PASS: COMPUTERIZED PREDICTION OF BIOLOGICAL ACTIVITY SPECTRA FOR CHEMICAL SUBSTANCES

D.A. FILIMONOV, V.V. POROIKOV*

National Research Center for Biologically Active Compounds, Staraya Kupavna, Moscow Region, 142450, Russia

ABSTRACT

Computerized system PASS (Prediction of Activity Spectra for Substances) is described. PASS predicts more than 100 pharmacological effects, mechanisms of action and specific toxicities simultaneously. The average accuracy of prediction is about 80% both in leave-one-out cross-validation procedure for the total training set and by prediction for independent sample of 5000 biologically active compounds. The effectiveness of PASS used for heterogenic sets of new drug-candidates possess different biological activities is 800% better than random guess-work and 300% better than estimation by skilled experts. PASS used in drug R&D might help to select compounds having structures dissimilar to well-known drugs and to estimate probable activity spectra for new compounds to optimize their synthesis and testing.

INTRODUCTION

Experimental determination of drug efficacy and safety is a time- and cost-consuming procedure. There exist standard tests for drug safety assessment (Maggon *et. al.*, 1992) and different strategies of search for new lead compounds (Walker, 1994).

Biological testing is organized taking into account "similarity/dissimilarity" of new compounds to the other known biologically active substances. Several similarity/dissimilarity suggestions are used both in drug design and screening to determine if particular tests are necessary and sufficient for comprehensive estimation of new compound activity.

1. If the structure of a new compound is similar to the structures of some other biologically active substances this compound may be supposed to exhibit the same activities.
2. If the pharmacological effect of a new compound is similar to the effect of some other substances with known mechanisms of action this compound may be assumed to have the same mechanisms of action.
3. If the pharmacological effect of a new compound is similar to the effect of other known biologically active substances but its structure is not similar to any of these substances a new mechanism of action may be found.
4. If the structure of a new compound with unknown biological activity is not similar to the structures of any known biologically active substance it is possible that by investigating its activity in all available tests a new chemical entity (NCE) can be found.

As usual similarity/dissimilarity for new compounds is estimated by skilled experts (medicinal chemists and/or pharmacologists) based on their own knowledge and experience. Computerized system that would be able to estimate similarity/dissimilarity of a new compound to other known biologically active substances provides at least two advantages.

1. This computerized system generates the estimates taking into consideration the data base of biologically active compounds significantly more than the knowledge of any of experts.
2. Computerized prediction is much more deterministic and obviously more objective than expert's guess-work.

Lastly, the results of prediction can be considered by experts at the final stage and the ultimate decisions are taken by humans.

The computerized system PASS is described below. This system provides a means to estimate a new compound's similarity/dissimilarity to well-known biologically active substances and to predict its probable biological activities in wide range covering simultaneously many pharmacological effects, mechanisms of action and specific toxicity.

GENERAL DESCRIPTION OF PASS

Basic elements of system PASS include: description of chemical structure, presentations of biological activity, training set consisting of well-known biologically active substances; a mathematical approach for estimation the probability of any activity in a new compound. These elements are detailed below.

Chemical structure description

There exist many characteristics of chemical compounds that are used as descriptors in structure-activity relationship analysis: sub-structural fragments, geometrical and topological indexes, physico-chemical characteristics, etc. For different kinds of biological activity in various chemical series particular descriptors appear more or less significant in appropriate SAR/QSAR equations. 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 random irregularities in simultaneous prediction of a wide range of activity for non-congeneric compounds. In PASS the Substructure Superposition Fragment Notation (SSFN) is used (Avidon, 1974; Avidon *et. al.*, 1982, 1983). Similar substructural approaches were successfully applied also in some other works (Golender *et.al.*, 1983; Franke *et.al.*, 1985)

The applicability of these descriptors to SAR analysis is justified a posteriori by satisfactory correlation between predicted and experimental data (Poroikov and Filimonov, 1994; Filimonov *et. al.*, 1995). Moreover, in special experiments for a heterogenic set of adrenergic compounds it was shown that addition of various physico-chemical characteristics to SSFN descriptors does not increase significantly the accuracy of prediction made by PASS (Shilova *et. al.*, 1995). SSFN modification used in PASS 3.05 is described in detail (Leibov, 1991).

Biological activity

Biological activity is the result of a chemical compound's interaction with a biological entity. In clinical studies the biological entity is represented by the human organism. In preclinical testing it comprises experimental animals (in vivo) and experimental models (in vitro). Biological activity depends on peculiarities of compound (structure and physico-chemical properties), biological entity (species, sex, age, etc.), mode of treatment (dose, route, etc.).

