

Discovery of new pharmaceutical agents using in silico approaches: PASS, GUSAR, PharmaExpert and Way2Drug platform”

Vladimir Poroikov

Institute of Biomedical Chemistry, Moscow, Russia

- [IBMC Home](#)
- [Administration](#)
- [Departments](#)
- [Publications](#)
- [Contacts](#)
- [Projects](#)

- [Search](#)
- [IBMC Links](#)

- [News and announcements](#)



Proteomics & Mass-spectrometry

Towards the human proteome



Personalized medicine

Digital "-omics" technologies: prevention and early diagnosis of socially significant diseases



Nanotechnology

Improving the efficiency of existing drugs



Bioinformatics

For identify biomarkers, pharmacological targets and the basic structures of new drugs



Biochemistry

Enzymes in pathological processes, biologically active compounds and vaccines

HUMAN PROTEOME PROJECT



Department for Bioinformatics

The Department of Bioinformatics carries out the investigations related to bioinformatics and computer-aided drug design & discovery. The skilled personnel, original and commercially available software & databases, and [high performance computation facilities](#) provide all pre-requisites for basic and applied research, covered the whole field "From genomes to drugs in silico".

Department for Bioinformatics currently includes five laboratories:

- [Laboratory of Postgenomic Data Analysis](#) (Head - Dr.Sc. Elena Ponomarenko),
- [Laboratory of Bioinformational Technologies](#) (Head - Dr.Sc. Andrey Lisitsa),
- [Laboratory of Molecular Graphics and Drug Design](#) (Head - Dr.Sc. Alexander Veselovsky),
- [Laboratory of Structure-Function Based Drug Design](#) (Head - Prof. Dr. Vladimir Poroikov),
- [Laboratory of Parallel Computing and Informational Technology](#) (Head - Ph.D. Vladlen Skvortsov).

The associates are taking active participation in teaching students (special courses and practicums) from Medical-Biological Faculty of Russian State Medical University and some other Moscow Universities.

Overview

1. Historical reflections
2. PASS (Prediction of Activity Spectra for Substances)
3. PharmaExpert
4. GUSAR (General Unrestricted Structure-Activity Relationships)
5. Some examples of applications
6. Our web-services based on PASS, GUSAR, etc.
7. Way2Drug.com: further progress

Overview

1. Historical reflections
2. PASS (Prediction of Activity Spectra for Substances)
3. PharmaExpert
4. GUSAR (General Unrestricted Structure-Activity Relationships)
5. Some examples of applications
6. Our web-services based on PASS, GUSAR, etc.
7. Way2Drug.com: further progress

System for Registration and Biological Testing of New Chemical Compounds

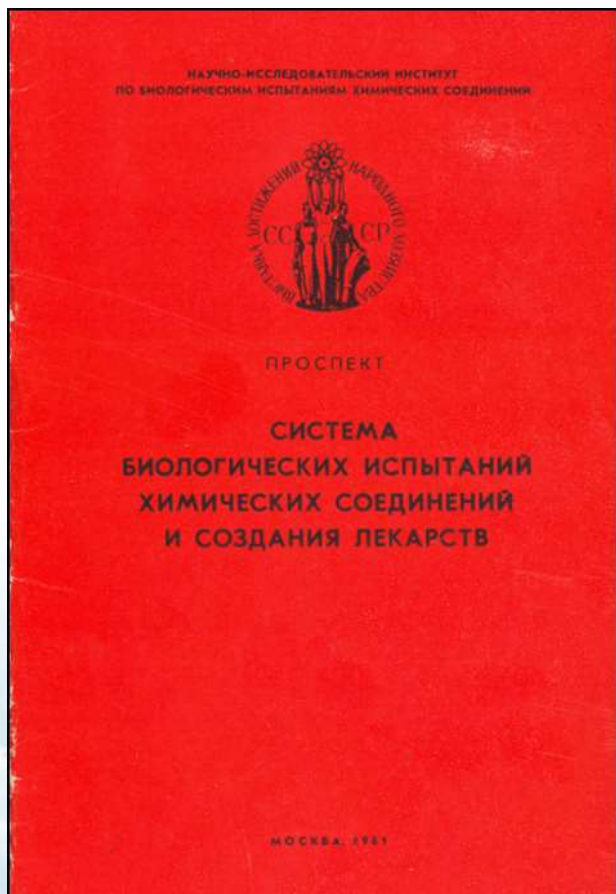
“The system of biological tests of physiologically active substances developed under the supervision of Piruzyan allows the use of relatively simple methods of investigation requiring minimal expenditures of time and substances to eliminate substances that are of little promise for pharmacology from a great number of new compounds and select those that may become medicinal drugs”.



Lev Aramovich Piruzyan (On his 70th birthday)

<http://link.springer.com/article/10.1134/S0362119707050192>

Some historical reminiscences



Медицинская библиотека. Биологические испытания химических соединений. М.: Издательство, 2005, № 1, стр.546-548

СИСТЕМА ГОСУДАРСТВЕННОЙ РЕГИСТРАЦИИ И БИОЛОГИЧЕСКИХ ИСПЫТАНИЙ ХИМИЧЕСКИХ СОЕДИНЕНИЙ: ВОСПОМИНАНИЯ О БУДУЩЕМ

В.В. Порошков

Около тридцати лет тому назад в Советском Союзе была создана Государственная система регистрации химических соединений.

В двадцати пяти километрах от Москвы, в поселке Старая Купавна, появился Научно-исследовательский институт по биологическим испытаниям химических соединений, который возглавил член-корреспондент Академии наук СССР Лев Арабаков Парузин. Основным заданием Института стали: регистрация всех синтезируемых в СССР и выделенных из природных источников химических соединений, организация и проведение их биологических испытаний. Эта деятельность планомерно осуществлялась почти двадцать лет, вплоть до распада Советского Союза в 1991 году. Впрочем и после этого в Институт, ставший проприетарным известным всей стране НИИЦ БИХС и носивший имя ВНИЦ БИВ, еще много лет поступали регистрационные карты на химические соединения, синтезированные во многих организациях бывшего СССР.

Задача государственной регистрации и биологических испытаний химических соединений, столь актуальная в семидесятые годы, не утратила своей значимости и до сих пор. Но какие обстоятельства препятствовали эффективной работе системы государственной регистрации? Почему многое из задуманного Л.А.Парузиным осуществилось в полной мере лишь спустя десятилетия и, в основном, за рубежом? Ответы на эти непростые вопросы мы попытаемся поискать в данном очерке.

