-0.02	4.76	4.02	3.81
				-0.21	8.57	5.96
					-0.15	5.28
						0.03

b)

C_2	C_3	C_5	C_8	C_9	O_{12}	O_{13}
-0.02	1.44	2.80	4.88	6.21	6.59	7.10
	-0.02	2.41	5.06	6.50	5.48	7.00
		0.18	2.83	4.32	5.09	4.97
			-0.03	1.03	2.38	2.36
				0.45	1.79	1.05
					-0.35	2.24
						-0.26

c)

Fig. (3). (a-c). ETSCs of pharmacophores **P1, P3** and anti-pharmacophore **AP1** (correspond to template compounds **89** (**a**), **6** (**b**), and **47** (**c**)).

its 7 atoms belong to a phenyl ring. Atoms Cl_{19} and C_{20} are also attached to the phenyl ring. Atom F_{17} enters the CF_3 group. Although **P3** contains neither sulfonylamide group nor methylsulfonyl group, it enters 29 active compounds. In other words, the developed system on the COX-2 inhibitor activity prognostication does not limit the search of compounds possessing anti-inflammatory activity by sulfonyl-containing structures. Comparing to **P1**, **P3** is less extended. The maximal distance between two negatively charged atoms Cl_{19} and F_{17} is 8.57Å. Atoms of carbon have small (positive and negative) charges. The exception is the C_2 atom; its charge (Q) is $0.10e^-$. Another example of an **ETSC** that corresponds to **AP1** is shown in Fig. (**3c**). The given anti-pharmacophore was found from the template compound **47** that belongs to series **I**. **AP1** consists of 7 atoms. Three of these atoms enter the carboxyl group, while the rest four atoms belong to the benzopyrane moiety. This anti-pharmacophore is realized in 20 inactive compounds, which possess diverse molecular skeletons (series **I-IV**). In spite of apparent similarity of the template structures **6** and **47** that have caused the **P3** and **AP1** identification, their **ETMC**s differ considerable. **AP1** is distinct from **P3** in that it includes carboxyl group with high negative charges on atoms of oxygen ($Q \approx -0.26e^- - -0.35e^-$). Two carbon atoms, C_5 and C_9, are charged positively (Q is $0.18e^-$ and $0.45e^-$, correspondingly). **P2, P5** and **AP2-AP6** realization is shown for template compounds **52**, **139** and **146** in Fig. (**4**).

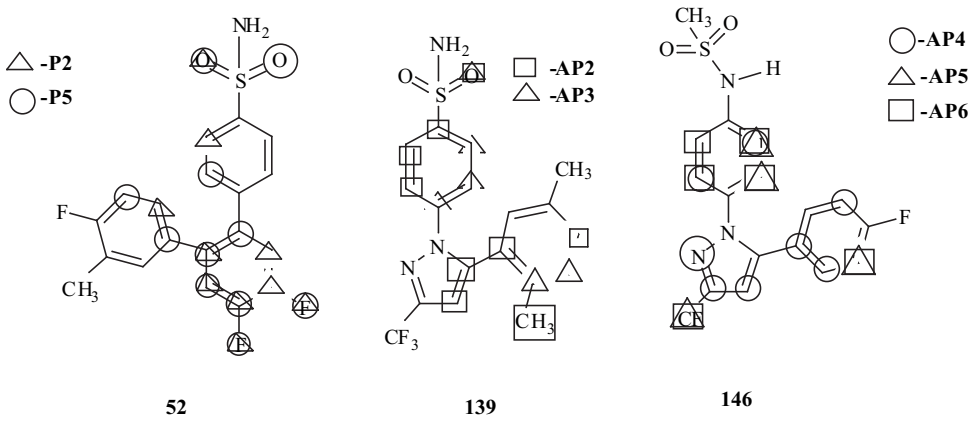


Fig. (4). The found pharmacophores **(P2, P5)** and anti-pharmacophores (**AP2-AP6)** realization in the compounds **52, 139** and **146.**