Any biologically active compound may reveal a variety of different effects. Some effects are useful in treatment of definite diseases but others may cause various side and toxic actions. The total complex of activities caused by the compound in biological entities can be called "biological activity spectrum of the substance". The biological activity spectrum of a compound should present every activity of compound despite the difference in essential conditions of its experimental determination. If the difference in species, sex, age, dose, route, etc. is neglected the biological activity can be identified only qualitatively. Thus, "biological activity

spectrum" may be defined as the "intrinsic" property of a compound depending only on its structure and physico-chemical characteristics.

There exists a hierarchy in biological activities that corresponds to the natural biological hierarchy: activities can be defined and determined in organism, organ/tissue, cellular, molecular levels. If the compound is found to have several activities different levels, the activities at lower levels can be considered as the cause (mechanism) of effects at levels (Poroikov, 1988). If some compound's activity is determined, e.g., as antiasthmatic, broncholytic and as a beta-2 adrenergic agonist, it can be suggested that broncholytic activity caused the antiasthmatic effect, and beta-2 agonist activity caused the broncholytic effect. Therefore, beta-2 agonist action is the primary cause of antiasthmatic effect.

It should be stressed that we cannot rely on a single cause for any effect. In the above case although we know that there exists the correlation between beta-2 agonists and antiasthmatic effects we never certain that the discussed compound does not have other known or unknown activity that may cause the antiasthmatic effect, e.g., it may be muscarinic cholinergic antagonist and/or an antiinflammatory agent both being known antiasthmatic mechanisms.
Version 3.05 of PASS covers 114 kinds of biological activities including basic pharmacological effects, action mechanisms and specific toxicities that are listed in table 1.

The training set

The predictions are carried out on the basis of analysis of a training set containing about 10000 drugs and biologically active compounds. This set consists of reference compounds for known chemical leads and different biological activities. The number of reference compounds for every activity is shown in table 1. On average, one compound exhibits 2.23 biological activities and includes 15 SSFN structural descriptors.

To provide up-to-date quality of the training set informational search and data supplementation should be done continuously. It is necessary to use different informational sources in aggregate because the information taken from one publication never covers all aspects of biological action of the described substance. For example, according to (Negwer, 1987) caffeine (CAS No 58-08-2) is "stimulant", "analeptic" and "diuretic". As a result of informational search it was found that caffeine is: psychotropic, psychostimulant, analeptic, respiratory analeptic, cardiotonic, diuretic, saluretic, immunomodulator, immunosuppressant, spasmolytic, spasmogenic, vasopressor, vasodilator, hypertensive, nucleotide metabolism regulator, a cAMP phosphodiesterase inhibitor and a teratogenic and embriotoxic agent (some apparent contradictions in terms may be attributed to the reversing of effects for various doses of caffeine).

The mathematical approach

The algorithm of activity spectrum prediction for a new chemical compound is presented below. Structural descriptors of compound are interrelated to binary vectors of every activity presence/absence. Therefore, it is possible to describe this relationship by regression equations. But the training set includes more than 5500 different SSFN descriptors, and the resulting matrix would include more than 30.000.000 cells. So this approach does not seem to be real. However, it is possible to use the peculiarity of SSFN language, in which descriptors deal only with a small number of fragments suggested to be biologically valid. Each compound in the training set can be described by up 200 SSFN descriptors but the average number is about 15, so in appropriate binary vectors only 15 elements from more than 5500 equal 1, and the other equal 0. Therefore the appropriate matrix

is sparse and almost diagonal it is possible to obtain analytical approximation for regression coefficients (equation 1).
The following designations are used: i is the number of SSFN descriptors; k is the number of activities; n is the total number of compounds in the training set; n_i is the number of compounds containing descriptor i; n_k is the number of compounds revealing activity k; n_{ik} is the number of compounds with activity k and descriptor i; x_{ik} is regression coefficient i for activity k; $\hat{y}_k$ is the estimate for probability of activity k of the compound described by binary vector with the components f_i; x_{0k}, x_{ik} which are the coefficients of regression. Then the first approximation for x_{ik} is:

$$x_{ik} = \frac{n_{ik} - \frac{n_i \bullet n_k}{n}}{n_i + 1}, \qquad (1)$$

Similar estimates have been used earlier (Avidon *et. al.*, 1978) but in our work instead the obvious estimate