Естественно, что все выделенное отражает субъективное мнение автора и ни в коей мере не является "истинной и последней истиной". Более того, размышления на данную тему с неизбежностью носят ретроспективный характер и, уже только потому, могут восприниматься с известной долей скепсиса. И хотя "история учит только тому, что ничему не учит", я надеюсь, что уроки создания и разрушения системы Государственной регистрации химических соединений

© В.В. Порошков, 2001

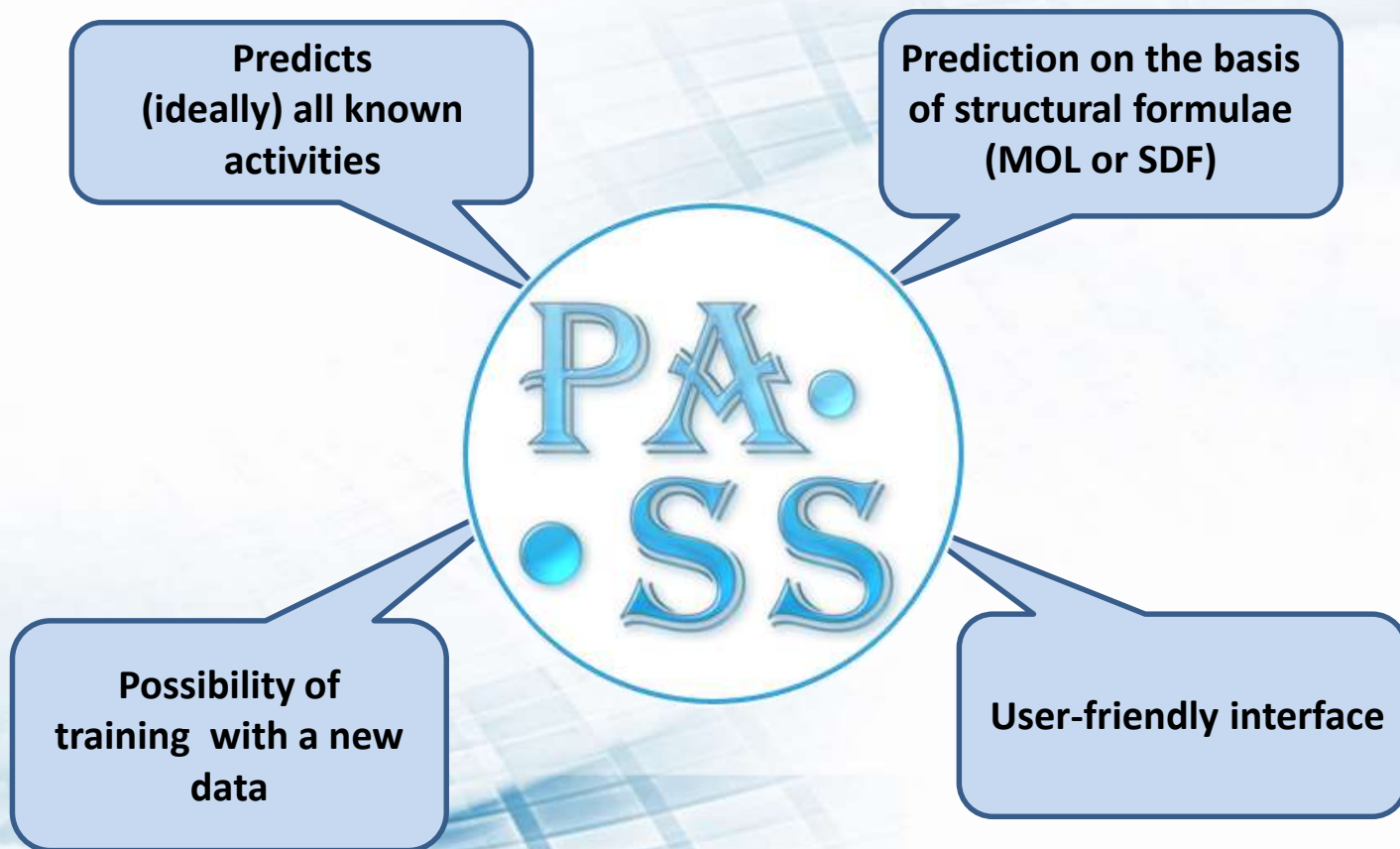
PASS History: Persistent updating and improvement

- 1972 *Collection of the training set started (USSR National System of New Chemical Compounds Registration).*
- 1976-1992 *Early versions of different computer programs for biological activity spectra prediction (V.A. Avidon, V.E. Golender & A.B. Rosenblit).*
- 1993 First version of PASS: 9,314 compounds; 114 activities, accuracy of prediction AP=76%.
- 1998 PASS C&T version 4.0: 30,537 compounds; 541 activities, AP=82%.
- 2005 PASS Pro 2005: >60,000 compounds; >2500 activities, AP=89%.
- 2009 PASS Pro v. 9.1: >200,000 compounds; >3500 activities, AP =95%.
- 2014 PASS Pro 2014: >950,000 compounds; >7000 activities, AP =95%.**

Overview

1. Historical reflections
2. **PASS (Prediction of Activity Spectra for Substances)**
3. PharmaExpert
4. GUSAR (General Unrestricted Structure-Activity Relationships)
5. Some examples of applications
6. Our web-services based on PASS, GUSAR, etc.
7. **Way2Drug.com: further progress**

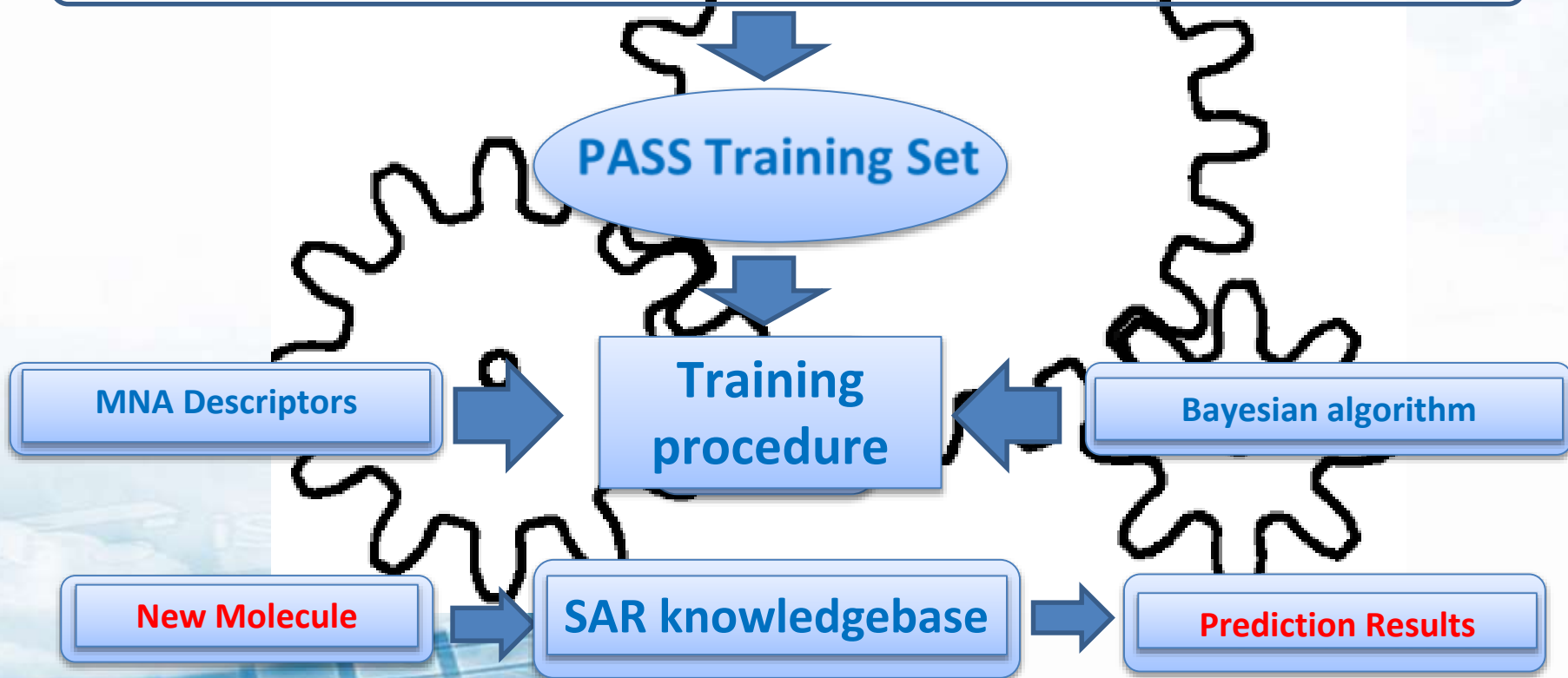
Requirements for a computer program evaluated biological activity profiles (spectra)



Computer program PASS: Prediction of Activity Spectra for Substances

Full text publications, databases, presentations at conferences etc.