The highest occupied molecular orbitals (HOMO) and the lowest unoccupied molecular orbitals (LUMO), called also *frontier orbitals*, may well play an important role in the donor-acceptor interaction of a substance with the corresponding receptors. Analysis of HOMOs for the compounds containing **P1, P3** and **AP1** has shown that set (29 compounds in all). To demonstrate the ability dramatically, this set was formed of three subsets, **T1, T2,** and **T3**, which correspond to three different base skeletons (see Fig. (**6**) and Table **6**) [12-14]. The peculiarity of the examining set is that all three skeletons differ from those from training set (series **I**-**IV**). The sub-series **T1, T2, T3**

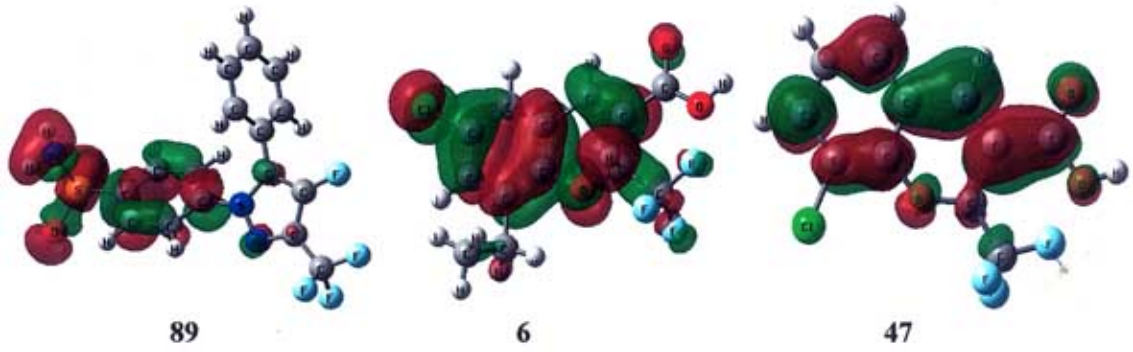


Fig. (5). A three-dimensional view of HOMO orbitals for template compounds 89, 6 and 47.

atoms with the highest values of the atomic orbital coefficients are mainly those atoms that enter into the fragments. Graphical representation of the HOMO orbitals is given in Fig. (**5**).

HOMO orbital for the template compound **89** Fig.(**5**) consists of orbitals of those atoms that form aminosulfonyl group and, partially, phenyl and pyrazole rings. The most part of atoms in **P1** are exactly those that deposit considerably to the HOMO orbitals. Similar situation can be observed in the case of template compounds **6** and **47**. As an example, HOMO orbital of compound **6** includes all atoms entering the **P3** pharmacophore. In contrast to **P3**, HOMO orbital of **AP1** consists of all atoms of carboxyl group and carbon atoms of phenyl ring. All the said suggests again an important role of these atoms in the substrate-receptor interaction.

Testing the Prognostication System

The forecasting ability of the system developed on the base of the **ETM** calculations was tested on an examining

Fig. (**6**) include oxazole, imidazole and pyrrole derivatives, respectively. By varying substituents in their skeletons, one can get a diverse set of chemical substances. In Table **6** the results of testing are given, where tested compounds are qualified as active (+) or inactive ones (–). The most of the compounds (i.e. 14 of 15 active compounds and 12 of 14 inactive ones) are identified correctly. Thus, there is an acceptable correlation with experimental data. Statistical

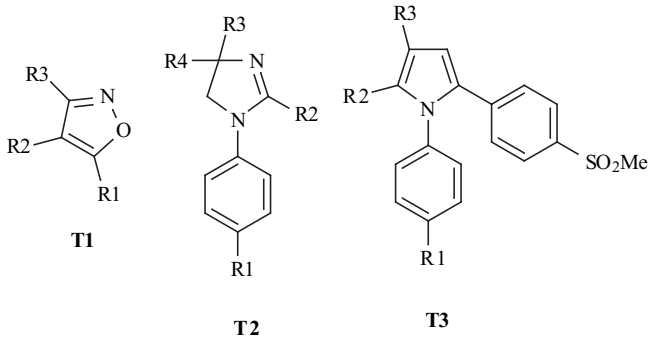


Fig. (6). Common skeletons of the tested compounds.

