Online resources for the prediction of biological activity of organic compounds*

D. S. Druzhilovskiy, A. V. Rudik, D. A. Filimonov, A. A. Lagunin, T. A. Gloriozova, and V. V. Poroikov*

V. N. Orehovich Institute of Biomedical Chemistry, 10/8 ul. Pogodinskaya, 119121 Moscow, Russian Federation. E-mail: dmitry.druzhilovsky@ibmc.msk.ru

Online resources (PASS Online, SuperPred, SwissTargetPrediction and DRAR-CPI) for the prediction of biological activity of organic compounds from their structural formulas were considered. Based on a test set of drugs approved by 2014, the accuracies of predictions were compared. The four web resources can be arranged with respect to the quality of prediction (sensitivity, S) as follows: SwissTargetPrediction ($S = 0.37$) < DRAR-CPI ($S = 0.41$) < SuperPred ($S = 0.53$) < PASS Online ($S = 0.95$). A conclusion was made that PASS Online employs superior machine learning algorithms based on MNA descriptors and Bayessian classifier in contrast to the similarity-based methods used in SuperPred and SwissTargetPrediction or the molecular docking methods used in DRAR-CPI. Possible reasons for the low prediction quality of SuperPred, SwissTargetPrediction, and DRAR-CPI are discussed and the development perspectives of this area of computational chemistry are given.

Keywords: structure-property relationship analysis, biological activity profile prediction, online resources.

Introduction

The number of synthesized and available organic compounds currently exceeds 60 million,[1] that of virtual (computer generated) organic compounds being more than 166 billion.[2] Theoretically, there can be an infinite number of organic molecules, because the addition of a methyl group to an existing molecule results in a new one.

As full member of RAS N. S. Zefirov repeatedly noted, there are two fundamental problems in chemistry, namely, to perform structural transformations (organic synthesis, studies of reaction mechanisms, search for new reactions and reactants, *etc.*) and to study the structure and properties of matter. "The object of medicinal chemistry is to search and create physiologically active compounds, to find the relationships between the chemical structure and physiological activity and, finally, to solve the inverse problem, that is, to design the desired structures with specified properties".[3]

Computational methods are now widely used to search for and to design novel physiologically active compounds with preset properties.[4] Computational chemistry methods aimed at evaluating the properties of organic compounds began to develop about six decades ago.[5—8] The search queries "structure—activity relationship" or, structure—property relationship or, molecular modeling" in the PubMed informational system give more than 200 thousand references to publications. Actually, there are much more studies in the field because many of them were carried out on commercial basis in chemical, pharmaceutical and biotechnological companies and the results obtained are subject to IP protection.

The scientific school headed by acad. Zefirov has made a considerable contribution to the development of computational chemistry, *viz.*, the addition of "Zefirov" to the query mentioned above results in 53 references (this is also incomplete data; for the full publication list see the web site[9]). These studies, in particular, concern the development of methods for the description of organic reactions through formal logics,[10] computer-aided solvation of molecular design problems,[11] computer generation of structural formulas of organic compounds,[12] virtual synthesis of organic compounds based on structural fragments,[13] conversion of molecular graphs to 3D structures,[14] solution of the inverse analysis problem of Quantitative Structure—Activity Relationships (QSAR) and Quantitative Structure—Property Relationships (QSPR),[15] approximate calculations of electrostatic potentials for QSAR,[16] application of graph theory in organic chemistry,[17] formation of algebraic criteria for chirality and their application for classification of rigid molecular structures,[18] ap-

* Dedicated to Academician N. S. Zefirov on the occasion of his 80th birthday.

Published in Russian in *Izvestiya Akademii Nauk. Seriya Khimicheskaya,* No. 2, pp. 0384—0393, February, 2016.

1066-5285/16/6502-0384

plication of artificial neural networks for QSAR studies,[19] molecular field topology analysis for QSAR,[20] generation of molecular graphs for the "structure—property" studies,[21] *etc*.

Recently, the group headed by N. S. Zefirov has not only been developing computational chemistry methods, but also applying them to solving the practical problems of medicinal chemistry. For example, they performed molecular modelling of the interaction of various ligands with GABA receptor,[22] molecular design of *o*-phosphorylated oximes of selective butyrylcholinesterase inhibitors,[23] the search for the flaviviruses reproduction inhibitors using molecular dynamics simulations,[24] the design of allosteric modulators of metabotropic glutamate receptors,[25] *etc*.

Among various properties of organic compounds that can be predicted by computational chemistry methods, biological activity occupies a particular position. Manifestation of certain pharmacological effects by particular substances allows one to use their structures as the basis for the development of novel drugs, but adverse effects and toxicity may restrict their utilization.[26] All currently registered physiologically active compounds are known to possess several (sometimes many) types of biological activity.[27] Therefore, it is important to perform a complex computational evaluation of the biological activity profile (or spectrum) based on the structural formula of an organic compound — a "single-point (Q)SAR" (according to N. S. Zefirov). Evaluation of biological activity on the basis of structural formulas is very important because it can be done for virtual (not yet synthesized) molecules, which allows one to considerably reduce the time and costs of synthetic studies.

Software for the prediction of biological activity profiles

The development of methods for evaluation of biological activity profiles began in the 1970s.[28,29] It was shown that the "decision rules" for the prediction of biological activity profiles for novel compounds can be formulated upon analysis of "structure—activity" relationship.[30,31]

Our research activity in this field was initiated by the need to solve some problems within the framework of the USSR State System for registration of chemical compounds[32] because attempts to use the approaches developed by other authors had failed.[26] The PASS (Prediction of Activity Spectra for Substances) Version one software was released in the early 1990s.[33,34] It was based on the analysis of "structure—activity" relationships for a training set including about 10 000 substances and allowed one to predict 114 biological activity types with an average accuracy of ~75%. We are continuosly upgrading the PASS software and currently it is based on a training set of about a million compounds and can predict more than 7000 types of biological activity with an average accuracy of more than 95%.[35] A detailed description of PASS, its validation, and many examples of its practical use were a subject of a separate publication.[35]

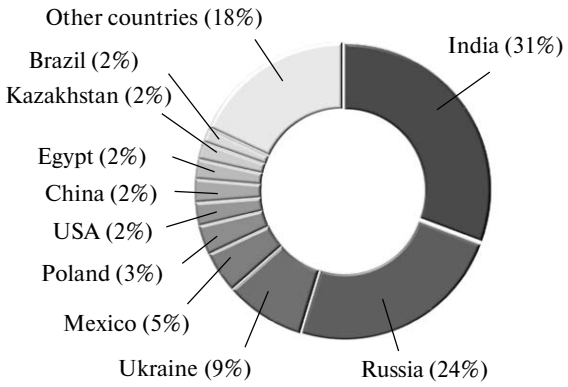


Fig. 1. Percentage of the PASS Online users from the world countries (12949 registered users by 01.06.2015).