Reliable data on structure and activity of drug-like molecules



PASS 2014 Characteristics

Training Set	959,801 drugs, drug-candidates, pharmacological and toxic substances comprise the training set
Biological Activity	7,158 biological activities can be predicted (Active vs. Inactive)
Chemical Structure	Multilevel Neighborhoods of Atoms (MNA) descriptors [1, 2]
Mathematical Algorithm	Bayesian approach was selected by comparison of many different methods [2]
Validation	Average accuracy of prediction in LOO CV for the whole training set is ~95% [2]; robustness was shown using principal compounds from MDDR database [3]

1. Filimonov D.A. et al. *J. Chem. Inform. Computer Sci.*, 1999, 39: 666-670.
2. Filimonov D.A., Poroikov V.V. *Chemoinformatics Approaches to Virtual Screening*, 2008, 182-216.
3. Poroikov V.V. et al. *J. Chem. Inform. Computer Sci.*, 2000, 40: 1349-1355.



Concept of biological activity spectrum

Biological Activity Spectrum is the intrinsic property of the compound reflected all biological activities, which can be found in the compound's interaction with biological entity.

Poroikov V.V., Filimonov D.A., Boudunova A.P. *Automatic Documentation and Mathematical Linguistics*. Allerton Press Inc., 1993, 27: 40-43.

Filimonov D.A., Poroikov V.V., Karaicheva E.I. et. al. *Experimental and Clinical Pharmacology*, 1995, 58: 56-62 (Rus).

Filimonov D.A., Poroikov V.V. In: *Bioactive Compound Design: Possibilities for Industrial Use*, BIOS Scientific Publishers, Oxford (UK), 1996. pp. 47-56.

Non-synonymous definitions in literature

Lewi P.J. Spectral mapping, a technique for classifying **biological activity profiles** of chemical compounds. *Arzneimittelforschung*. 1976; **26** (7):1295-1300.

Battistini A. et al. **Spectrum of biological activity** of interferons. *Annali dell'Istituto Superiore di Sanità*. 1990; **26** (3-4):227-253.

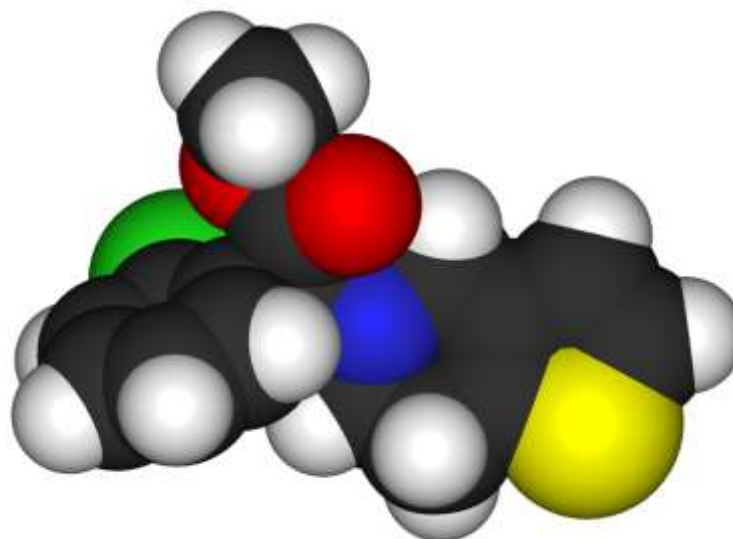
Gringorten J.L. et al. **Activity spectra** of Bacillus thuringiensis delta-endotoxins against eight insect cell lines. *In Vitro Cell. Dev. Biol. Anim.* 1999; **35** (5):299-303.

Fleri A.F. et al. **Biological spectra** analysis: Linking **biological activity profiles** to molecular structure *Proc. Natl. Acad. Sci. USA*. 2005; **102** (2): 261-266.

Rana A. Benzothiazoles: A new **profile of biological activities**. *Indian J. Pharm. Sci.* 2007; **69**:10-17.

Fedichev P., Vinnik A. **Biological Spectra** Analysis: Linking **Biological Activity Profiles** to Molecular Toxicity. 2007; <http://www.q-pharm.com>.

Chemical Structure Representation

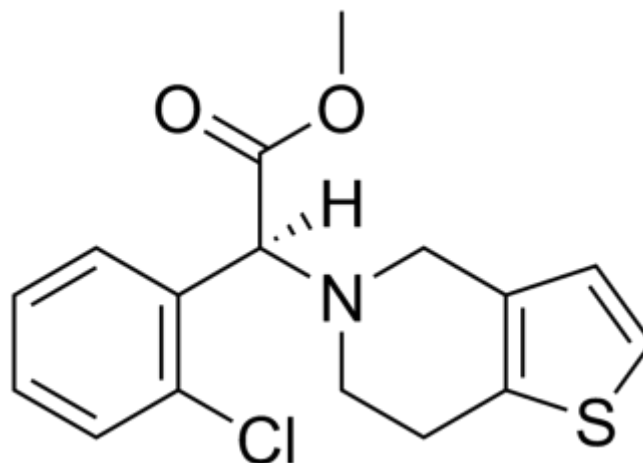


Spatial configuration of a free uncharged molecule in the ground state in vacuum is the necessary and sufficient description of its structure.

To use this molecular structure description one needs the substantial computational resources for molecular modeling and/or quantum-chemical calculations.

However, the basis of all calculations is the traditional structural formula.

Chemical structure representation



The structural formula unambiguously determines all properties of the organic molecule.

Environment? – Structural formula determines at least potential, “intrinsic” properties.

Neighborhoods of atoms descriptors

The most biological activities of organic compounds are the result of molecular recognition, which in turn depends on the correspondence between the particular atoms of the ligand and the target.

