Problems of Medical Chemistry (Moscow), 1997, v.43, No 1, pp 41-57.

DISCOVERY OF NEW ANTIULCER AGENTS BY COMPUTER AIDED PREDICTION OF BIOLOGICAL ACTIVITY

Trapkov V.A.*, Boudunova A.P.*, Burova O.A.**, Filimonov D.A., Poroikov V.V.#

Institute of Biomedical Chemistry RAMS, Russia, 119832, Moscow, Pogodinskaya Str., 10; National Research Center for Biologically Active Compounds*, Kupavna (Moscow Region), 142450, Russia; Chemical Pharmaceutical Institute**, Novokuznetsk, Russia

Abstract.

Specialised system for computer assisted prediction of antiulcer activity of chemical compounds on the basis of their structural formulae is described. Predicted activity spectrum includes antiulcer activity; antisecretory and gastroprotective effects; H_2- and M_1-receptors blockade, H^+, K^+-ATPase inhibition as the mechanisms of antisecretory effect. It is shown that the average prediction accuracy is 83 % in leave-one-out cross validation and 82 % for the independent test set. Prediction made for about 300 new chemical compounds provides the selection of 20 potential antiulcer agents, 9 of them were synthesised and tested, 5 compounds have potent antiulcer activity. The discovered antiulcer agents are classified as New Chemical Entry because antiulcer and close effects have not been earlier found for any compound of this chemical series.

About 5-10 % of the different age people in the world have the ulcer. Relapses of this disease accompanied by a dangerous complications occur in 70% patients with duodenal ulcers and 30% patients with gastric ulcers during 1 year after the remission. Therefore, the research and development (R&D) of new potent anti-ulcer drugs are still topical problem of modern pharmacology [1, 2].

In the past years Computer Aided Drug Design (CADD) is widely used in new drug R & D [3]. Recently we developed the computerised system PASS (Prediction of Activity Spectra for Substance) that estimates simultaneously the probability of more than 100 pharmacological effects and mechanisms [4,5]. The effectiveness of this computer aided approach application in screening has been shown to be 800% more than the random guess-work [5] and 300% more than the estimation by skilled experts [6]. However, the antiulcer action has not been covered by the initial version of PASS. Therefore, in this work we extend PASS prediction's area on antiulcer activities and use this specialised system to discover some new antiulcer agents.

Method of Computer Aided Prediction of Antiulcer Activity

We describe here the computer aided system for prediction of antiulcer actions of chemical compounds. The experience of development and application of PASS [4-6] has been taken into account. New specialised antiulcer systems have to comprehensively cover known antiulcer effects and mechanisms and include the appropriate reference antiulcer agents. As a consequence, the system provides more accuracy and reliability of the prediction's results [5]. The system can be also used as the model for further development of general approach to biological activity spectra's prediction [5].

To whom all correspondence should be addressed.

Basic components of any computer aided prediction system include: presentation of biological activity; description of chemical structure; training set consisted of well-known antiulcer agents and chemically related compounds which do not possess the antiulcer effects; mathematical method for structure-activity relationships analysis. These components are described below.

Biological Activity Description. Many different pathogenic factors cause the gastric and duodenal ulcer. The ulcer is considered to occur due to the disbalance between the aggressive intragastric components and resistance of upper gastrointestinal mucous. The acid secreted by gastric parietal cells is the main aggressive factor. For many years the clinical statement "no acid -- no ulcer" is the basis in pharmacotherapy of peptic ulcer by gastric acid's neutralisation or inhibition. This approach is sufficiently effective and antisecretory agents are widely used now for the treatment of ulcer, for example, the blockers of H_2-receptors (cimetidine, ranitidine, famotidine), the blocker of M_1-receptors pirenzepine, the inhibitor of H^+, K^+-ATPase omeprazole, *etc.* [7].
Research of agents that can enhance the resistance of upper gastrointestinal mucous to the action of ulcerogenic factors is an alternative trend in development of antiulcer drugs. Although the gastroprotection phenomenon has been described recently and its mechanisms are not studied in details, some new agents, for example, misoprostol, cetraxate, teprenone, sofalcone, gefarnate and plaunatol are discovered on its basis. These agents are launched as remedies not only for peptic ulcer treatment but also for decrease or prevention of nonsteroidal antiinflammatory drug's induced gastrotoxicity [8].
Both antisecretory and gastroprotective effects are covered by the antiulcer prediction system:

1. Antiulcer Activity
    1.1. Antisecretory Activity
        1.1.1. H_2-Receptors Blocker
        1.1.2. M_1-Receptors Blocker
        1.1.3. H^+, K^+-ATPase Inhibitor
    1.2. Gastroprotective Activity

Any compound that has antisecretory and/or gastroprotective action possesses antiulcer activity. Antisecretory agents have different mechanisms (1.1.1-1.1.3), while all the other antiulcer compounds have gastroprotective action.
We employ the qualitative representation of biological activities (presence or absence) which allows to use the results obtained by different authors in various tests in uniform manner [5].

Chemical Structure Description. There exist many characteristics of chemical compounds used as descriptors in SAR/QSAR: sub-structural fragments, geometric and topological indexes, physico-chemical properties, *etc.* For different kinds of biological activity in different chemical series particular descriptors appear more or less significant in appropriate SAR/QSAR relationships [3]. Thus, it is necessary to use the description that would be sufficiently exact to achieve consistent prediction but would not be so sensitive as to measure some random regularities.
We use the Substructure Superposition Fragment Notation (SSFN) proposed initially by V.V. Avidon and co-workers [9] and modified recently by V.G. Blinova and A.E. Leibov [10]. The applicability of these descriptors to the SAR analysis is justified a posteriory due to the high correlation between predicted and experimental data [4-6, 11].