Since 2000 we are maintaining the PASS Online website,[36—38] which allows registered users to obtain the online prediction results for about 4000 biological activity types.[39]. The system is currently used by 12949 researchers, graduate students, and students from 91 country, and more than 500 000 structural formulas have been analyzed (Fig. 1). Over 250 publications with experimental proofs of computational predictions have been published (see a review[35]).

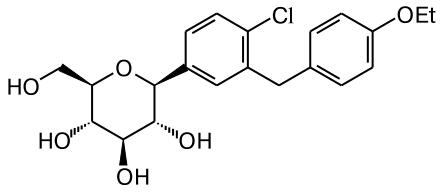
The creation of online databases of physiologically active compounds (PubChem,[40] ChEMBL,[41] DrugBank,[42] *etc*.) made it possible to develop a number of internet resources aimed at evaluating biological activity profiles of organic compounds. They employ computational methods that either evaluate the similarity of the analyzed structure to the structures of the compounds from the training set (for example, SuperPred[43,44] and SwissTargetPrediction[45,46]) or perform molecular docking of the analyzed structure to a pre-selected set of 3D models of target proteins and subsequent analysis of associations compared to the reference set of biologicaly active compounds (for instance, DRAR-CPI[47,48]). More detailed information on the evaluation methods used and references to relevant publications is available at corresponding websites.[43,45,47]

Evaluation of the accuracy of biological activity prediction using different online resources

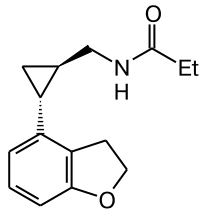
The existence of different online resources causes the need to compare the accuracy of predictions using the biological activity profiles. We constructed the test set on the basis of information on the structure and biological activity of medicinal drugs approved by the FDA in 2014.[49] The total number of drugs was 41; however, computed predictions can not be obtained for 11 of them and four

compounds are mixtures of two or more medicinal substances. Taking into account this fact, the test set included 36 different mono-component structural formulas (see below).

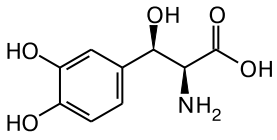
Table 1 summarizes the data on the known molecular action mechanisms and pharmacotherapeutic applications of these drugs.[49,50] The action mechanism is unknown for three drugs (Doxidopa, Miltefosine, Pirfenidone).[46] However, it is obvious that Doxidopa exhibits an adrenergic effect because it is a precursor of adrenaline (a prodrug). As to the other two drugs, the Integrity database contains information that Miltefosine is an angiogenesis inhibitor (it is not clear whether this is related to its antileishmanial effect), while Pirfenidone inhibits transforming growth fac-