MOLECULAR BIOLOGY
QUANTUM CHEMISTRY
QUANTUM FIELD THEORY

$$M = V + VgM = V + VgV + VgVgV + VgVgVg + \dots$$

$$M_i = V_i + V_i g M = V_i + V_i g (M_1 + M_2 + \dots + M_m)$$

Descriptors are based on the concept of atoms' of molecule taking into account the influence of the neighborhoods:

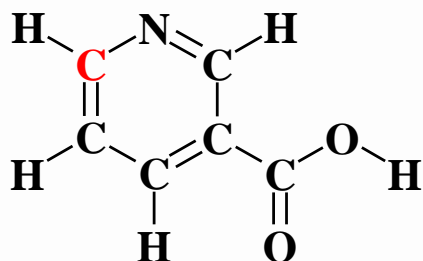
MNA - multilevel neighborhoods of atoms

QNA - quantitative neighborhoods of atoms

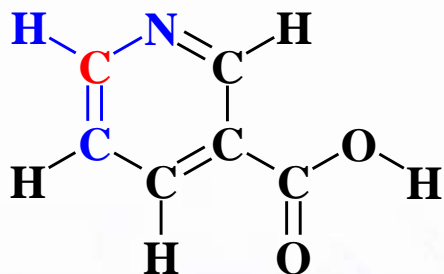
Filimonov D.A., Poroikov V.V. Probabilistic approach in activity prediction. *In: Chemoinformatics Approaches to Virtual Screening*. Eds. Alexandre Varnek and Alexander Tropsha. Cambridge (UK): RSC Publishing, 2008, 182-216.

Filimonov D.A., Zakharov A.V., Lagunin A.A., Poroikov V.V. QNA based 'Star Track' QSAR approach. *SAR and QSAR Environ. Res.*, 2009, 20 (7-8), 679-709.

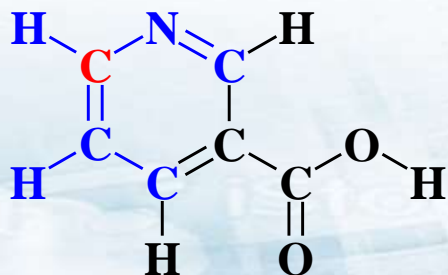
MNA: Multilevel Neighborhoods of Atoms



MNA/0: C



MNA/1: C(CN-H)

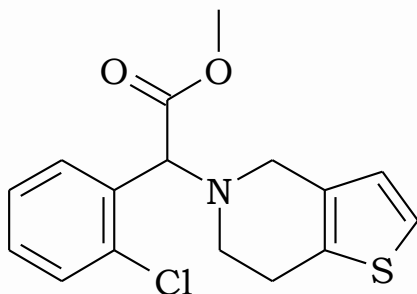


MNA/2: C(C(CC-H)N(CC)-H(C))

Filimonov D.A. et al. *J. Chem. Inform. Computer Sci.*, 1999, **39**: 666-670.

Substance representation

Structural formula of Clopidogrel



Activity Spectrum of Clopidogrel

Abdominal pain
Acute neurologic disorders treatment
Agranulocytosis
Allergic reaction
Anaphylaxis
Anemia
Angioedema
Angiogenesis inhibitor
Antianginal
Antiarthritic
Anticoagulant
Antineoplastic
Antipsoriatic
Antithrombotic

...

112 known activities in PASS SAR Base

MNA Descriptors of Clopidogrel

HC	C(C(CCC)C(CC-H-H)S(CC))
CHHHO	C(C(CCC)C(CS-H)-H(C))
CHHCC	C(C(CCC)N(CC-C)-H(C)-H(C))
CHHCN	C(C(CCS)C(CC-H)C(CN-H-H))
CHCC	C(C(CCS)C(CN-H-H)-H(C)-H(C))
CHCCN	C(C(CC-H-H)N(CC-C)-H(C)-H(C))
CHCS	C(C(CC-H)C(CC-H)-H(C))
CCCC	C(C(CC-H)C(CC-C)-H(C))
CCCS	C(C(CC-H)C(CC-C)-Cl(C))
CCCCI	C(C(CC-H)C(CC-Cl)-H(C))
CCOO	C(C(CC-H)C(CC-Cl)-C(CN-H-C))
NCCC	C(C(CC-H)S(CC)-H(C))
OC	N(C(CN-H-H)C(CN-H-H)-C(CN-H-C))
OCC	S(C(CCS)C(CS-H))
SCC	-H(C(CC-H))
CIC	-H(C(CC-H-H))
	-H(C(CN-H-H))
	-H(C(CS-H))
	-H(-C(CN-H-C))
	-H(-C(-H-H-H-O))
	-C(C(CC-C)N(CC-C)-H(-C)-C(-C-O-O))
	-C(-H(-C)-H(-C)-H(-C)-O(-C-C))
	-C(-C(CN-H-C)-O(-C)-O(-C-C))
	-O(-C(-H-H-H-O)-C(-C-O-O))
	-O(-C(-C-O-O))
	-Cl(C(CC-Cl))

Prediction of Biological Activity Spectra

According to the Bayes' theorem, the probability $P(A|S)$ that the compound S has activity (or inactivity) A , equals to:

$$P(A|S) = P(S|A) \cdot P(A) / P(S)$$

If the descriptors of organic compound D_1, \dots, D_m are independent, then:

$$P(S|A) = P(D_1, \dots, D_m|A) = \prod_i P(D_i|A)$$

$P(A)$ and $P(A|D_i)$ are calculated as sums through all compounds of the training set:

$$P(A | D_i) = \frac{\sum_k g_k(D_i) w_k(A)}{\sum_k g_k(D_i)}$$

$$P(A) = \frac{\sum_i \sum_k g_k(D_i) w_k(A)}{\sum_i \sum_k g_k(D_i)}$$

Biological activities predicted by PASS

- **Pharmacotherapeutic effects**
(antihypertensive, hepatoprotective, antiinflammatory etc.);
- **Mechanisms of action**
(5-HT_{1A} agonist, cyclooxygenase 1 inhibitor, adenosine uptake inhibitor, etc.);
- **Specific toxicities**
(mutagenicity, carcinogenicity, teratogenicity, etc.);
- **Interaction with Antitargets**
(HERG channel blocker, etc.);
- **Metabolic terms**
(CYP1A substrate, CYP3A4 inhibitor, CYP2C9 inducer, etc.);
- **Influence on gene expression**
(APOA1 expression enhancer, NOS2 expression inhibitor, etc.);
- **Action on transporters**
(Dopamine transporter antagonist, Sodium/bile acid cotransporter inhibitor, etc.).

PASS algorithm description

Filimonov D.A., Lagunin A.A., Glorizova T.A., Rudik A.V., Druzhilovskiy D.S., Pogodin P.V., Poroikov V.V. (2014). Prediction of the biological activity spectra of organic compounds using the PASS online web resource. *Chemistry of Heterocyclic Compounds*, 50: 444-457.

Filimonov D.A., Poroikov V.V. (2008). Probabilistic Approach in Virtual Screening. In: *Chemoinformatics Approaches to Virtual Screening*. Alexander Varnek and Alexander Tropsha, Eds. RSC Publishing, 182-216.

Filimonov D.A., Poroikov V.V. (2006). Prediction of biological activity spectra for organic compounds. *Russian Journal of General Chemistry*, 50: 66-75.