Table 6. Compounds of the Examining Set and their Activity

Comp N	Series N	R1	R2	R3	R4	IC_{50}, μM	Activity	
							exper.	calcul.
193	T1	CH_3	4-SO_2NH_2Ph	Ph	-	<0.1	+	+
194	T1	CH_2-*t*-Butyl	4-SO_2NH_2Ph	Ph	-	<0.1	+	+
195	T1	CH_2OCH_3	4-SO_2NH_2Ph	Ph	-	<0.1	+	+
196	T1	CH_3	4-SO_2NH_2Ph	4-FPh	-	<0.1	+	+
197	T1	Ph	4-SO_2NH_2Ph	CH_3	-	<0.1	+	+
198	T2	SO_2CH_3	4-ClPh	H	CF_3	<0.1	+	+
199	T2	SO_2CH_3	4-ClPh	H	4-FPh	0.1	+	+
200	T2	SO_2CH_3	3-Cl-4-CH_3Ph	H	CF_3	0.1	+	+
201	T2	SO_2NH_2	4-ClPh	H	CF_3	0.1	+	+
202	T2	SO_2NH_2	3-Cl-4-CH_3Ph	H	CF_3	<0.1	+	+
203	T3	F	CH_3	H	-	<0.1	+	–
204	T3	CF_3	CH_3	H	-	<0.1	+	+
205	T3	CH_3	CH_3	H	-	<0.1	+	+
206	T3	F	CH_3	SO_2CF_3	-	<0.1	+	+
207	T3	F	CH_3	CH_2O-3-ClPh	-	<0.1	+	+
208	T1	CH_3	4-SO_2NH_2Ph	4-Pyridyl	-	>100	–	–
209	T1	CO_2H	4-SO_2NH_2Ph	Ph	-	>100	–	+
210	T1	CH_3	4-SO_2NH_2Ph	3-Cl-4-OCH_3Ph	-	86.7	–	–
211	T1	*n*-Pr	4-SO_2NH_2Ph	Ph	-	51.0	–	+
212	T1	CO_2H	4-SO_2NH_2Ph	3-F-4-OCH_3Ph	-	35.7	–	–
213	T2	SO_2CH_3	4-Pyridyl	H	CF_3	>100	–	–
214	T2	F	4-SO_2CH_3Ph	OH	CF_3	40	–	–
215	T2	SO_2CH_3	2-Me-4-Pyridyl	H	CF_3	9.6	–	–
216	T2	F	4-SO_2CH_3Ph	H	CF_3	4.7	–	–
217	T2	SO_2CH_3	4-ClPh	OH	CF_3	4.0	–	–
218	T3	$C(O)CH_3$	CH_3	H	-	2.9	–	–
219	T3	F	CH_3	CHO	-	3.2	–	–
220	T3	F	CH_3	CH_2OH	-	3.9	–	–
221	T3	F	H	H	-	10.2	–	–

analysis of the results obtained on the training (192 molecules) and examining (29 molecules) sets allowed to draw the graph of the frequencies of the pharmacophores realization in all these compounds taken together Fig. (**7**). These frequencies vary in the limits of 78-99% in dependence on different groups of active compounds.

Relative to inactive compounds, the frequencies of the pharmacophores realization lie in the range of 4-20%. A pharmacophore appearance in the structures of compounds belonging to the inactive class can be explained by structural affinity of active and inactive compounds. As it was noted before, an attempt had been done to model novel COX-2 inhibitors [4]. Hypothetical structures being potential COX-2 inhibitors had been proposed there. It was interesting to examine them by means of the prognostication system described above. From the results of our test, most compounds indeed contain the pharmacophores found from the **ETM** application, and the prognostication system qualifies them as active compounds. It can be seen from the examples of two compounds taken from this work Fig. (**8**).