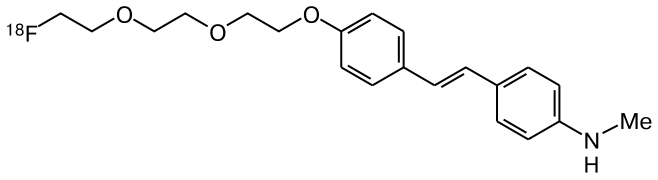
Dapagliflozin



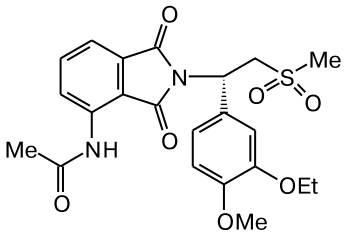
Tasimelteon



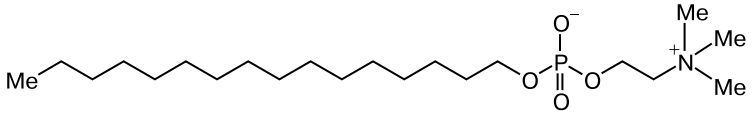
Droxidopa



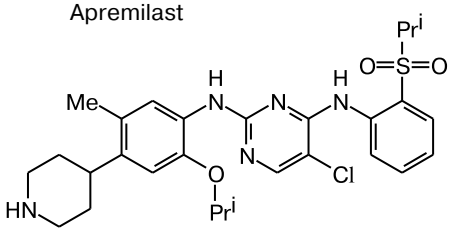
Florbetaben-F^{18}



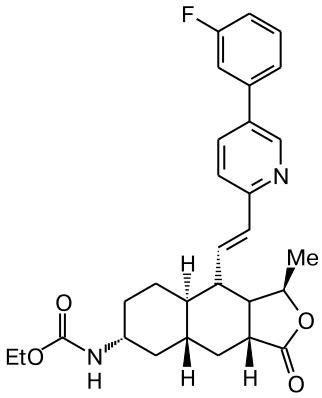
Apremilast



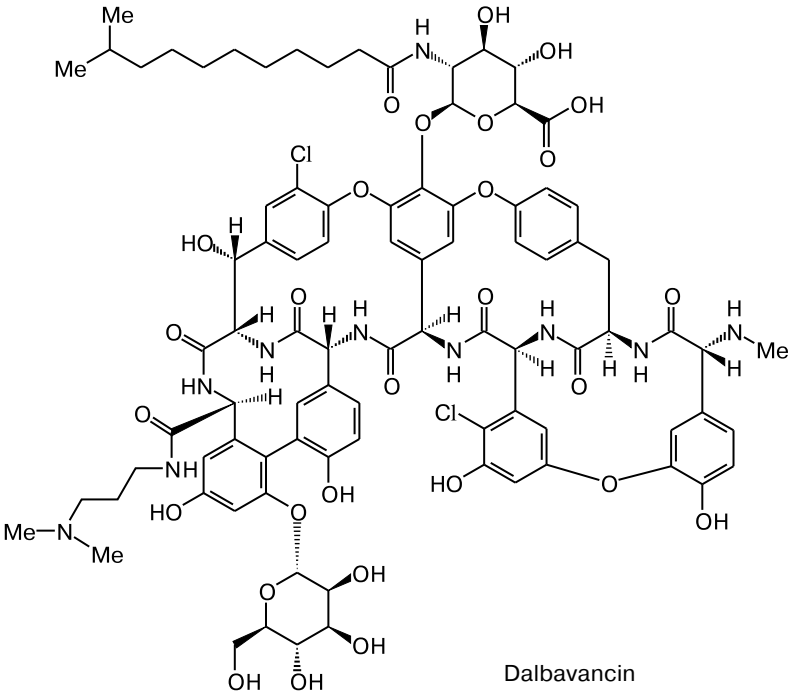
Miltefozin



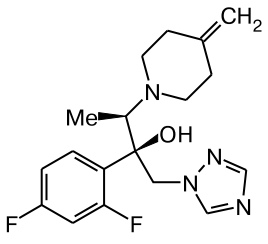
Ceritinib



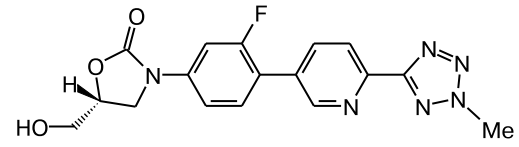
Vorapaxar



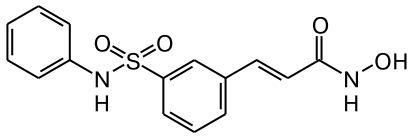
Dalbavancin



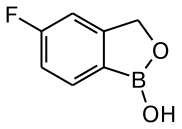
Efinaconazole



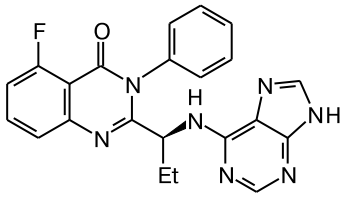
Tedizolid



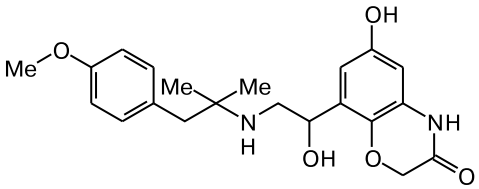
Belinostat



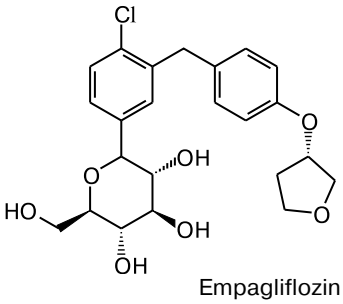
Tavaborole



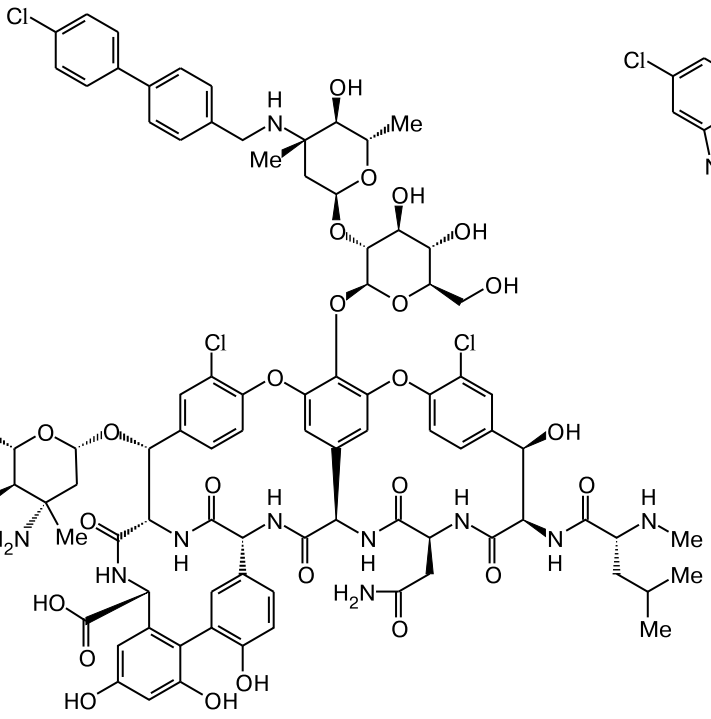
Idelalisib



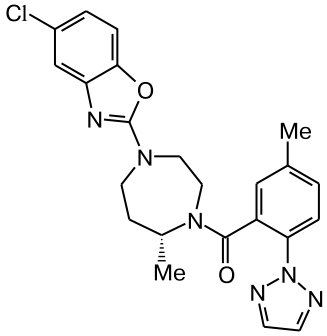
Olodaterol



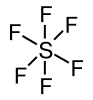
Empagliflozin



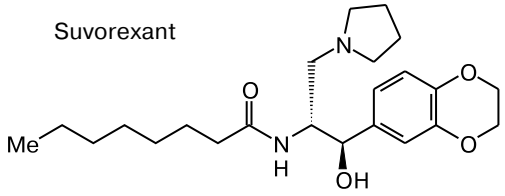
Orivatancin



Suvorexant



Lumason

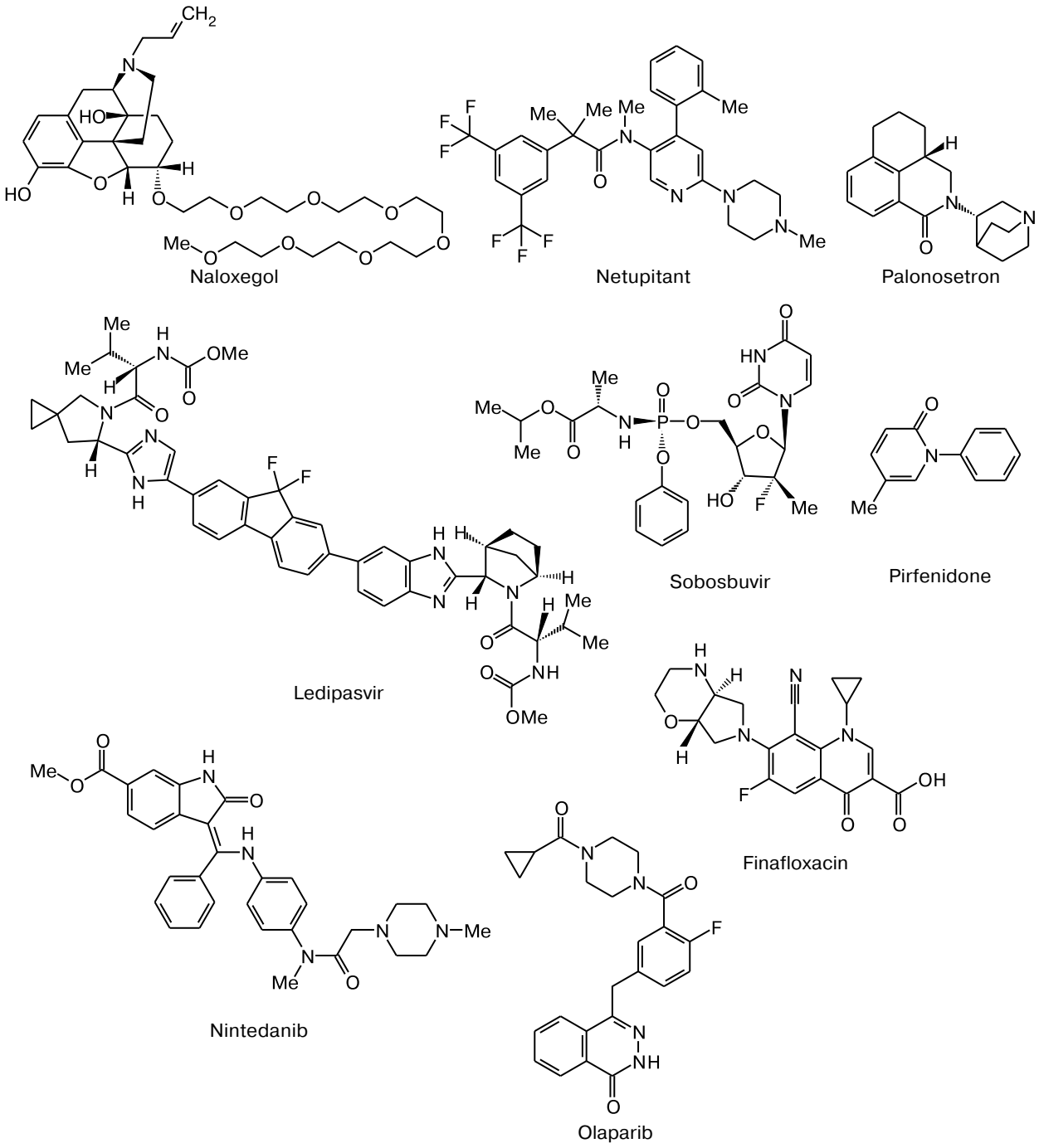


Eliglustat

tor beta and the production of tumor necrosis factor (these mechanisms may be related to therapy of idiopathic pulmonary fibrosis). For some drugs (Efinaconazole, Tedizolid, Tavaborole, Ceftolozane), only data on the chemical classes of antimicrobial drugs are available,[46] while the data on their mechanisms of action we retrieved from the Integrity database. Data presented in Table 1 demonstrate the variety of chemical structures and biological activity types for the compounds of the test set, thus indicating a good representativity of this set. Since our research was aimed to evaluate the quality of prediction, whether