$$\hat{y}_k = x_{0k} + \sum_i f_i \bullet x_{ik} \qquad (2)$$

the special approach was applied. For descriptors of each compound the estimates (1) can be considered as a random sample of probability of each activity. Every compound consists a different number of descriptors, which we designate m. The m estimates of x_{ik} for each k are transformed in ascending order and three quantitative attributes for any kind of activity are calculated:

$$t_{1k} = \frac{1}{m+1} \sum_{j=1}^{m} x_{jk},$$

$$t_{2k} = \frac{1}{m+1} \sum_{j=1}^{m} (2j - m - 1) \bullet x_{jk} \qquad (3)$$

$$t_{3k} = \frac{1}{(m+1)^2} \sum_{j=1}^{m} (2j - m - 1)^2 \bullet x_{jk}.$$

These quantities are further used in equation (4) to estimate the probability of a compound possessing every activity by 4 parameters.

$$\hat{y}_k = a_{0k} + a_{1k} \bullet t_{1k} + a_{2k} \bullet t_{2k} + a_{3k} \bullet t_{3k} \qquad (4)$$

The factors a_{lk} (l = 0, 1, 2, 3) are found by least squares method:

$$a_{lk} = arg \min_{a_{lk}} \sum_{q=1}^{m} (y_{qk} - \hat{y}_{qk})^2, \qquad (5)$$

where: y_{qk} = 1, if compound q has activity k; and y_{qk} = 0 otherwise.

For realization of this algorithm in system PASS the table of n_k, n_i and n_{ik}, values is used where the SSFN descriptors are organized in binary tree. The appropriate set of values n_i and n_{ik} is extracted from the table and estimates x_{ik} are calculated for any of SSFN descriptor after its search in this tree. After the ordering of estimates x_{ik} the estimates of t_{1k}, t_{2k} and t_{3k} according to the equation (3) and the estimates of $\hat{y}_k$ according to the equation (4) are calculated. In case $\hat{y}_k$ <0 it is transformed into $\hat{y}_k$ = 0, and in case $\hat{y}_k$ >1 it is transformed into $\hat{y}_k$ = 1.

TABLE 1. List of biological activities predicted by PASS version 3.05

Biological Activity	N	E	Biological Activity	N	E
Adrenergic			*Analgesics and NSAIDs*		
Adrenergic stimulator	219	16	Analgesic	658	26
Alpha-1,2 adrenergic stimulator	87	13	Morphine-like analgesic	182	15
Beta-1,2 adrenergic stimulator	95	15	Non-morphine analgesic	299	18
Adrenergic blocker	279	18	NSAID	698	27
Alpha-1,2 adrenergic blocker	117	24	General anesthetics	40	23
Beta-1,2 adrenergic blocker	127	13	Local anesthetics	336	20
Sympatholytic	44	23	*Anticonvulsive and muscle-relaxants*		
Cholinergic			Anticonvulsant	400	28
Cholinergic stimulator	561	21	Antiparkinsonian, rigidity relieving	36	37
M-cholinergic stimulator	340	21	Antiparkinsonian, tremor relieving	116	25
N-cholinergic stimulator	261	22	Muscle relaxant	274	27
Cholinergic blocker	91	24	Curare-like muscle relaxant	112	16
M-cholinergic blocker	26	18	CNS-acting muscle relaxant	97	26
N-cholinergic blocker	63	24	*Spasmolytics*		
Peripheral M-cholinergic blocker	274	22	Bronchodilator	217	28
Ganglion blocker	167	21	Vasodilator	486	29
Histaminergic			Peripheral vasodilator	118	25
Histaminergic blocker	366	25	Coronary vasodilator	171	23
H1-histaminergic blocker	31	25	Papaverin-like spasmolytic	200	26
H2-histaminergic blocker	11	24	Spasmolytic	1099	30
Dopaminergic			*Smooth muscle-contraction stimulators*		
Dopaminergic stimulator	44	14	Spasmogenic	261	25
Dopaminergic blocker	35	22	Vasopressor	115	18
5HT-receptors			Abortion inductor	90	31
5HT-receptors blocker	173	28	*Blood pressure regulators*		
GABAergic			Hypertensive	137	25
GABA receptors stimulator	8	32	Hypotensive	738	34
Hormones, their analogs and antagonists			*Diuretics*		
Androgen	26	7	Diuretic	100	17
Androgen antagonist	20	23	Saluretic	70	17