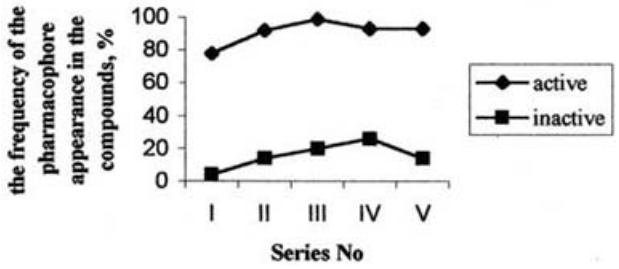
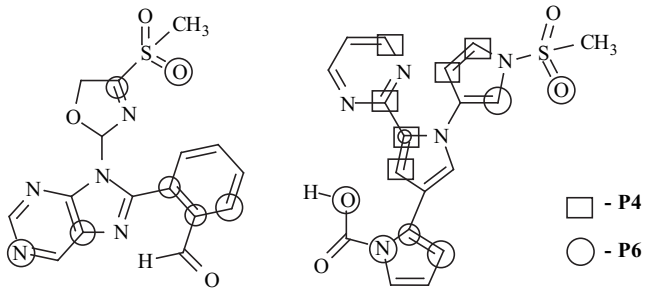


Fig. (7). Summary frequencies of all six pharmacophores realization in different types of skeletons representing active and inactive compounds in series **I-IV** combined with examining set **V**.

As seen from Fig. (**8**), both compounds (**222, 223**) contain pharmacophores **P4, P6**. It is worth to be noted that none of 6 anti-pharmacophores listed above has been found in the given compounds. Thus, they can be considered as potential inhibitors of COX-2.



222 **223**

Fig. (8). The found pharmacophores realization in two hypothetical structures taken from [4].

CONCLUSIONS

Based on the study of conformational and electron properties of a large series of 192 compounds belonging to different structural classes and possessing different levels of activity, a system for the COX-2 inhibitor activity prognostication was developed. The system includes 6 pharmacophores, 6 anti-pharmacophores, and conditions necessary for the activity demonstration by a compound. The system identifies active compounds with high enough probability (~0.96 in average) under condition that activities of compounds in the examining set are same as those used in this study ($IC_{50} \leq 0.1$ μM for active compounds and $IC_{50} \geq 2.0$ μM for inactive ones). The developed system was tested on an examining set of 29 compounds with a few types of skeletons, which differ from skeletons of the training set. Submatrices (**ETSC**s) that correspond to the found pharmacophores contain data on their spatial and electron characteristics and do not depend on the sorts of their atoms. It stems from the fact that **ETMC**s take into account three-dimensional (3D) compound structures and electron density distribution in the 3D space to describe compound structures. This circumstance appears to be very important when exploring compounds that belong to different structural groups. The results of this study agree with experimental data on biological activities of the studied compounds well enough.

ACKNOWLEDGEMENTS

A gratitude is expressed to TÜBITAK by E.S. for the financial support. All authors are also grateful to Dr. M. Göktepe and H. Gündüz from FAKO pharmaceuticals for their interest to this study and help.

ABBREVIATION

NSAIDs = Non-steroidal anti-inflammatory drugs

COX = Cyclooxygenase

ETM = Electronic-topological method

ETMC = Electronic-topological matrix of contiguity

ETSCs = *Submatrices* of ETMCs

REFERENCES

[1] Seibert, K.; Zhang, Y.; Leahy, K.; Hauser, S.; Masferrer, J.; Perkins, W.; Lee, L.; Isakson, P. *Proc. Natl. Acad. Sci. USA Pharmacology,* **1994**, *91,* 120013.

[2] Copeland, R.A.; Williams, J.M.; Giannaras, J.; Nurnberg, Sh.; Covington, M.; Pinto, D.; Pick, S.; Trzaskos, J.M. *Proc. Natl. Acad. Sci. USA Biochemistry,* **1994**, *91.* 11202.

[3] Wilkerson, W.W.; Copeland, R.A.; Covington, M.; Trzaskos. J.M. *J. Med. Chem.,* **1995**, *38,* 3895.

[4] Marnett, L. J.; Kalgutkar, A. S. *TiPS,* **1999**, *20*, 465.

[5] Zhu, J.; Yu, H.; Fan, H.; Liu, H.; Shi, J. *J. Comp.-Aid. Mol. Design,* **2001**, *15,* 447.