The Training Set. The training set contains 160 antiulcer agents and 139 their congeners that have no appropriate activity. This sample represents the reference compounds for known chemical series and different antiulcer actions. It is necessary to notice that the accuracy of predictions depends strongly on the quality of the training set. Thus, informational search and data supplement should be provided permanently to improve the training set quality and to increase the reliability of the system.

The Mathematical Approach. We have studied about 200 different mathematical algorithms in estimating the possibilities to predict simultaneously presence/absence for many kinds of activity in heterogenic set of compounds [12]. One of the most effective and robust algorithm is described in [5]. This algorithm was used for structure-activity relationship's analysis in antiulcer prediction system.

As a result of prediction for every compound the estimate of a posteriory and a priory probability for each activity are calculated. If the a posteriory probability is more than the a priory probability the compound is suggested to have appropriate activity. However, in particular case the researcher must decide whether he prefers: to lose an active substance as a result of reducing the number of experiments or to test every compound with non-zero probability of appropriate activity.

The robustness of prediction results is provided by the exclusion of every compound from the training set during the training procedure [5]. Therefore, during the training we get also the estimates of prediction's quality by cross-validation (leave-one-out). The 1st kind error is done when for active compound a posteriory probability is less than a priory probability (the frequency of occurrence of the activity in compounds of the training set). The 2nd kind error is done when for inactive compound a posteriory probability is more than a priory probability. The average value of errors for every activity is given in table 1.

Table 1

Cross-validation of antiulcer actions prediction accuracy

Activity names	Number of compounds in the training set	Accuracy of active compound's prediction, %	1st kind errors, %	2nd kind errors, %
Antiulcer Activity	135	64	36	29
Antisecretory Activity	91	73	27	23
H_2-Receptors Blocker	44	64	36	27
M_1-Receptors Blocker	32	94	6	6
H^+, K^+-ATPase Inhibitor	4	100	0	0
Gastroprotective Activity	10	100	0	3
In Average:		83	17	14

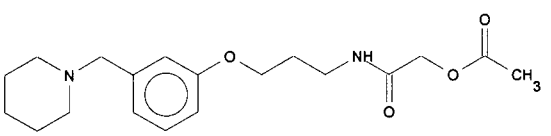
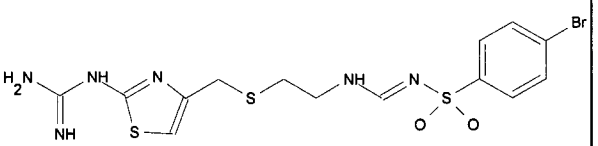
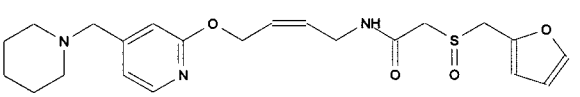
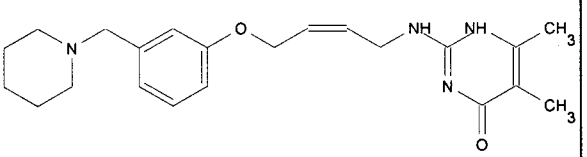
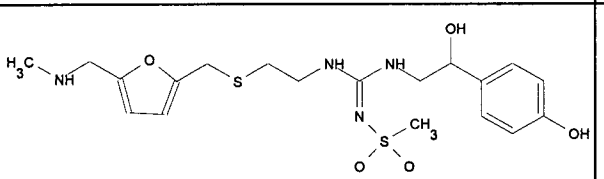
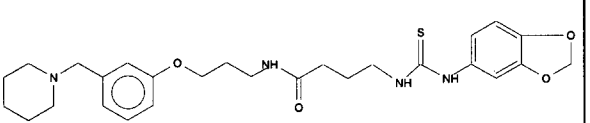
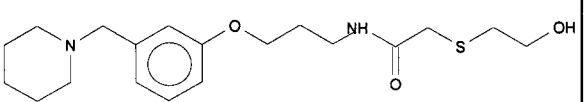
In average the 1st kind errors are more than 2nd kind errors. Mean accuracy of active compound's prediction is about 83%. Overprediction is about 14%. These characteristics are close to the appropriate mean values for all the 114 different kinds of activity [4,5,12].