Hestagen	38	4	Saluretic, reabsorbtion inhibitor	48	13
Hestagen antagonist	7	18	*Cardiovascular*		
Estrogen	89	10	Antiarrhythmic	270	22
Estrogen antagonist	30	12	Cardiotonic	114	22
Glucocorticoid	49	3	Cardiodepressant	143	31
Mineralocorticoid	7	4	*Antitussive*		
Thyroid hormone	6	4	Antitussive	132	25
Thyroid hormone antagonist	27	11	Narcotic antitussive	31	18
Prostaglandin	20	23	*Antitumor*		
Metabolism regulators			Antitumor	675	22
Anabolic	40	14	Antitumor-cytostatic	215	20
Nucleotide metabolism regulator	316	32	*Immune system regulators*		
Xanthine oxidase inhibitor	14	14	Immunomodulator	163	27
Uricosuric agent	44	15	Immunodepressant	129	25
cAMP PDE inhibitor	71	25	Interferon inductor	9	18
MAO inhibitor	146	24	Antiallergic	211	26
MAO inhibitor reversible	5	27	Blood clotting regulators		
MAO inhibitor irreversible	5	50	Coagulant	25	13
Lipid metabolism regulator	331	29	Anticoagulant	62	14
Cholesterol lowering	135	27	*Antibacterial, antiviral, etc.*		
Hypoglycemic	139	19	Antibacterial	890	25
Antioxidant	27	27	Antimycobacterial	262	18
Antihypoxic	30	38	Antiviral	365	27
ACHE inhibitor	58	32	Antihelmintic	223	23
ACHE inhibitor reversible	6	50	Antifungal	482	25
Psychotropic			Antiprotozoal	375	20
Antidepressant	282	24	Antitrichomonadal	71	21
Imipramin-like antidepressant	61	21	Antimalarial	116	16
Psychostimulant	82	24	Antispirochetal	30	18
Neuroleptic	227	13	*Miscellaneous*		
Sedative	451	31	Anorectic	75	16
Hypnotic	171	18	Choleretic	82	19
Tranquilizer	253	24	Radioprotective	139	28
Psychotropic (others)	1265	28	*Specific Toxicity*		
Narcotic or narcotic antagonist	88	17	Carcinogenic	37	35
Analeptics			Mutagenic	105	29
Analeptic	88	27	Teratogenic and/or embryotoxic	90	45
Respiratory analeptic	65	30			
Cardiovascular analeptic	17	36			

N is the appropriate number of compounds in the training set; E is the mean of 1st and 2nd kind error in per cent.

Since PASS should predict activities for a new compound, the estimation of parameters x_{ik} for every substance is made after its exclusion from the training set. The structure of data used at present allows us to do this easily: during the calculation of x_{ik} estimates of n_{ik} are reduced by 1 an n_i, n_k are reduced by 1 if the appropriate substance has activity k.

As a result of prediction the estimates of $\hat{y}_k$ are calculated which correspond to the a posteriori probability of activity k for appropriate compound (C_{conf}) and the ratio of a posteriori to a priori probability $C_{eff} = C_{conf}\, n/n_k$ (Figure 1).

INTERPRETATION OF PREDICTION RESULTS

The general quality of PASS prediction is determined by completeness/incompleteness of the training set; approximations of chemical structure description, biological activity presentation and mathematical model used in calculations. Thus, when considering prediction results it is necessary to bear in mind not only the values of C_{conf} and C_{eff} but also the "cost" of possible mistakes. When a new compound with novel activity (in relation to the training set) is studied it probably should be tested experimentally despite a low value of C_{conf}. The skilled 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.

All these reasons are valid if the training set is relatively complete in relation to the considered activity. So, if any doubts concerning the prediction result have arisen it is first necessary to check the training set and to add compounds with appropriate activity if necessary.

The experience of using PASS allows us to conclude that if $C_{conf} > 60\%$, the probability of finding the appropriate activity in experimental testing is high. But the compound may be the analog of a well-known drug from the training set. If $C_{conf} = 20\text{-}60\%$ the probability to find this activity for the compound is lower, but the compound's similarity to well-known drugs is also lower. If $C_{conf} < 20\%$, the probability of finding appropriate activity for the compound is rather low. But, if it will be found, it may be new. The most interesting is the case, when the value of C_{eff} is also high because it corresponds to the activity represented by a small number of substances in the training set.