Descriptors (1999)



Robustness (2000)



PASS Online (2000)

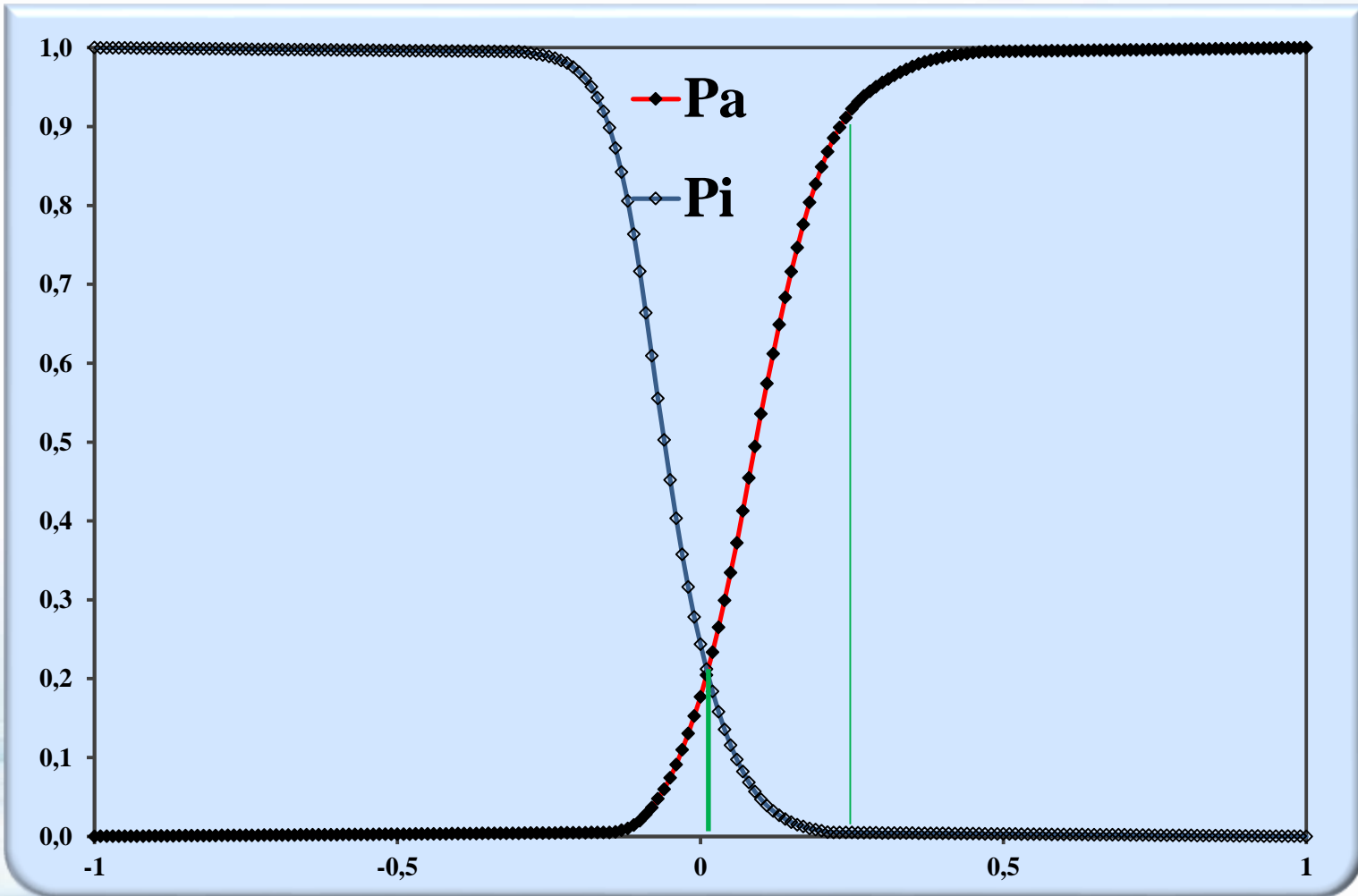


Drug-likeness (2001)



NCI Browser (2003)

Threshold selection: active/inactive?



PASS: Selection of the desirable activities

The screenshot shows the PASS software interface. The 'Base' menu is open, and the 'Selection ...' option is highlighted. The main window displays the file path 'C:\V\ACTIVITY-SELECTION\PASS2014-Reumatoid-Arthritis.SAR'. At the bottom, a table lists 'Unused Activity Type' with columns for 'Group', 'Number', and 'IAP'.

Unused Activity Type	Group	Number	IAP
(N-acetylneuraminyl)-galactosylglucosylceramide N-acetylgalactosaminyltransferase inhibitor	M	6	0.9714
(R)-3-amino-2-methylpropionate-pyruvate transaminase inhibitor	M	24	0.9980
(R)-6-hydroxynicotine oxidase inhibitor	M	3	0.9435

Buttons: Include ..., Load ..., Save ..., Ok, Cancel

Selected Activity Types: 0 of 7474 Av. IAP: 0.0000

Rheumatoid arthritis treatment: PASS activities

5-Lipoxygenase inhibitor
Adenosine A3 receptor agonist
Adenosine deaminase inhibitor
Angiogenesis inhibitor
Antiinflammatory
Antimetabolite
Antioxidant
Apoptosis agonist
Autoimmune disorders treatment
Beta amyloid protein antagonist
Bisphosphonate
Bruton tyrosine kinase inhibitor
Calcineurin inhibitor
Cannabinoid receptor agonist
Cathepsin K inhibitor
CC chemokine 1 receptor antagonist
CC chemokine 2 receptor antagonist

Cell adhesion inhibitor
Cell adhesion molecule inhibitor
Chemokine receptor antagonist
Collagenase inhibitor
Complement inhibitor
Corticosteroid-like
Corticotropin releasing factor antagonist
CXC chemokine 2 receptor antagonist
CXC chemokine 4 receptor antagonist
Cyclin-dependent kinase 1 inhibitor
Cyclin-dependent kinase 2 inhibitor
Cyclin-dependent kinase 5 inhibitor
Cyclin-dependent kinase 7 inhibitor
Cyclin-dependent kinase 9 inhibitor
Cyclooxygenase 2 inhibitor
...
VLA-4 antagonist

Totally: 110 activities

Selection of activities for RA treatment

Select Activity Types to be Predicted

Predictable Activity Type	Group	Number	IAP
5-Lipoxygenase inhibitor	M	2834	0.9811
Adenosine A3 receptor agonist	M	163	0.9944
Adenosine deaminase inhibitor	MA	225	0.9903
Angiogenesis inhibitor	EM	4489	0.9223
Antiinflammatory	E	5983	0.8471
Antimetabolite	EM	732	0.9908
Antioxidant	EM	843	0.9494
Apoptosis agonist	E	1299	0.8877
Autoimmune disorders treatment	E	3972	0.8875
Beta amyloid protein antagonist	M	122	0.9805

Unused Activity Type	Group	Number	IAP
(N-acetylneuraminyl)-galactosylglucosylceramide N-acetylgalactosaminyltransferase inhibitor	M	6	0.9714
(R)-3-amino-2-methylpropionate-pyruvate transaminase inhibitor	M	24	0.9980
(R)-6-hydroxynicotine oxidase inhibitor	M	3	0.9435
(R)-Pantolactone dehydrogenase (flavin) inhibitor	M	8	0.8780
(R)-aminopropanol dehydrogenase inhibitor	M	10	0.9939
(R)-limonene 6-monooxygenase inhibitor	M	3	0.9980
(R,R)-butanediol dehydrogenase inhibitor	M	3	0.9966
(S)-2-Methylmalate dehydratase inhibitor	M	4	0.9980
(S)-2-hydroxy-acid oxidase inhibitor	M	28	0.9336
(S)-3-amino-2-methylpropionate transaminase inhibitor	M	8	0.9607