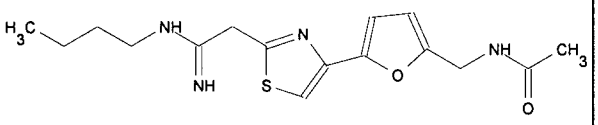
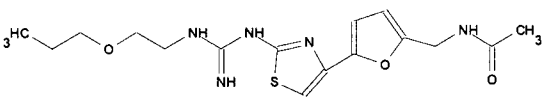
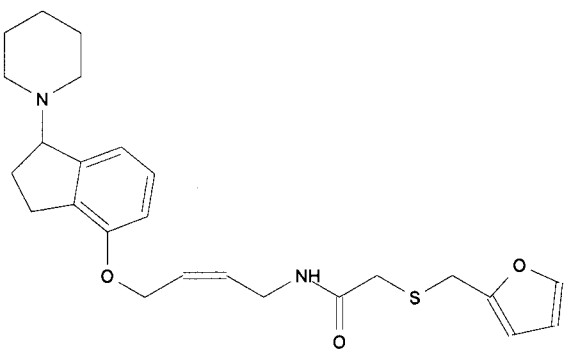
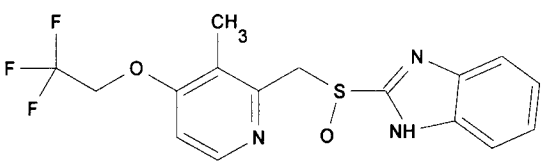
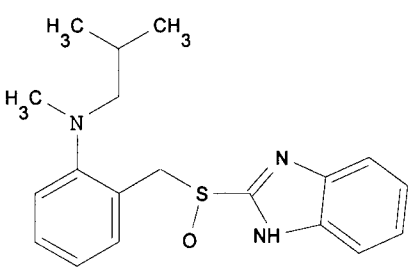
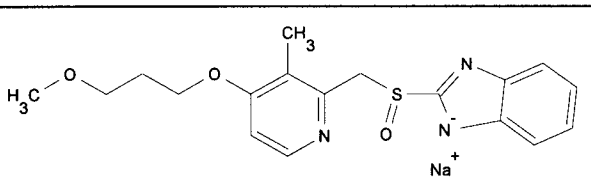
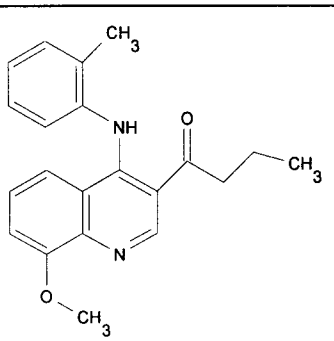
Validation of Prediction Ability on Independent Test Set of Compounds

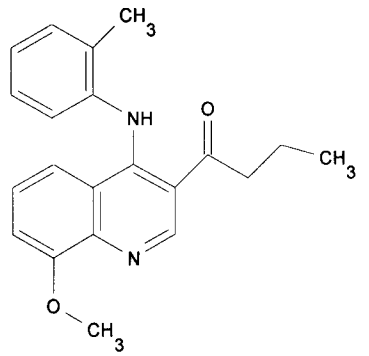
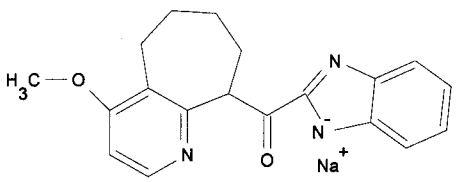
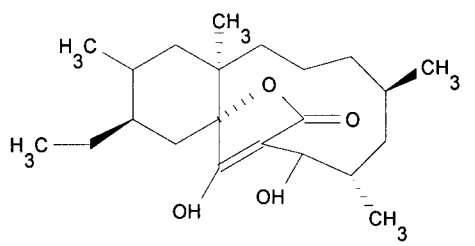
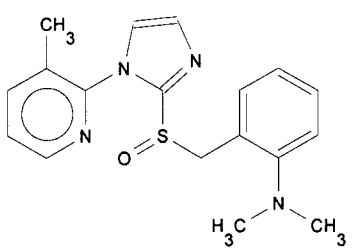
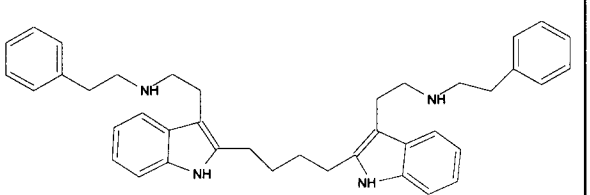
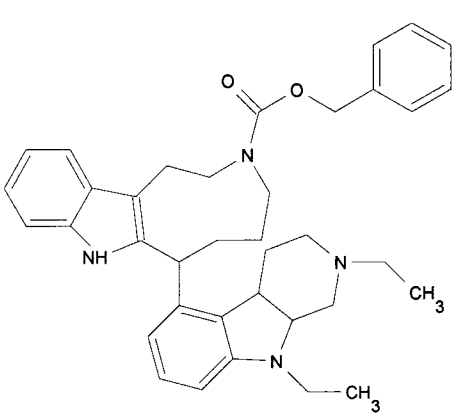
In order to check the validity of prediction we prepared the independent test set included 55 new antiulcer agents recently published in [13]. Structures and activities of these compounds are shown in table 2. Activity codes correspond to the above meanings. P is prediction code; E is experimental code. The presence of activity is marked by "+"; the absence of activity is marked by "-"; while unverified information is marked as "?".

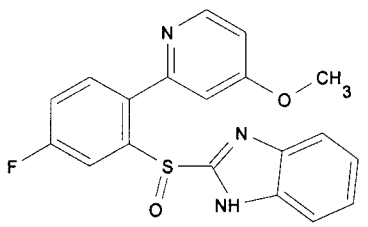
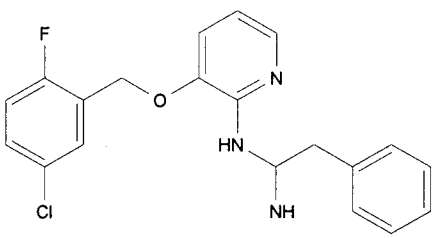
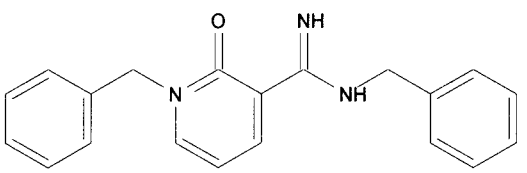
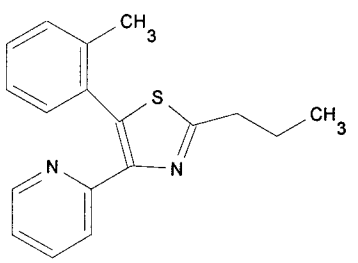
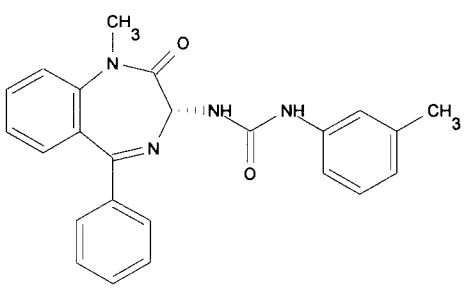
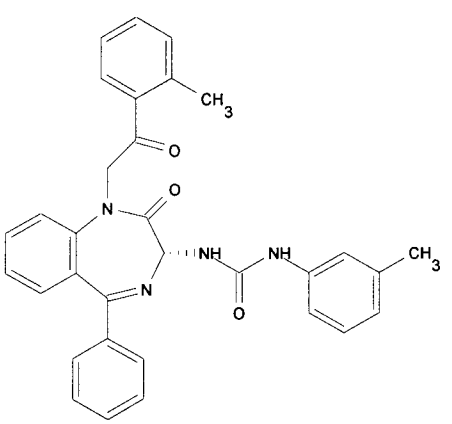
Table 2