the compounds from the test set were or were not included in the training sets of the online resources studied was nonessential — all of them were tested under identical conditions. Moreover, we showed earlier[51] that exclusion of equivalent structures from the training set of the PASS software does not significantly influence the quality of prediction.

Input of the structural formulas in order to get the predictions was performed in the format of the MOL files (PASS Online), SMILES notation (SuperPred and SwissTargetPrediction), and MOL2 files (DRAR-CPI).[52] Conversion from the MOL format to other formats was made by the MarvinSketch software.[53]

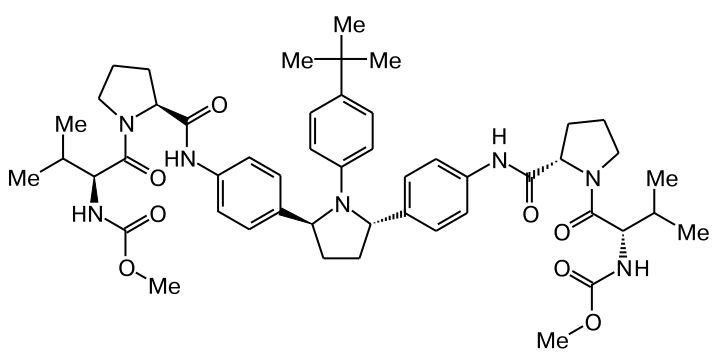
The computed predictions of biological activity of 36 molecules from the test set and comparison with the



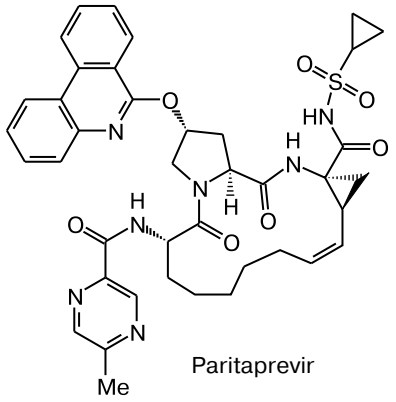
experimental data are given in Table 1. Possible predicted actitity types were chosen based on the default criteria used by the particular websites.

An analysis of the data in Table 1 shows that for PASS Online the number of prediction rejections was 3 of 36. Since the drug Lumason contains less than 3 carbon atoms, while Dalbavancin and Orivatancin have a molecular weight exceeding 1250, both cases are treated by the PASS system as exceptions going beyond the applicability range determined by the training set. The DRAR-CPI system rejected a prediction for two drugs — Dablavancin and Orivatancin, because the size of the corresponding MOL2 files exceeded 12 KB (this limitation was stated on the website). The SwissTargetPrediction system rejected a prediction only for Lumason.