VALIDITY AND EFFECTIVENESS TESTING

The prediction accuracy was estimated on the basis of the leave-one-out cross validation procedure. Average accuracy for all predicted activities was shown to be about 80%. The results of cross-validation for every activity is shown in Table 1. E is the average of 1st and 2d kind error obtained during cross-validation. It can be concluded that the average value of errors is about 20% and lower for many activities. Thus, the average accuracy of prediction is satisfactory.

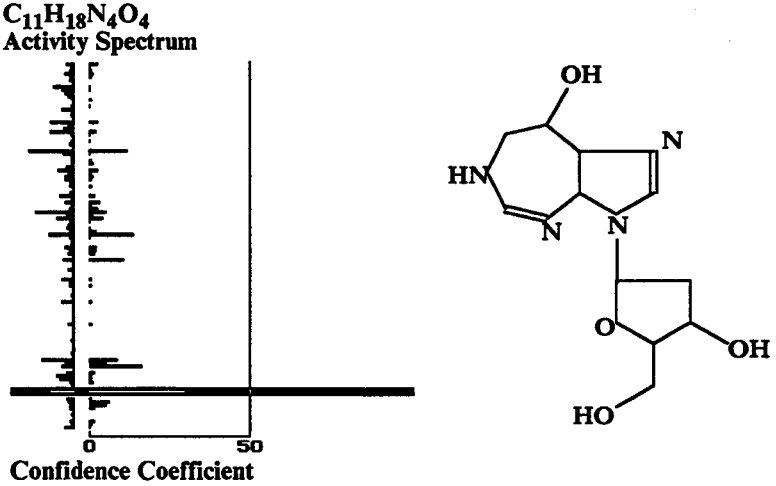

Fig.1. Activity spectrum predicted for pentostatin. The left spectrum illustrates the a priori probability (the frequency of every activity occurrence in the training set) and right spectrum illustrates a posteriori probability of each activity ($\hat{y}_k$). Their values are 7.2 and 30% respectively for antitumor action marked with black line. The ratio of a posteriori to a priori probability is 4.2. The presence of antitumor activity is confirmed by experimental data (Johnson *et. al.*, 1993).

We compared the accuracy of prediction by PASS with guess-work made by skilled experts (Poroikov *et. al.*, 1993). An independent sample including 33 new substances having 13 different activities was used for testing. The average accuracy of computer predictions about 64% (1st kind error is about 36%). Trials in which 10 medicinal chemists and pharmacologists guessed probable activities for this sample demonstrated an average accuracy below 20% (1st kind error is almost 80%). This shows a significant advantage of PASS over the experts in estimating probable activities.

The effectiveness of PASS use in preclinical testing of new compounds was estimated on the basis of a heterogenic set of 50 new drug-candidates which are related to about 100 various chemical classes and possessed 36 different biological activities (Prous, 1994). It was assumed that two strategies for screening may be used: blind-testing when each activity is studied systematically in order of activity numbers in Table 1 and predictions-oriented testing when the order of activities corresponds to the descending order of their a posteriori probabilities according to the PASS prediction. It was found that in blind-testing, to reveal real activities of every compound it is necessary to provide 2663 different experiments. If the testing is organized according to PASS prediction only 316 experiments are necessary to achieve the same purpose. In this case the effectiveness of PASS exceeds 800% (2663/316 = 8.43). It is necessary to mention that PASS predicts 41 out of 50 real activities giving an accuracy of prediction of 82%.

PASS APPLICATION IN DRUG R&D

PASS has been used in new drug R&D for several years both in the National Research Center for Biologically Active Compounds and in some other Institutes (Poroikov *et. al.*, 1994; Filimonov *et. al.*, 1995). There exist several examples when new information concerning biologically active compounds was found on the basis of PASS prediction. Among them: the discovery of new mechanism of action for substance with known effect, the finding of new chemicals with antiulcerogenic activity, the discovery of a non-mutagenic analog of a new antiarrhythmic drug-candidate which has mutagenic activity; etc. (Poroikov *et. al.*, 1994; Filimonov *et. al.*, 1995). The list of successful PASS applications might be more extensive if the possibilities of chemical synthesis and biological testing were not so limited.

CONCLUSIONS

The application of the computerized system PASS in the process of new drugs R & D provides in many cases the possibility to select compounds with desirable spectra of therapeutic effects and minimal side actions, prior to experimental testing or even synthesis. The system cannot predict all the possible properties for every compound because its possibilities are limited in particular by using an appropriate training set and list of activities. But PASS is open to further development oriented in the specific fields of interest of any researcher or company.