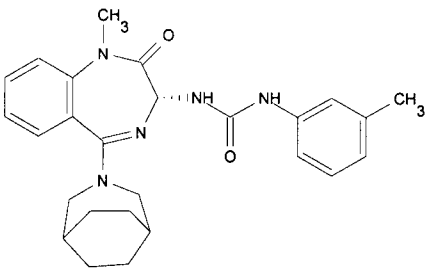
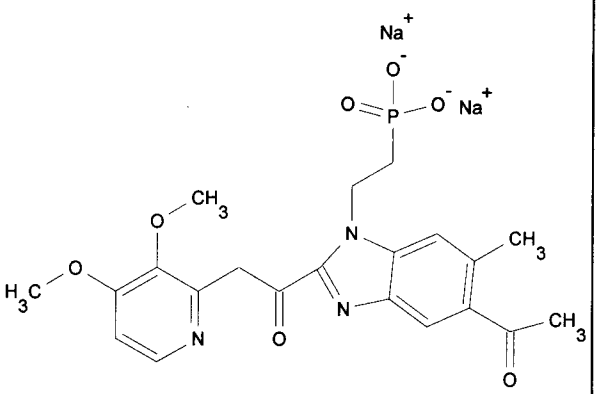
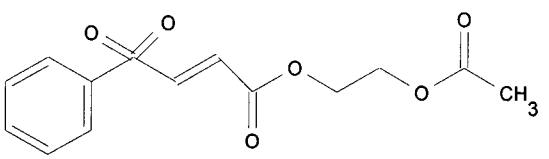
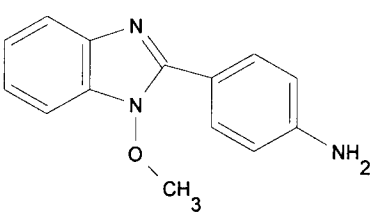
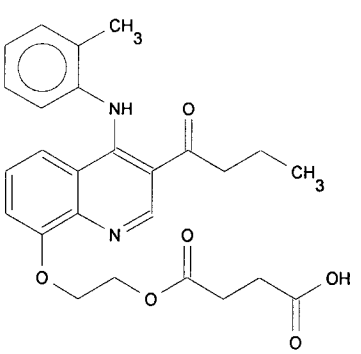
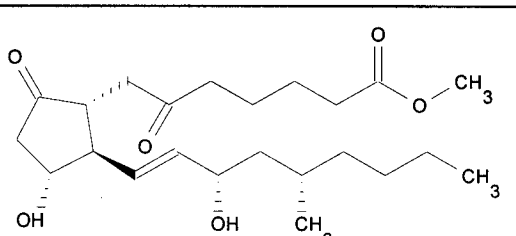
Comparison of Antiulcer Prediction for the Test Set with Experimental Data

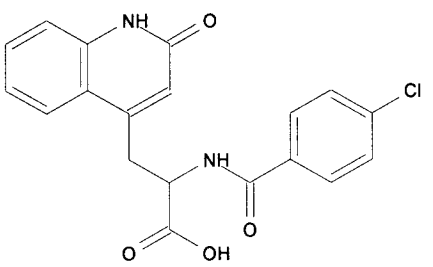
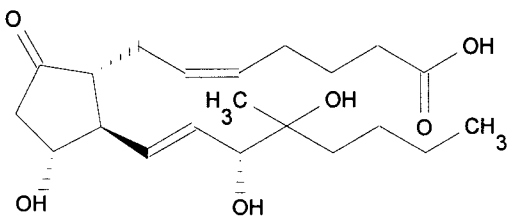
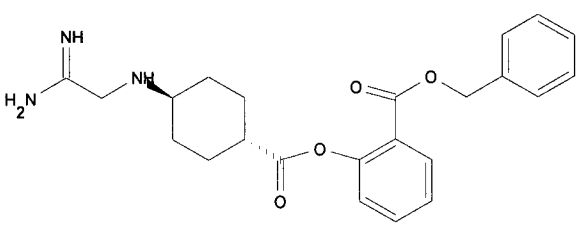
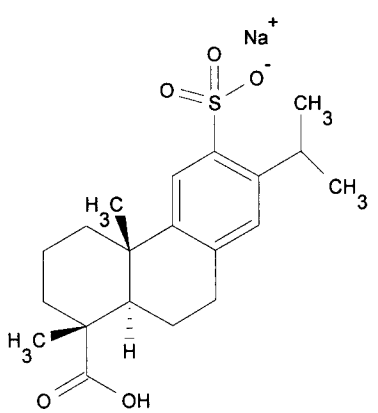
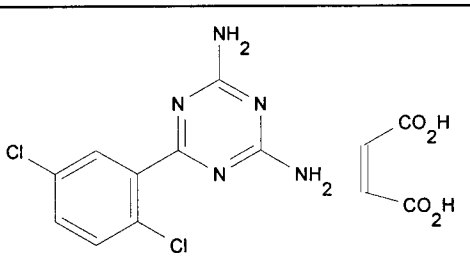
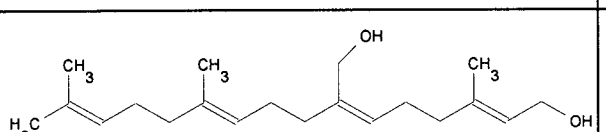
No	Drug Name, Firms	Structural Formula	Act. Codes	P/E
1	Roxatidine acetate, Teikoku Hormone		1 1.1 1.1.1 2	-/+ +/+ +/+ +/?
2	Ebrotidine, Ferrer		1 1.1 1.1.1	+/+ +/+ +/+
3	FRG-8813, Fujirebio; Taiho		1 1.1 1.1.1	+/+ +/+ +/+
4	IGN-2098, Grelan		1 1.1 1.1.1	+/+ +/+ +/+
5	T-593, Toyama		1 1.1 1.1.1	-/+ +/+ +/+
6	TRM-115, Terumo		1 1.1 1.1.1	+/+ +/+ -/+
7	Z-300, Zeria		1 1.1 1.1.1 2	-/+ +/+ +/+ +/?