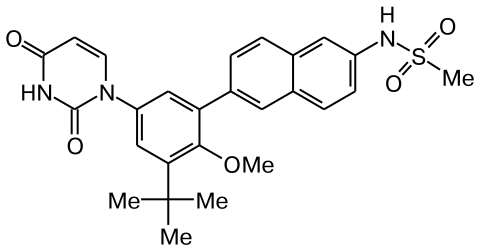
Below we compare the output data for the PASS Online, SuperPred, SwissProt and DRAR-CPI online resources taking Tasimelteon as an example, because all



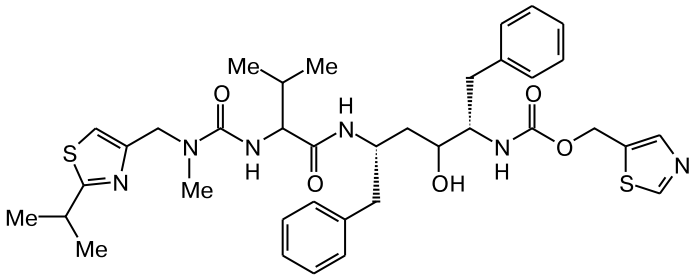
Ombitasvir



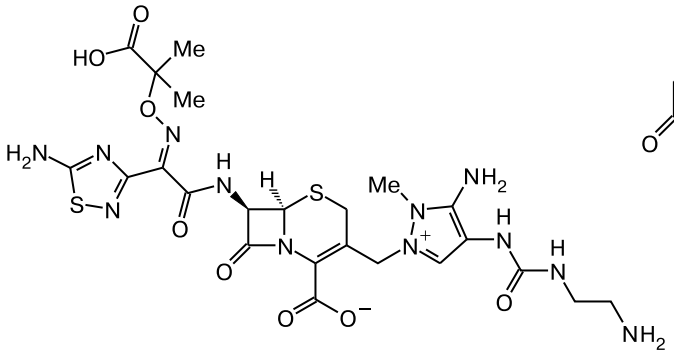
Paritaprevir



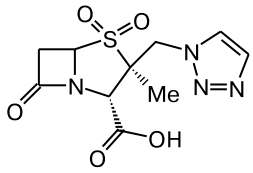
Dazabuvir



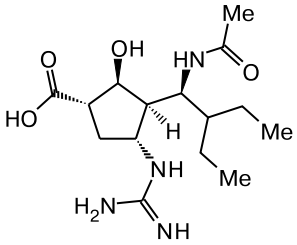
Ritonavir



Ceftozolan



Tazobaktam



Peramivir

four prediction systems gave proper evaluation of the known activitiy types.

The PASS Online gives the result of prediction in the form of a table (Table 2) containing the possible actitity types with the corresponding probabilities of activity (*Pa*) and inactivity (*Pi*).

As can be seen, the mechanism of action of this drug is melatonin receptor agonist, which is predicted with *Pa* in the range of 0.618—0.686, while its application for treatment of sleep-awakening disorders, is predicted with *Pa* = = 0.307. By default, all activity types with *Pa* > *Pi* are considered possible in the PASS system.

The SuperPred resource[43,44] performs the prediction in two modes, drug classification and target prediction. In the first mode, Tasimelteon was predicted with a probability of 70.59% (similarity by the Tanimoto coefficient[52] is 0.44) to be a melatonin receptor agonist (N05CH) by the ATC drug classification. The second mode predicted two targets, namely, melatonin receptor 1B (*E*-value: $1.972 \cdot 10^2$) and melatonin receptor 1A (*E*-value: $2.712 \cdot 10^2$). Therefore, both the molecular target and the pharmacotherapeutic application for this drug were predicted correctly.

The SwissTargetPrediction system[45,46] evaluates only the interaction of a particular structure with molecular targets. Along with the target name it gives its identifiers in the UniProt and ChEMBL databases, the probability of assignment of the compound to the particular target, and the number of similar compounds having the same target (based on the 2D and 3D analyses). The targets (83.3%) for Tasimelteon are melatonin receptor 1A, melatonin receptor 1B, melatonin receptor (by homology). It should be noted that in all three cases there are 450 and 147 similar compounds according to the 2D and 3D analy-

Table 1. Comparison of the biological activity profile prediction accuracy for the test compounds set by different online web resources