[6] Suh, Y.-G.; Kim, Y-H.; Park, M-H.; Choi, Y-H.; Lee, H-K.; Moon, J-Y.; Min, K-H.; Shin, D-Y.; Jung, J-K.; Park, O-H.; Jeon, R-O.; Park, H-S.; Kang, S-A. *Bioorg. Med. Chem. Lett.,* **2001**, *11,* 559.

[7] Hashimoto, H.; Maeda, K., Ozawa, K.; Haruta, J.; Wakitani, K. *Bioorg. Med. Chem. Lett.,* **2002**, *12*, 65.

[8] Wesolowski, S. S.; Jorgensen, W. L. *Bioorg. Med. Chem. Lett.,* **2002**, *12*, 267.

[9] Dimoglo, A. S. *Chim. Pharm. Zh. (Russ.)*, **1985,** 19, 438.

[10] Carter, J.S.; Obukowicz, M.J.; Davedas, B.; Talley, J.J.; Brown, D.L.; Graneto, M.J.; Bertenshaw, S.R.; Rogier, J.D.J.; Nagarajan, S.R.; Hanau, C.E.; Hartman, S.J.; Ludwig, C.L.; Metz, S. *US Patent,* **2000**, 6 034 256.

[11] Li, J.J.; Norton, M.B.; Reinhard, E.J.; Anderson, G.D.; Gregory, S.A.; Isakson, P.C.; Coboldt, C.M.; Masferrer, J.L.; Perkins, W.E.; Seibert, K.; Zhang, Y.; Zweifel, S.; Reitz, D.B. *J. Med. Chem.,* **1996**, *39,* 1846.

[12] 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.M.; Rogers, R.S.; Rogier, D.J.; Yu, S.S.; Anderson, G.D.; Burton, E.G.; Cogburn,

J.N.; Gregory, S.A.; Koboldt, C.M.; Perkins, W.E.; Seibert, K.; Veenhuizen, A.W.; Zhang, Y.Y.; Isakson, P.C. *J. Med. Chem.,* **1997***, 40,* 1347.

[13] Weier, R.M.; Lee, L.F.; Partis, R.A.; Koszyk, F.J. *US Patent,* **1999***, 5,* 916 905.

[14] Talley, J.J.; Brown, D.L.; Nagarajan, S.; Carter, J.S.; Weier, R.M.; Stealey, M.A.; Collins, P.W.; Rogers, R.S.; Seibert, K. *US Patent,* **1997***, 5,* 633 272.

[15] Khanna, I.K.; Weier, R.M.; Collins, P.W.; Yu, Y.; Xiangdong, X.; Partis, R.A.; Koszuk, F.J. *US Patent,* **1997***, 5,* 616 601.

[16] Khanna, I.K.; Richard, R.M.; Yu, Y. *US Patent,* **1999***, 5,* 935 990.

[17] Shvets, N. M.; Dimoglo, A. S. *Nahrung,* **1998,** *42,* 364.

[18] Dimoglo, A.S.; Beda, A.A.; Shvets, N.M.; Gorbachov, M.Yu.; Kheifits, L.A.; Aulchenko, I.S. *New. J. Chem.,* **1995***, 19,* 149.

[19] Dimoglo, A.S.; Vlad, P.F.; Shvets, N.M.; Coltsa, M.N. *New J. Chem.,* **2001**, *25,* 283.

[20] Dimoglo, A.S.; Shvets, N.M.; Tetko, I.V.; Livingstone, D.J. *QSAR,* **2001***, 20,* 31.

[21] Shvets, N.M. *Comp. Sci. J. of Moldova,* **1993,** *1*, 101.

[22] Shvets, N.M. *Comp. Sci. J. of Moldova*, **1997,** *5*, 309.

[23] Gilbert, K.; Gaevski, J. A MMPi Molecular Mechanics Program, Indiana University: Indiana, **1985**.

[24] Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P. *J. Am. Chem. Soc.,* **1985***, 107,* 3902.

[25] Pople, J.A.; Santry, D.P.; Segal, G.A. *J. Chem. Phys.,* **1972,** *43***,** 129.