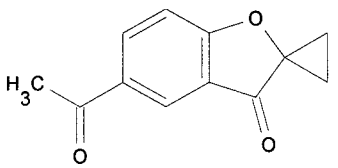
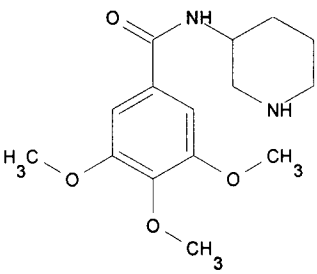
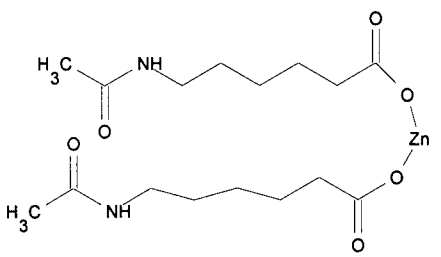
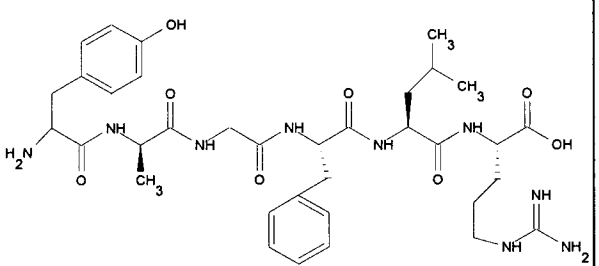
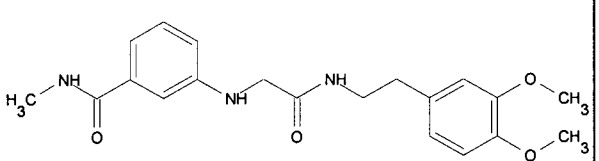
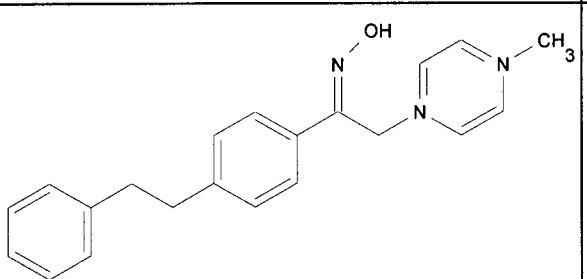
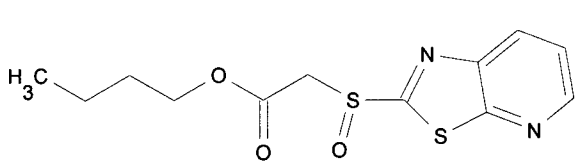
8	WO 9303028, Fujisawa		1 1.1 1.1.1	+/+ +/+ +/+
9	WO 9403450, Fujisawa		1 1.1 1.1.1	+/+ +/+ +/+
10	JP 93097837, Kyorin		1 1.1 1.1.1	+/+ +/+ +/+
11	Lansoprasole, Takeda		1 1.1 1.1.3	+/+ +/+ +/+
12	Leminoprazole, Nippon Chemipar		1 1.1 1.1.3	+/+ +/+ +/+
13	Rabeprazole, Eisai		1 1.1 1.1.3	+/+ +/+ +/+
14	SK&F-96067, SmithKline Beecham		1 1.1 1.1.3 2	-/+ -/+ -/+ +/?