Drug name	Action mechanism	Pharmacotherapeutic effect (application area)	Prediction accuracy (mechanism/effect)			
			PASS Online	Super-Pred	Swiss-Target*	DRAR CPI
Dapagliflozin	Sodium glucose 2nd type cotransporter inhibitor	Second type diabetes	+/+	+/+	+	–/+
Tasimelteon	Melatonin receptor agonist	Sleep-awakening disorders	+/+	+/+	+	+/+
Droxidopa	Synthetic amino acid noradrenalin precursor	Precollaptoid state of patients with neurogenic orthostatic hypotension	+/+	+/–	–	+/+
Florbetaben-F^{18}	Radioactive diagnostic agent	Brain studies for evaluation of beta-amyloid hemolyth density	–/–	–/–	–	–/–
Miltefozin	Angiogenesis inhibitor	Leishmaniosis	+/+	–/–	–	–/–
Apremilast	Phosphodiseterase 4 inhibitor	Psoriatic arthritis and psoriatic plaque	+/+	–/–	–	+/+
Ceritinib	Inhibitor of ALK kinase, insulin-like growth factor 1 receptors, insulin receptors and protooncogenic tyrosin kinase ROS	ALK-positive metastatic parvicellular lung cancer	+/+	+/–	+	+/+
Vorapaxar	Protease-activated receptors 1 antagonist	Thrombotic cardiovascular complications	+/+	+/–	+	–/+
Dalbavancin	Semisynthetic lypoglicopeptide	Acute bacterial skin infections	NP	–/+	–	NP
Efinaconazole	Sterol-14-alfa-demethylase inhibitor	Onychomicosis of feet nails	+/+	+/+	–	–/+
Tedizolid	Antibacterial agent of the oxazolidinone class, protein biosynthesis inhibitor	Acute bacterial skin infections	+/+	–/+	–	–/+
Belinostat	Hystone deacytilase inhibitor	Peripheral T-cell lymphoma	+/+	+/+	+	–/+
Tavaborole	Antimycotic of the oxaborole class, leycyl tRNA syntase inhibitor	Onychomicosis of feet nails	+/+	–/–	–	–/+
Idelalisib	Phosphoinosytyl-3-OH kinase inhibitor	Chronic lymphocytic leukimia, B-cell non-Hodgkin´s lymphoma, small lymphocytic lymphoma	+/+	+/–	+	+/+
Olodaterol	Beta-2 adrenoreceptor prolonged agonist	Chronic obstructive lung disease	+/+	+/+	–	–/+
Empagliflozin	Sodium glucose 2nd type cotransporter inhibitor	Second type diabetes	+/+	+/+	+	–/+
Orivatancin	Semicynthetic lypoglicopeptide	Acute bacterial skin infections	NP	–/+	–	NP
Suvorexant	Orexine receptor antagonist	Insomnia	+/+	+/–	+	–/–
Eliglustat	Glucosylceramide synthase inhibitor	Gaucher disease	+/+	+/+	+	–/+
Naloxegol	Opiod receptor antagonist	Opioid-induced ileus	+/+	+/+	+	–/–
Lumason	Ultrasound contrast agent	Cardiovascular ultrasound studies	NP	+/+	NP	–/–
Netupitant	Neurokinyne receptor NK1 antagonist	Antiemetic	+/+	+/–	–	–/+
Palonosetron	5-HT(3) receptor antagonist	Prevention of induced nausea and vomitting	+/+	+/+	+	–/+
Ledipasvir	Viral hepatitis C protein NS5A inhibitor	Viral hepatitis C	+/+	–/–	–	–/–
Sobosbuvir	RNA-dependent RNA polymerase (NS5B) inhibitor	Viral hepatitis C	+/+	–/–	–	–/–
Pirfenidone	Beta tumor growth factor and tumor necrosis factor inhibitor	Idiopathic pulmonary fibrosis	+/+	–/–	–	+/–

(*to be continued*)

Table 1 (*continued*)

Drug name	Action mechanism	Pharmacotherapeutic effect (application area)	Prediction accuracy (mechanism/effect)			
			PASS Online	Super-Pred	Swiss-Target*	DRAR CPI
Nintedanib	Kinase (PDGFR, FGFR, VEGFR, and FLT3) inhibitor	Idiopathic pulmonary fibrosis	+/–	+/–	+	+/–
Finafloxacin	Fluoroquinolone type antimicrobial (swimmer´s ear)	Acute external otitis	+/+	+/+	–	–/–
Olaparib	poly-ADP ribosepolymerase inhibitor	BRCA-mutant ovary cancer	+/+	+/–	–	+/+
Ombitasvir	NS5A protein inhibitor	Viral hepatitis C	+/+	–/–	–	–/–
Paritaprevir	NS3 protease inhibitor	Viral hepatitis C	+/+	–/–	–	–/–
Dazabuvir	RNA-depentent RNA polymerase (NS5B) inhibitor	Viral hepatitis C	+/+	–/–	–	–/–
Ritonavir	HIV protease inhibitor	HIV infection	+/+	+/–	–	+/+
Ceftozolan	Antibacterial, bacterial cell wall synthesis inhibitor	Complicated intraabdominal infections and complicated urinary tract infections	+/+	+/+	–	–/–
Tazobaktam	Beta-lactamase inhibitor	Complicated intraabdominal infections and complicated urinary tract infections	+/+	+/+	–	–/–
Peramivir	Neuraminidase inhibitor	Flu	+/+	–/–	–	–/–

* SwissTargetPrediction predicts only the interaction with molecular targets.
NP — not predicted; «+» — prediction conforms to the experiment; «–» — prediction does not agree with the experiment.

Table 2. Prediction results of Tasimelteon drug biological activity by PASS Online resource

Pa	*Pi*	Activity name
0.707	0.011	Neurotransmitter uptake inhibitor
0.686	0.001	Melatonin receptor MT1 agonist
0.674	0.000	Melatonin receptor MT2 agonist
0.618	0.001	Melatonin receptor agonist
0.571	0.021	Trombocyte adhesion inhibitor
		...
0.307	0.011	Sleep disorder treatment
0.303	0.009	Serotonin uptake inhibitor

ses, respectively. With a lower probability (60%) NADH-dehydrogenase 1 (by homology) is identified as the target, as well as ribosylnicotinamide dehydrogenase *etc*.

The DRAR-CPI resouce[47,48] evaluates the association of the analyzed molecule with the possible pharmacotherapeutic effects by computing its interaction with a set of target proteins by the molecular docking methods. The reference set includes 254 ligands that interact with 385 targets. The evaluation is given as real numbers from -1 to +1 and represents the consistency/inconsistency of the reference interaction profiles with the interaction profile of the analyzed molecule. The Kolmogorov-Smirnov statistical analysis is used to calculate the «*p*-values» which range from 0 to 1 (the lower the «*p*-value» the higher the probability for association to exist). The results of Tasimelteon evaluation are given in Table 3.

Since the evaluations made for Tasimelteon by the DRAR-CPI system show an association with Ramelteon which is a melatonin receptor agonist used to treat sleep disorders, we concluded the prediction to agree with the experimental data.

A comparison of the evaluations given by all four online resources are given in Table 1, which demonstrates that the PASS Online system gives a correct prediction of the action mechanisms for 32 and 31 compounds of 33, respectively. Predicions made by the SuperPred system for the action mechanisms were correct for 22 of 36 test compounds, while the effect predictions were correct in 16 cases. The SwissTargetPrediction resource correctly evaluates the mechanisms for 13 compounds out of 35, while the DRAR-CPI system predicted correct mechanisms in 9 cases and correct effects in 19 cases out of 34.