Include ... Load ... Save ... Ok Cancel

Selected Activity Types: 110 of 7474 Av. IAP: 0.9697

PASS Pro: Creating new SAR Base

The screenshot shows the PASS software interface. The 'File' menu is open, and 'New Base' is selected. A dialog box titled 'SAR Base Information' is displayed, showing the following data:

Substances	0
Descriptors	0
Activity Types	0
Selected Activity Types	0
Average IAP	
Prediction	Disabled

The screenshot shows the PASS software interface after adding data. The 'File' menu is open, and 'Add ...' is selected. A dialog box titled 'SAR Base Information' is displayed, showing the following data:

Substances	100	Modified
Descriptors	1092	Modified
Activity Types	54	Modified
Selected Activity Types	0	
Average IAP		
Prediction		Disabled

PASS Pro: Training

The screenshot shows the PASS Pro software interface. The main window has a menu bar with 'File', 'Base', 'Predict', 'View', 'Options', and 'Help'. The 'File' menu is open, showing options like 'New Base', 'Open Base ...', 'Save Base', 'Save Base As ...', 'Add ...', 'Exclude ...', 'Training ...', 'Selection ...', and 'Validation ...'. The 'Training ...' option is highlighted.

Two dialog boxes are open:

- SAR Base Information:**

Substances	100	Modified
Descriptors	1092	Modified
Activity Types	54	Modified
Selected Activity Types	54	Modified
Average IAP	0,8938	
Prediction		Disabled
- Select Activity Types to be Predicted:**

Predictable Activity Type	Group	Number	IAP
Teratogen	T	40	0,6263
Antineoplastic	E	13	0,7073
Immunosuppressant	EM	10	0,7178
Antipruritic, allergic	E	5	0,7621
Convulsant	T	6	0,7695
Carcinogenic, rat, male	T	6	0,7695
Cardiotonic	EM	6	0,7801
Phosphodiesterase inhibitor	MA	5	0,7958
Antifungal	E	5	0,8042
Vasodilator	EM	6	0,8138

Unused Activity Type	Group	Number	IAP

Buttons: Include ..., Load ..., Save ..., Ok, Cancel

Selected Activity Types: 54 of 54 Av. IAP: 0,8938

PASS Pro: Selection procedure

Select Activity Types to be Predicted

Predictable Activity Type	Group	Number	IAP
Nucleotide metabolism regulator	M	4	0,8958
Antipruritic, non-allergic	E	3	0,9003
Psychotropic	E	11	0,9019
HERG channel blocker	A	5	0,9179
Potassium channel (Voltage-sensitive) blocker	M	5	0,9179
Antibacterial	E	15	0,9184
Atherosclerosis treatment	E	5	0,9242
QT interval prolongation	T	3	0,9244
Diuretic	EM	10	0,9244
Antiallergic	E	18	0,9248

Unused Activity Type	Group	Number	IAP
Spasmolytic	E	8	0,8859
Cyclic AMP phosphodiesterase inhibitor	M	4	0,8958

Selected Activity Types: 32 of 54 Av. IAP: 0,9512

PASS Pro: Ready for Prediction

The screenshot shows the PASS software window with the following details:

- Window Title: PASS
- Menu Bar: File, Base, Predict, View, Options, Help
- Toolbar: Includes icons for file operations and a dropdown menu showing "Pa > Pi".
- Current Base: PASS_SAR_Base1
- Dialog Box: SAR Base Information

Substances	100	Modified
Descriptors	1092	Modified
Activity Types	54	Modified
Selected Activity Types	32	Modified
Average IAP	0.9512	
Prediction		<input type="checkbox"/> Enabled

PASS Pro: 20-fold validation

The screenshot shows the PASS software interface with the 'Validation ...' option selected in the 'File' menu. A 'SAR Base Validation' dialog box is open, displaying a table of validation results for 32 activity types. The table includes columns for ID, status, mechanism, count, and two IAP values, along with the activity name. Summary statistics at the bottom indicate 32 activity types were successfully validated, with an average IAP of 0.9512 and a 20-fold average IAP of 0.9200.

ID	Status	Mechanism	Count	IAP 1	IAP 2	Activity Name
10	Ok	E	5	0.9242	0.9179	Atherosclerosis treatment
11	Ok	Z	3	0.9450	0.8557	CYP2D6 substrate
12	Ok	MA	3	0.9622	0.9622	Carbonic anhydrase inhibitor
13	Ok	EM	5	0.9979	0.9958	Cell wall synthesis inhibitor
14	Ok	M	3	0.9313	0.9003	Cholinergic
15	Ok	M	3	0.9381	0.9347	Cortisol sulfotransferase inhibitor
16	Ok	MA	6	0.9415	0.9450	Cyclooxygenase inhibitor
17	Ok	EM	10	0.9244	0.9111	Diuretic
18	Ok	EM	5	0.9874	0.9895	Diuretic inhibitor
19	Ok	MA	4	0.9766	0.9740	Dopamine D2 antagonist
20	Ok	M	5	0.9937	0.9937	Dopamine antagonist
21	Ok	M	5	0.9874	0.9895	Electrolyte absorption antagonist
22	Ok	MA	15	0.9922	0.9922	Glucocorticoid agonist
23	Ok	A	5	0.9179	0.9158	HERG channel blocker
24	Ok	M	6	0.9362	0.9344	Lipocortins synthesis agonist
25	Ok	M	3	0.9759	0.9759	Microtubule formation inhibitor
26	Ok	EM	5	0.9853	0.9853	Non-steroidal antiinflammatory agent
27	Ok	M	4	0.8958	0.5182	Nucleotide metabolism regulator
28	Ok	M	5	0.9179	0.9158	Potassium channel (Voltage-sensitive) blocker
29	Ok	E	11	0.9019	0.9009	Psychotropic
30	Ok	T	3	0.9244	0.8935	QT interval prolongation
31	Ok	EM	6	1.0000	1.0000	Saluretic
32	Ok	M	6	0.9255	0.5160	Vitamin

End : 14.01.2016 6:47:14
 Validation SAR Base time: 0:00:00.950

32 Selected Activity Types
 32 Validated Activity Types
 32 Successfully Validated Activity Types
 0.9512 Average IAP
 0.9200 20-Fold Average IAP

PASS Pro: Quality Control

The screenshot displays the PASS Pro software interface. The main window is titled "PASS" and has a menu bar with "File", "Base", "Predict", "View", "Options", and "Help". The "File" menu is open, showing options like "New Base", "Open Base...", "Save Base", "Save Base As...", "Add...", "Exclude...", "Training...", "Selection...", and "Validation...".