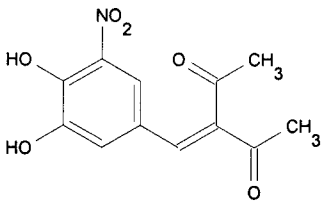
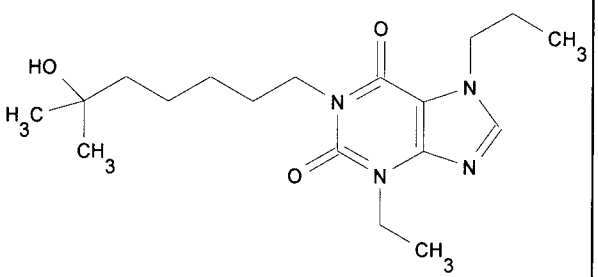
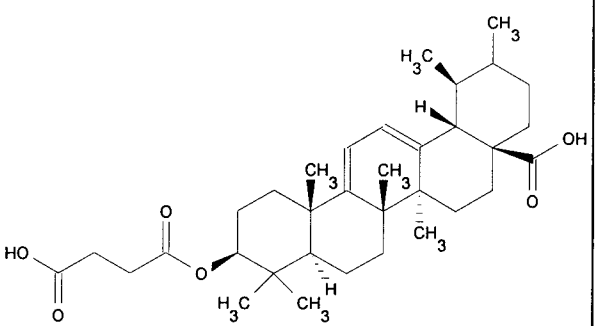
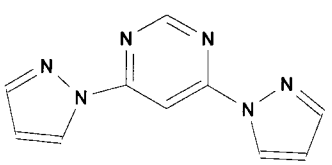
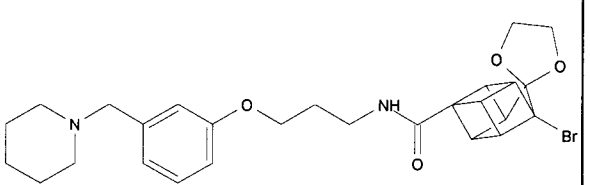
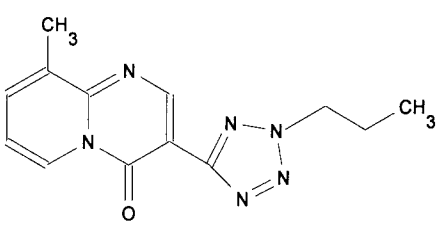
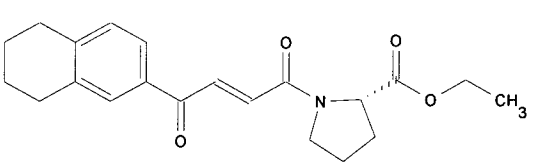
15	SK&F-97574, SmithKline Beecham		1 1.1 1.1.3 2	-/+ -/+ -/+ +/?
16	TY-11345, Toa Eiyo		1 1.1 1.1.3	+/+ +/+ +/+
17	A-88696, Lilly		1 1.1 1.1.3 2	+/+ -/+ -/+ +/?
18	T-330, Tanabe Seyaku		1 1.1 1.1.3 2	-/+ +/+ +/+ +/?
19	EP 537532, JP 93097801, Nisshin Flour Milling		1 1.1 1.1.3	-/+ -/+ -/+
20	EP 535529, JP 93092973, Nisshin Flour Milling		1 1.1 1.1.3	+/+ -/+ -/+

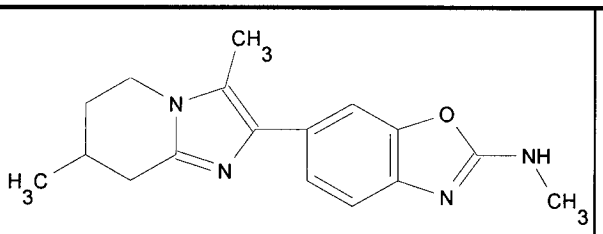
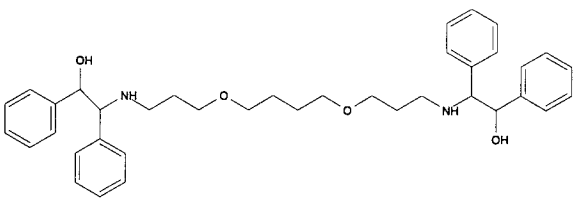
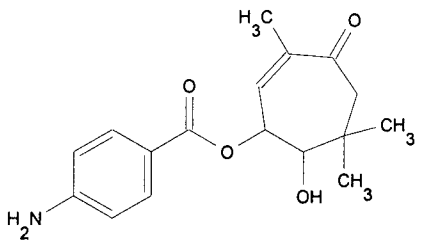
21	EP 526033, WO 9302077, Fisons		1 1.1 1.1.3	+/+ +/+ +/+
22	WO 9315055, SmithKline Beecham		1 1.1 1.1.3	+/+ -/+ -/+
23	WO 9315056, SmithKline Beecham		1 1.1 1.1.3 2	-/+ -/+ -/+ +/?
24	WO 9315071, SmithKline Beecham		1 1.1 1.1.1 1.1.3	-/+ +/+ +/? -/+
25	L-365260, Merck&Co.		1 1.1	+/+ +/+
26	YM-022, Yamanouchi		1 1.1	+/+ +/+

27	(-)-L-740093, Merck&Co.		1 1.1	+/+ +/+
28	WO 9312124, Astra		1 1.1 1.1.3	+/+ -/+ +/+
29	WO 9312079, Gedeon Richter		1 1.1	-/+ -/+
30	JP 302514079, Sankyo		1 1.1 1.1.3 2	-/+ -/+ +/? +/?
31	WO 9312090, SmithKline Beecham		1 1.1 2	-/+ -/+ +/?
32	Ornoprostil, Alloca		1 2	+/+ +/+

33	Rebamipide, Otsuka		1 2	+/+ +/+
34	Nocloprost, Schering AG		1 2	+/+ +/+
35	Benexate, Teikoku Chem		1 2	+/+ +/+
36	Ecabet Sodium, Tanabe Seyaku		1 2	+/+ +/+
37	Irsogladine maleate, Nippon Seyaku		1 2	+/+ +/+
38	Plaunotol, Sankyo		1 2	+/+ +/+

39	Spisofurone, Takeda		1 1.1 2	+/+ +/? +/+
40	Troxipide, Kyorin		1 2	+/+ +/+
41	Zink acexamate, Vinas		1 2	+/+ +/+
42	Dalargin, RAMS		1 2	+/+ +/+
43	Ecabapide, Daichi Seyaku		1 2	-/+ +/+
44	MCI-727, Mitsubishi Kasey		1 1.1	+/+ +/?
45	ME-3407, Meiji Seika		1 1.1 1.1.1	-/+ +/+ +/?