Since the information on the biological activity of the test set compounds is presented in terms of the major pharmacotherapeutic effect and the molecular action mechanism and we have no screening data for these drugs for other activity types, the only attribute for the averaged quality assessment is "sensitivity":

$$S = \mathrm{TP}/\mathrm{NA},$$

where TP is the number of correctly predicted activity types and NA is the number of the known activity types.

Thus, the prediction quality for the test set analyzed is as folliws: SwissTargetPrediction (S = 0.37) < DRAR-CPI (S = 0.41) < SuperPred (S = 0.53) < PASS Online (S = 0.95).

Table 3. Evaluation of pharmacotherapeutic effects of Tasimelteon drug by DRAR-CPI web resource

Drug name	Application area	Association evaluation	«*P*-value»
Irinotecan	For the treatment of metastatic colorectal cancer (1st line therapy in combination with 5-fluorouracil and leucovorin)	−1	0.268
Metipranolol	For the treatment of patients with elevated intraocular pressure or open-angle glaucoma ...	0.999996	0
Ramelteon	For the treatment of insomnia due to sleep rythm disorders ...	0.925099	0

Conclusion

We reviewed a number of free online resouces for the prediction of biological activity profiles of organic compounds from their structural formulas. The accuracy of computer predictions depends both on the quality of the training set and on the evaluation criteria. Obviously, the training sets used by the four online resources studied are significantly different from one another.

At the same time, if there existed a perfect training set containing information about all the known biologically active compounds (in our case, about all medicinal drugs), the methods based on similarity should recognize the structures of the known drugs. An example is provided in this study, namely, a contrast agent for ultrasonic studies was correctly identified by the SuperPred system as belonging to the "Ultrasound contrast agent" (V08DA) category of the ATC classification. The PASS Online and SwissTargetPrediction systems gave no biological activity predictions in this case, because the substance was out of the applicability domain, while the DRAR-CPI resource gave an incorrect prediction. Therefore, our results indicate that the training sets of SwissTargetPrediction and SuperPred are far from being perfect. However, it is known that the similarity methods of evaluation of biological activity do not possess high predictive power.[54]

Seemingly, low prediction quality for the compounds of the test set given by the DRAR-CPI system can be explained by the limitations of the docking procedures (the difference of the 3D structure of the target protein in crystal and in solution, errors in binding energies evaluation) and by the imperfection of the training set (small number of the reference ligands and target proteins). It should be noted that molecular docking methods are time-consuming (prediction for a single molecule performed by PASS Online, SuperPred, or SwissTargetPrediction takes less than a minute, while DRAR-CPI needs several hours).

Based on the accuracy comparison performed in this study, a conclusion can be made that the PASS Online system uses superior machine learning methods, based on the MNA descriptors and Bayessian classifier in comparison to the similarity-based methods used in SuperPred and SwissTargetPrediction or molecular docking employed by DRAR-CPI. However, the training set and functionality of the PASS Online resource still require further development because new data on physiologically active compounds are obtained continuosly and the user demands to web resources tend to grow.

As new online resources for the prediction of biological activity and properties of organic compounds appear,[55—59] the question of integration arises, because the results obtained from multiple sources can be used to achieve higher prediction quality and extended functionality, which would be appreciated first of all by the academic researchers involved in the search and development of novel medicinal drugs.[60,61]

This work was financially supported by the Fundamental Research Program for 2013—2020 of the Russian Academy of Sciences.

References

1. ChemNavigator URL, http://www.chemnavigator.com.
2. L. Ruddigkeit, L. C. Blum, J.-L. Reymond, *J. Chem. Inf. Model.*, 2013, **53**, 56.
3. N. S. Zevirov, O. N. Zefirova, *Moscow Univ. Chem. Bull., Ser. Khim.*, 2000, **41**.
4. G. Kubinji, *Mendeleev Chem. J.*, 2006, **50**, 5.
5. C. Hansch, P. Maloney, T. Fujita, *Nature*, 1962, **194**, 178.
6. C. Hancsh, T. Fujita, *J. Am. Chem. Soc.*, 1964, **86**, 1616.
7. S. M. Free, J. W. Wilson, *J. Med. Chem.*, 1964, **7**, 395.
8. L. B. Kier, *Mol. Pharmacol.*, 1967, **3**, 487.
9. http://elibrary.ru/author_items.asp?authorid=5599.
10. S. S. Trach, N. S. Zefirov, *J. Org. Chem. USSR* (*Engl. Transl.*), 1982, **18**.