A secondary window titled "SAR Base History" is open, showing the following text:

```

Data from C:\Program Files (x86)\2014 07 20 Standard\Samples\Errors Example\Errors_Example (P)
Activity Field : <PASS_ERROR>

Begin   : 14.01.2016 6:51:11

No 1 - Error: Number of actual atoms and bonds is not enough: 0 0
No 2 - Error: Components: 2
No 3 - Error: Carbons: 2 < 3.
No 4 - Error: Valence error.
No 5 - Error: MolCharge: 1.
No 6 - Error: Molweight: 2802.4 > 1250.0.
No 7 - Error: R atom is found.
No 8 - Error: X atom is found.
No 9 - Error: Ala atom is found.
No 10 - Error: V2000 or V3000 is not found.
No 11 - Error: V2000 or V3000 is not found.
No 12 - Error: Unexpected 'M END' in structure.
No 13 - Error: Molfile format error.
No 14 - Error: Molfile format error.
No 15 - Error: Molfile format error.
No 16 - Error: Molfile format error.
No 17 - Error: Bond error.

End     : 14.01.2016 6:51:11

Data from C:\Program Files (x86)\2014 07 20 Standard\Samples\Errors Example\Errors_Example (P)
0 of 17 Substances were added to SAR Base.
0 new Substances were added to SAR Base.
0 new Activity Data were added to SAR Base.
Adding new Data time: 0:00:00.022.
    
```


PASS predictions for Clopidogrel

C:\PASS 2012\Drugs_Example.sdf

5x5 4x4 3x3 2x2 Molecular Structure MNA

1

2

3

4

5

6

Antithrombotic

Effects Mechanisms Toxicity Antitargets Metabolism Gene Exp

45 of 464 Possible Pharmacological Effects at Pa > Pi

0.951	0.004	Neuroprotector
0.886	0.005	Acute neurologic disorders treatment
0.723	0.006	Antithrombotic
0.712	0.004	Platelet aggregation inhibitor
0.618	0.019	Antianginal
0.553	0.013	Atherosclerosis treatment
0.463	0.048	Analgesic
0.385	0.009	Platelet antagonist
0.361	0.027	Stroke treatment
0.352	0.026	Angiogenesis stimulant
0.332	0.017	Anticoagulant
0.366	0.083	Diabetic neuropathy treatment
0.292	0.013	Analgesic, opioid
0.324	0.049	Antiinflammatory, ophthalmic
0.341	0.116	Spasmolytic, urinary
0.290	0.102	Cell adhesion molecule inhibitor
0.301	0.135	Neurodegenerative diseases treatment
0.261	0.098	Antipsoriatic
0.167	0.005	Acetylcholine release stimulant
0.199	0.057	Fibromyalgia syndrome treatment
0.236	0.104	Age-related macular degeneration treatment
0.202	0.075	Pancreatic disorders treatment
0.228	0.104	Amyotrophic lateral sclerosis treatment
0.375	0.254	Vasodilator, cerebral
0.176	0.058	Lipoprotein disorders treatment
0.156	0.047	Diabetic retinopathy treatment
0.257	0.150	Psychotropic

42 Substructure Descriptors; 0 new.

246 of 6400 Possible Activities
45 of 464 Possible Pharmacological Effects
79 of 3850 Possible Mechanisms of Action
106 of 321 Possible Toxic and Adverse Effects
5 of 118 Possible Antitargets
12 of 195 Possible Metabolism-Related Actions
17 of 1610 Possible Gene Expression Regulation
4 of 68 Possible Transporters-Related Actions

> <NAME> (0)
Clopidogrel

1/129 0.723 0.006 Antithrombotic

Clopidogrel: predicted vs. known activities

Abdominal pain
Acute neurologic disorders
treatment
Agranulocytosis
Allergic reaction
Anaphylaxis
Anemia
Angioedema
Angiogenesis inhibitor
Antianginal
Antiarthritic
Anticoagulant
Antineoplastic
Antipsoriatic
Antithrombotic
Anxiety
Arthralgia
Atherosclerosis treatment
Back pain
Behavioral disturbance
Blindness
Bronchoconstrictor
Cardiotoxic
Cataract
CCL4 expression enhancer
CCL5 expression enhancer
Chest pain
Colic
Colitis

Conjunctivitis
Consciousness alteration
Constipation
Cough
CYP2 substrate
CYP2C substrate
CYP2C19 inhibitor
CYP2C19 substrate
CYP2C9 inhibitor
CYP3A substrate
CYP3A4 substrate
Cytochrome P450 inhibitor
Dermatitis
Dermatologic
Dizziness
Drug eruption
Dyspepsia
Emetic
Eosinophilia
Erythema
Erythema multiforme
Exanthema
Flatulence
GP IIb/IIIa receptor antagonist
Hallucinogen
Headache
Heart failure
Hematotoxic
Hemorrhage

Henoch-Schonlein purpura
Hepatic failure
Hepatitis
Hepatotoxic
Hypertensive
Hyperthermic
Hypotension
Infection
Insomnia
Lassitude
Leukopenia
Lichen planus
Lichenoid eruption
Malaise
Menstruation disturbance
Myalgia
Nausea
Necrosis
Nephrotoxic
Neuroprotect
Neutropenia
Ocular toxicity
Pain
Pancreatitis
Pancytopenia
Platelet aggregation inhibitor
Platelet antagonist
Pruritus
Pulmonary embolism

Purinergic P2 antagonist
Purinergic P2T antagonist
Purinergic P2Y antagonist
Purinergic P2Y12 antagonist
Purinergic receptor antagonist
Purpura
Renal colic
Reproductive dysfunction
Rhinitis
Sensory disturbance
Serum sickness
Shock
Sinusitis
Sleep disturbance
Stomatitis
Syncope
THBS1 expression enhancer
Thrombocytopenia
Toxic
Toxic epidermal necrolysis
Toxic, gastrointestinal
TP53 expression enhancer
Urticaria
Vasculitis
Vertigo
Vision disturbance

Blue – predictions coincided with the experiment.

Black – unpredictable activities. **Red** – unpredicted activities.

Overview

1. Historical reflections
2. PASS (Prediction of Activity Spectra for Substances)
3. **PharmaExpert**
4. GUSAR (General Unrestricted Structure-Activity Relationships)
5. Some examples of applications
6. Our web-services based on PASS, GUSAR, etc.
7. **Way2Drug.com: further progress**

PharmaExpert: Tool for analysis of PASS predictions

The screenshot displays the PharmaExpert interface. At the top, it shows the 'Prediction & Integration' window with a list of predicted activities for a compound (likely Celecoxib). The main window is divided into several panels:

- Substrate & Receptor:** Shows chemical structures for various targets like ACE2, Cyclooxygenase 1, and Cyclooxygenase 2.
- Activity List:** A table listing predicted activities with their predicted values. For example, 'Cyclooxygenase 1 inhibitor' has a predicted value of 0.996.
- Information Panel:** Provides detailed data for the selected activity, including:
 - Activities:** 10860
 - Synonyms:** 17367
 - Mechanisms:** 6239
 - Effects:** 707
 - Toxicity:** 306
 - Antitargets:** 131
 - Metabolites:** 343
 - Gene Expression:** 2908
 - Transporters:** 104
 - Relationships:** 12785
 - All Times:** 28185
 - Proteins:** 9829
 - KEGG pathway target pairs:** 15468
 - NO₂ pathway target pairs:** 4936
 - Reaction pathway target pairs:** 30038
 - NO₂ pathway target pairs:** 192188
- Pathway Analysis:** Shows KEGG pathways and reactions associated with the activity, such as 'Cyclooxygenase 1 inhibitor' and 'Cyclooxygenase 2 inhibitor'.
- Therapeutic Effects:** Lists potential therapeutic effects like 'Alzheimer's disease treatment' and 'Cardiac muscle contraction'.
- Side Effects:** Lists potential side effects like 'Abdominal pain' and 'Anemia'.