46	Nitecapone, Orion		1 1.1	+/+ +/?
47	A-90.6119, Hoechst AG		1 1.1	+/+ +/?
48	ISF-3401, ISF		1 2	+/+ +/+
49	IO-21, Nissin Food Prod.		1 1.1.1 2	-/+ +/? +/+
50	SWR-00104, Sawai		1 2	-/+ +/+
51	WO 9304065, Chinoin		1 1.1 2	+/+ +/? +/+
52	WO 9312070, Gedeon Richter		1 1.1 2	+/+ +/? +/+

53	JP 93025169, Fujisawa		1 1.1 1.1.1 2	-/+ +/? +/? +/+
54	US 5171753, A.H. Robins		1 2	-/+ -/+
55	EP 528699, Tokyo Tanabe		1 2	+/+ +/+

Results of validation are summarized in table 3.

Table 3

Validation of antiulcer action's prediction accuracy on the basis of independent test set

Activity names	Number of compounds in test set	Accuracy of active compound's prediction, %	1st kind errors (-/+), %	2nd kind errors (+/?), %
Antiulcer Activity	55	67	33	-
Antisecretory Activity	30	67	33	13
H_2-Receptors Blocker	10	90	10	7
H^+, K^+-ATPase Inhibitor	14	50	50	2
Gastroprotective Activity	19	95	5	16
In Average:		74	26	10

It is clear from table 3 that the average validity of prediction is satisfactory. For independent heterogenic chemical set it is shown that the frequency of 1st kind errors is 26%. The accuracy of active compound's prediction is slightly lower in comparison with the cross-validation (table 1) but still satisfactory to use this computerised system in practice.

The best results (95 and 90% accuracy) are shown for gastroprotective activity and H_2-receptors blockers; the worst result (50%) accuracy) is shown for H^+, K^+-ATPase inhibitors. The last one is explained by relatively small number of appropriate compounds in the training set and that all these compounds are from only one chemical series (benzimidazoles). Unfortunately, among the new antiulcer compounds in [13] there are no

M_1-blockers, therefore this set could not be used for validating the predictions of this activity.
It is necessary to emphasise that no antiulcer activity has been predicted only for two compounds (19 and 54 in table 2). If we try to use the system in selection of new substances for antiulcer screening we will test 53 from 55 compounds. Therefore, 96% of new compounds are exactly classified in there higher level of activity (53/55=0.96).

Prediction of Antiulcer Activity for New Compounds and Their Experimental Testing

To discover new antiulcer agents we predict the activities of about 300 compounds synthesised in Chemical Pharmaceutical Institute (Novokuznetsk). No one of these compounds have been suggested to have antiulcer activity prior the computer aided prediction. 20 compounds were recommended for testing as potential antiulcer agents on the basis of prediction. Example of the prediction for one of these compounds with registry number 8443688 [14] is given below.

Activity Name	Probability, %	Frequency, %
Antiulcer Activity	70	45
Antisecretory Activity	39	30
Gastroprotective Activity	12	15

Its antiulcer probability is 70% that is significantly more than 45% frequency to find antiulcer compound by the random search in the training set. This compound has been synthesised at first already in 1988 but it has not been never tested for antiulcer activity before the computerised prediction. Similar predictions are obtained for every of 20 compounds selected for antiulcer testing. Now 9 of them are synthesised and studied for antiulcer action as described below.

Model of acute ethanol-induced ulcer in rats. The method of acute ethanol induced injury is widely used for antiulcer agent's testing after the work of Robert et.al. [15]. It is presented below in more details.
Male albino rats weighing 170-220 g are fasted in individual cages with raised mesh bottoms for 24 h prior to the experiments, but allowed free access to water. Minimum 5 animals per treatment group are used in each experiment. Acute gastric haemorrhaged lesions are induced by intragastrical administration of 1 ml of 96% ethanol to each rat. The test compounds are suspended in 0.1% Tween-80 and are given intragastrically in dose 100 mg per 1 kg of body weight in 60 min prior the ethanol administration. Control animals are given the vehicle alone. The animals are killed in 1h after ethanol administration. The stomach of each animal is removed, inflated by injecting 8 ml of 2% formaline, immersed in 2% formaline for 10 min to fix both the inner and outer layers of gastric walls, and opened along the greater curvature. Thereafter, it is mounted on a cork plate to minimise mucous folding. The length of each lesion in mm is measured under a dissecting microscope (x 8), summed, and used as a lesion index. The person measuring the lesions did not know about the treatment given to animals. Enhancement or inhibition of ethanol

induced haemorrhages injury is calculated as percentage in comparison to the control (100%). Significance of differences is evaluated by Student's unpaired *t*-test.
Two known antiulcer agents Sucralfate and Cimetidine are used as the reference compounds with the same dosage and mode of treatment as for the test compounds.
The results of testing are given in table 4. The data in table 4 demonstrates that 5 from 9 compounds that were synthesised and tested have potent antiulcer effect. It is necessary to stress that the discovered antiulcer compounds can be classified as New Chemical Entries (NCE) [13] because antiulcer effect was not found earlier in this chemical series.