11. S. S. Trach, E. V. Gordeeva, N. S. Zefirov, *Moscow Univ. Chem. Bull.*, *Ser. Khim.*, 1984 (*Engl. Transl.*), **24**.
12. N. S. Zefirov, E. V. Gordeeva, S. S. Tratch, *J. Chem. Inform. Comput. Sci.*, 1988, **28**, 188.
13. D. E. Lushnikov, N. S. Zefirov, *J. Chem. Inform. Comput. Sci.*, 1992, **32**, 317.
14. E. V. Gordeeva, A. R. Katritzky, V. V. Shcherbukhin, N. S. Zefirov, *J. Chem. Inform. Comput. Sci.*, 1993, **33**, 102.
15. M. I. Skvortsova, I. I. Baskin, O. L. Slovokhotova, V. A. Palyulin, N. S. Zefirov, *J. Chem. Inform. Comput. Sci.*, 1993, **33,** 630.
16. D. B. Kireev, V. I. Fetisov, N. S. Zefirov, *J. Mol. Struct.*, 1994, **304**, 143.
17. I. I. Baskin, M. I. Skvortsova, I. V. Stankevich, N. S. Zefirov, *J. Chem. Inform. Comput. Sci.*, 1995, **35,** 527.
18. S. S. Tratch, N. S. Zefirov, *J. Chem. Inform. Comput. Sci.*, 1995, **36**, 448.
19. I. I. Baskin, V. A. Palyulin, N. S. Zefirov, *J. Chem. Inform. Comput. Sci.*, 1997, **37**, 715.
20. V. A. Palyulin, E. V. Radchenko, N. S. Zefirov, *J. Chem. Inform. Comput. Sci.*, 2000, **40**, 659.
21. A. A. Melnikov, V. A. Palyulin, N. S. Zefirov, *J. Chem. Inform. Comput. Sci.*, 2007, **47**, 2077.
22. D. I. Osolodkin, V. I. Chupakhin, V. A. Palyulin, N. S. Zefirov, *J. Mol. Graph. Model.*, 2009, **27**, 813.
23. E. V. Radchenko, A. A. Melnikov, G. F. Mahaeva, V. A. Palyulin, N. S. Zefirov, *Dokl. Chem.* (*Engl. Transl.*), 2012, **443**.
24. E. V. Dueva, D. I. Osolodkin, L. I. Kozlovskaya, V. A. Palyulin, V. M. Pentkovski, N. S. Zefirov, *Molecular Informatics*, 2014, **33**, 695.
25. E. V. Radchenko, D. S. Karlov, V. A. Palyulin, N. S. Zefirov, *Dokl. Chem.* (*Engl. Transl.*), 2014, **454**.
26. Yu. V. Burov, L. V. Korolchenko, V. V. Poroikov, *Bull. USSR Bioactive Compound Security Center* (*Engl. Transl.*), 1990, № 1, 4.
27. J.-U. Peters, *J. Med. Chem.*, 2013, **56**, 8955.
28. V. V. Avidon, *Pharm. Chem. J.* (*Engl. Transl.*), 1974, **8**.
29. V. E. Golender, A. B. Rosenblit, *Avtomatika i Telemehanika*, 1974, **11**, 99 (in Russian).
30. V. V. Avidon, V. S. Arolovich, S. P. Kozlova, L. A. Piruzyan, *Pharm. Chem. J.* (*Engl, Transl.*), 1978, **12**.
31. V. E. Golender, A. B. Rosenblit, *Computational Methods in Drug Design*, Zinatne, Riga, 1978, 232 pp. (in Russian).
32. V. V. Poroikov, in *Medicinal biophysics. Biological Screening of Chemical Compounds.* Medicine, Moscow, 2005, 546 (in Russian).
33. V. V. Poroikov, D. A. Filimonov, A. P. Budunova, *Nauchno-Tehn. Inform.*, 2, 1993, № 6, 11 (in Russian).
34. D. A. Filimonov, V. V. Poroikov, E. I. Karaichava, R. K. Kazaryan, A. P. Budunova, E. M. Mihailovskii, A. V. Rudnitskih, L. V. Goncharenko, Yu. V. Burov, *Experim. Clin. Pharmac.*, 1995, **58**, 56 (in Russian).
35. D. A. Filimonov, A. A. Lagunin, T. A. Gloriozova, A. V. Rudik, D. S. Druzhilovskii, P. V. Pogodin, V. V. Poroikov, *Chem. Heterocycl. Compd.*, 2014, **50**, 444.
36. A. Lagunin, A. Stepanchikova, D. Filimonov, V. Poroikov, *Bioinformatics*, 2000, **16**, 747.
37. A. Sadym, A. Lagunin, D. Filimonov, V. Poroikov, *SAR & QSAR Environ. Res.*, 2003, **14**, 339.
38. A. Geronikaki, D. Druzhilovsky, A. Zakharov, V. Poroikov, *SAR & QSAR Environ. Res.*, 2008, **19**, 27.
39. PASS Online URL, http://www.way2drug.com/passonline.
40. PubChem URL, https://pubchem.ncbi.nlm.nih.gov.
41. ChEMBL URL, https://www.ebi.ac.uk/chembl.
42. DrugBank URL, http://www.drugbank.ca.
43. SuperPred URL, http://prediction.charite.de.
44. J. Nickel1, B.-O. Gohlke, J. Erehman, P. Banerjee, W. W. Rong, A. Goede, M. Dunkel1, R. Preissner, *Nucl. Acids Res.*, 2014, **42**, W26.
45. SwissTargetPrediction URL, http://www.swisstargetprediction.ch.
46. D. Gfeller, A. Grosdidier, M. Wirth, A. Daina, O. Michielin, V. Zoete, *Nucl. Acids Res.*, 2014, **42**, W32.
47. DRAR-CPI URL, http://cpi.bio-x.cn/drar.
48. H. Luo, J. Chen, L. Shi, M. Mikailov, H. Zhu, K. Wang, L. He, L. Yang, *Nucl. Acids Res.*, 2011, **39**, W492.
49. A. Mullard, *Nat. Rev. Drug Discov.*, 2015, **14**, 77.
50. Thomson Reuters Integrity URL, https://integrity.thomsonpharma.com.
51. D. A. Filimonov, V. V. Poroikov, in *Chemoinformatics Approaches to Virtual Screening*, RSC Publishing, Cambridge, 2008, 182.
52. *Chemoinformatics*, Eds J. Gasteiger, T. Engel, Wiley-VCH, Weinheim, 2006, 680 pp.
53. ChemAxon URL, http://www.chemaxon.com.
54. Y. Martin, J. I. Kofron, L. M. Traphagen, *J. Med. Chem.*, 2002, **45**, 4350.
55. A. Lagunin, S. Ivanov, A. Rudik, D. Poroikov, *Bioinformatics*, 2013, **29**, 2062.
56. L. Wang, C. Ma, P. Wipf, H. Liu, W. Su, X.-Q. Xie, *AAPS J.*, 2013, **15**, 395.
57. P. V. Pogodin, A. A. Lagunin, S. M. Ivanov, V. I. Konova, D. A. Filimonov, V. V. Poroikov, *Rus. Med. Univ. Bull.*, 2014, № 4, 69 (in Russian).
58. A. V. Rudik, A. V. Dmitriev, A. A. Lagunin, D. A. Filimonov, V. V. Poroikov, *J. Chem. Inform. Model.*, 2014, **54**, 498.
59. E. V. Radchenko, P. V. Karpov, S. B. Sosnin, V. A. Palyulin, N. S. Zefirov, in *XXII Russian National Congress. «Humanity and Medicines»*, Moscow, 2015, p. 250 (in Russian).
60. J. Frearson, P. Wyatt, *Expert Opin. Drug Discov.*, 2010, **5**, 909.
61. D. M. Huryn, L. O. Resnick, P. Wipf, *J. Med. Chem.*, 2013, **56**, 7161.

Recieved June 16, 2015;
in revised form October 4, 2015