ACKNOWLEDGMENTS

This work is supported in part by the Russian Ministry of Science and Technical Politics (Scientific Program "R & D of New Drugs by Methods of Chemical and Biological Synthesis", Branch 04 "Computerized Drug Design", Project No 04-02-06). We are grateful to Dr. J.R. Prous, the President of Prous Science Publishers, for his kind gift of "*The Year's Drug News, 1994*". We thank DFG, Deutsche Forschungsgemeinschaft for supporting of our participation in the Conference "Design of Bioactive Compounds: Possibilities for Industrial Applications", Potsdam (Germany) 4-7 September, 1995.

REFERENCES

Avidon, V.V. (1974) The criteria of chemical structures similarity and the principles for design of description language for chemical information processing of biologically active compounds. *Khimiko-Pharmacevticheskii Zhurnal (Rus)*, **8**, 22-25.

Avidon, V.V., Arolovich, V.S., Kozlova, S.P., Piruzian, L.A. (1978) Statistical study of information file on biologically active compounds. II. Choice of decision rule for biological activity prediction. *Khimiko-Pharmacevticheskii Zhurnal (Rus)*, **12**, 88-93.

Avidon, V.V., Pomerantsev, I.A., Rozenblit, A.B., Golender, V.E. (1982) Structure-activity relationship oriented languages for chemical structure representation. *Journal of Chemical Information and Computational Sciences*, **22**, 207-214.

Filimonov, D.A., Poroikov, V.V., Karaicheva, E.I., Kazaryan, R.K., Boudunova, A.P., Michailovsky, E.M., Rudnitskich, A.V., Goncharenko, L.V., Burov, Yu.V. (1995) Computerized prediction of biological activity spectra for chemical compounds on the basis of their structural formulae: the system PASS. *Experimental and Clinical Pharmacology (Rus)*, **58**, 56-62.

Franke, R., Huebel, S., Streich W.J. (1985) Substructural QSAR Approaches and Topological Pharmacophores. *Environmental Health Perspectives*, **61**, 239-255.

Golender, V.E., Rosenblit, A.B. (1983) *Logical and Combinatorial Algorithms for Drug Design*, Research Studies Press, Wiley&Sons, 352 pp.

Johnson, J.B., Williams, M., Verburg, L., Israels, L.J., Begleiter, A. (1993) Combination therapy with deoxyadenosine plus 2'-deoxycoformycin and alkylating agents. *Proceedings of American Association of Cancer Research*, **34**, Abstr.1765.

Leibov, A.E. (1991) Automatic coding of chemical structures in SSFN codes. In: *The Achievements of Science and Technology. Ser.Informatic. (Rus)*. Moscow: VINITI (Russian Institute of Scientific & Technical Information), **15**, pp 141-158.

Maggon, K.K., Mechkovski, A. (1992) Total Quality Management. *Drug News and Perspectives*, **5**, 261-270.

Negwer, M. (1987) *Organic Chemical Drugs and Their Synonyms*, Berlin: Akademie-Verlag, **1-3**, 2470 pp.

Poroikov, V.V. (1988) Biomolecular structures data banks application to the prediction of new chemical compounds biological activities. In: *Proc. 3rd Conference "Theoretical Investigations and Data Banks in Molecular Biology and Genetics" (Rus)*. Novosibirsk, pp 62-63.

Poroikov, V.V., Filimonov, D.A., Boudunova, A.P. (1993) Comparison of the results of prediction of the spectrum of biological activity of chemical compounds by experts and the PASS computer system. *Automatic Documentation and Mathematical Linguistic*, New York: Allerton Press Inc., **27**, 40-43.

Poroikov, V.V., Filimonov, D.A. (1994) Computerized Prediction of Biological Activity Spectra for Chemical Structure - New Approach to Effective Drug Design. In: *The Abstracts of 10th European Symposium of Structure-Activity Relationships: QSAR and Molecular Modelling*, Barcelona (Spain), A136.

Prous, J.R. (1994) *The Year's Drug News. Therapeutic Targets,* Barcelona: Prous Science Publishers, 549 pp.

Shilova, E.V., Filimonov, D.A., Poroikov, V.V. (1995) PASS: Prediction of activity spectra for substances. Adrenergic compounds set. In: *Abstracts of the Conference "Design of Bioactive Compounds: Possibilities for Industrial Applications"*, London: SCI, pp 28.

Walker, S. (1994) The future of the Japanese pharmaceutical industry. - How can R&D be really international. *Pharma Japan, No 1424,* Tokyo: Yakugyo Jiho Co., Ltd., pp 7-10.