Table 4

Antiulcer Activity of New Compounds

Compound No	Lesion Index, mm	Damage Inhibition, %
Control	38 ± 4	
8445288	8 ± 2	80
Control	52 ± 10	
8949688	39 ± 9	6
10518693	59 ± 14	-13
10518893	49 ± 13	6
Control	57 ± 8	
10518793	16 ± 8	55
Control	41 ± 9	
8442588	13 ± 2	69
Control	41± 9	
8443688	12 ± 1	72
10571493	10 ± 3	76
10571593	27 ± 7	35
Control	71 ± 10	
Sucralfate	34 ± 9	52
Control	53 ± 9	
Cimetidine	42 ± 7	22

Notice. Active agents are marked in bold. The structures of compounds are not disclosed because of probable patenting.

Therefore, rather than synthesise and study 300 compounds, computer aided prediction allows to select only 20 compounds with high probability of activity. 5 new potent antiulcer agents were discovered by testing only 9 of these 20 compounds. The effectivity of computer aided prediction's use in this work is about 1500% (300/20=15).

Conclusions

1. Computerised system for prediction of antiulcer activity of chemical compounds of heterogenic structure is developed.
2. Average accuracy of antiulcer prediction is 83% in cross-validation and 74% for independent test set.
3. New antiulcer agents that can be classified as NCE are discovered by use this computer aided prediction system.

This work is supported in part by the Russian Ministry of Science and Technical Politics (Scientific Program "R & D of New Drugs by the Methods of Chemical and Biological Synthesis, Branch 04 "Computer Aided Drug Design", Project No 04-02-06).
We are grateful for Dr.J.R.Prous, the President of Prous Science Publishers, and Mr.J.Prous, Jr., the Director on Computer Applications of Prous Science Publishers, for their kind gift of "The Year's Drug News, 1994".

References

1. Emas S. Medicinal principles for treatment of peptic ulcer. // Scand. J. Gastroenterol. - 1987.-Vol. 22, Suppl. N 137.-P. 28-32.
2. Grigor'ev P.Ya., Shlevkov B.A., Agafonova N.A., Isakov V.A., Kolodkin A.M., Pryanishnikova A.S., Timoshenko G.D. // Pharmacotherapy of gastric and duodenal ulcers relapses. -Moscow: Medicine, 1990.-82 pp.
3. Computer Aided Drug Design in Industrial Research. / Franke R., Herrmann E.C. Eds. // Berlin: Springer-Verlag, 1994. - 252 pp.
4. Filimonov D.A., Poroikov V.V., Karaicheva E.I., Kazaryan R.K., Boudunova A.P., Michailovsky E.M., Rudnitskih A.V., Goncharenko L.V., Burov Yu.V. Computer Aided Prediction of Biological Activity Spectra for Chemical Compounds on the Basis of Their Structural Formulae. // Experimental and Clinical Pharmacology (Rus), 1995 - V.58 - N 2 - P.56-62.
5. Filimonov D.A., Poroikov V.V. PASS: Computerized Prediction of Biological Activity Spectra for Chemical Substances. // In: Design of Bioactive Compounds. Oxford (UK): BIOS Science Publishers, 1996. - In press.
6. Poroikov V.V., Filimonov D.A., Boudunova A.P.. Comparison of Prediction Results for the Chemical Compounds Spectrum of Biological Activity by Experts and the PASS System. // Automatic Documentation and Mathematical Linguistics. Allerton Press, Inc. - 1993 - V.27 - No 3 - P.40-43.
7. Bays D.E., Finch H. Inhibitors of gastric acid secretion // Natural Products Reports.- 1990.-Vol. 7.-N 5.-P. 409-445.
8. World's pharmaceutical market in 1980-1990. // Moscow: JV "Inpharmchem", 1991. - 93 pp.
9. Avidon V.V., Pomerantsev I.A., Rozenblit A.B., Golender V.E. Structure-activity relationship oriented languages for chemical structure representation. // J. Chem. Inf. Comput. Sci. - 1982 - V.22 - P.207-214.
10. Leibov A.E. Automatic coding of chemical structures in SSFN codes. // In: The Achivements in Science and Technology. Ser. Informatic. (Rus). Moscow: VINITI, 1991. - V.15 - P.141-158.
11. Shilova E.V., Filimonov D.A., Poroikov V.V. PASS: Prediction of Activity Spectra for Substance. Adrenergic Compounds Set. // In: Abstr. SCI Conference "Design of Bioactive Compounds: Possibilities for Industrial Application", L.:SCI, 1995. - P.28.
12. Filimonov D.A. Comparison of different algorithms for computerized biological activity spectra prediction on the basis of substance's structural formula. // In: Abstr. 2nd Russian National Congress "Man and Drug". Moscow, 1995. - P.62.
13. Prous J.R. The Year's Drug News. Therapeutic Targets.// Barcelona: Prous Science Publishers, 1994. - 549 P.
14. Burov Yu.V., Poroikov V.V., Korolchenko L.V. National system for registration and biological testing for chemical compounds: facilities for new drugs search. // Bull. Natl. Center for Biologically Active Compounds (Rus) - 1990 - No 1 - P.4-25.
15. Robert A., Nezamis J.E., Lancaster C., Hanchar A.J. Cytoprotection of prostaglandins in rats: prevention of gastric necrosis produced by alcohol, HCl, NaOH, hypertonic NaCl and thermal injury. // Gastroenterology. - 1979. - V.77. - P.433-443.