PharmaExpert: Statistics of predicted activities

Prediction & Interpretation - G:\DATABASES\PRESTWICK\PRESTWICK-4\prestwick_chemical_library_cured (PASS2014)-20.SDF, 1/1074

Atagurans (I) Alantoin Acetazolamide Metformin hydrochloride Isolupredone acetate Analoxide hydrochloride dihydrate Hydrochlorothiazide Sulfaguanidine Metoclopramide

Save TXT Save SD Clipboard Exclude

Pa Pi chemical_name

Statistics

Numbers descending Pa:Pi Save TXT

No.	Pa:Pi	Pa:30%	Pa:50%	Pa:70%	Types of Activity
4232	116	46	6	1	BTF3 expression inhibitor
4233	38	22	6	1	3-Isocyanophthalate dehydrogenase inhibitor
4294	38	18	6	0	Hydroxymethylglutaryl-CoA reductase inhibitor
4295	64	35	6	2	CD40 expression enhancer
4296	9	6	6	5	Opical partial agonist
4297	64	22	6	2	Trehalose-phosphatase inhibitor
4298	9	9	6	4	H ⁺ /K ⁺ -transporting ATPase inhibitor
4299	40	20	6	3	Dopachrome isomerase inhibitor
4300	26	17	6	1	Postcoital contraceptive
4301	48	18	6	1	Aconitine decarboxylase inhibitor
4302	97	53	6	1	Formaldehyde dehydrogenase inhibitor
4303	16	15	6	2	Thymidine phosphorylase inhibitor
4304	54	21	6	2	RAPGEF2 expression enhancer
4305	17	11	6	3	Gluconokinase inhibitor
4306	94	49	6	0	Pyruvate kinase inhibitor
4307	102	15	6	3	Polar-amino-acid-transporting ATPase inhibitor
4308	10	9	6	6	Pyrimidine antagonist
4309	27	17	6	1	MAD inhibitor
4310	94	24	6	2	Nitric oxide antagonist
4311	67	16	6	4	Axylamine N-acetyltransferase inhibitor
4312	40	13	6	2	Beta-N-acetylcholinesterase inhibitor
4313	30	13	5	3	Phosphoribulokinase inhibitor
4314	30	15	5	1	UGT2B4E substrate
4315	30	21	5	1	Signs receptor antagonist
4316	31	8	5	2	Oxalifene-oxo acid transaminase inhibitor
4317	31	16	5	0	Cytokine production inhibitor
4318	32	20	5	3	NAD ⁺ nucleoside inhibitor
4319	30	15	5	1	UGT2B4D substrate
4320	32	12	5	3	Chondroitin 6-sulfotransferase inhibitor
4321	40	15	5	0	Inositol monophosphatase inhibitor
4322	26	6	5	4	Narcotic antagonist
4323	26	15	5	5	CYP1B1 inhibitor
4324	23	11	5	5	Axalkylamine dehydrogenase inhibitor
4325	38	16	5	3	Nicotinamide nucleotide adenylyltransferase inhibitor
4326	39	18	5	0	Intestinal peptide-proton cotransporter inhibitor
4327	27	14	5	2	Methionine adenosyltransferase inhibitor
4328	48	17	5	0	Aromatic-amino-acid transaminase inhibitor
4329	46	13	5	3	2-Dehydroepiandrosterone reductase (beta-specific) inhibitor

Gene Ontology

- cytogenetic inhibitor
- metabolism
- abolition
- in
- abolition
- stimulant
- cytokine receptor interaction
- basic cell lineage
- inhibitory pathway
- T signaling pathway
- signaling pathway
- orlistat regulated kinase 1 inhibitor
- ne-regulated sodium reabsorption
- signaling pathway
- inhibitor
- relaxation
- inhibition
- in
- ne-regulated sodium reabsorption
- in disease
- inhibitor
- receptor signaling pathway
- inhibitor
- inhibitor (American lysosomocidal)
- in signaling pathway
- in synapse
- inhibitory leukinemia
- inhibition
- in cancer
- inhibitory cell formation
- in cancer
- inhibitory pathway
- in PI3K signaling pathway
- in PI-mediated phagocytosis
- inhibitor
- in
- inhibitory synapse
- inhibitory pathway
- B
- C
- inhibitory pathway
- A
- inhibitory pathway
- inhibitor
- in

PharmaExpert: Search for multitargeted antineoplastic agents

The screenshot displays the PharmaExpert software interface. At the top, there are navigation menus (File, Tools, View, Help) and a search bar with the query "Pa: 0.100". Below this, a row of chemical structures is shown, including Aspirin, Alantoin, Acetazolamid, Metformin hydrochloride, Isufipredone acetate, Amide hydrochloride diphosphate Hydrochlorothiazide, Sulfiguanidine, Metcram, and Benzonalate.

The main window is divided into two panes. The left pane shows a list of activities with columns for Pa, Pi, and Types of Activities. The right pane, titled "Multitargeted actions", shows a list of effects and a table of results.

No	Pa	Number	Activity type	Activity type	Activity type
462	0.632	2	Cyclic GMP phosphodiesterase inhibitor	Epidermal growth factor receptor kinase inhibitor	Vascular endothelial growth factor 2 antagonist
463	0.632	2	Cyclic GMP phosphodiesterase inhibitor	ErbB 2 antagonist	Vascular endothelial growth factor 2 antagonist
464	0.632	2	Cyclic GMP phosphodiesterase inhibitor	MAP kinase kinase inhibitor	Vascular endothelial growth factor 2 antagonist
465	0.632	2	Cyclic GMP phosphodiesterase inhibitor	Platelet activating factor beta antagonist	Vascular endothelial growth factor 2 antagonist
466	0.632	1	Cyclic GMP phosphodiesterase inhibitor	Transcription factor NF kappa B inhibitor	Vascular endothelial growth factor 2 antagonist
467	0.632	3	Cyclic GMP phosphodiesterase inhibitor	Tyrosine kinase inhibitor	Vascular endothelial growth factor 2 antagonist
468	0.632	1	Cyclic GMP phosphodiesterase inhibitor	Vascular endothelial growth factor 1 antagonist	Vascular endothelial growth factor 2 antagonist
469	0.632	3	Cyclic GMP phosphodiesterase inhibitor	Vascular endothelial growth factor antagonist	Vascular endothelial growth factor 2 antagonist
470	0.322	1	Cyclin-dependent kinase 2 inhibitor	Tyrosine kinase inhibitor	Vascular endothelial growth factor 2 antagonist
471	0.322	1	Cyclin-dependent kinase 2 inhibitor	Vascular endothelial growth factor antagonist	Vascular endothelial growth factor 2 antagonist
472	0.632	2	EgrB2 kinase inhibitor	Epidermal growth factor antagonist	Vascular endothelial growth factor 2 antagonist

PharmaExpert: Link to the KEGG Pathway (example)

The screenshot displays the PharmaExpert software interface. At the top, there is a search bar and a list of chemical structures from a library. Below this is a table of search results with columns for predicted values and activities. A table below the search results lists various activities and their predicted values.

Pa	Pi	Activity	Predicted value
0.969	0.000	Plain deaminase inhibitor	
0.931	0.003	Pale red cell aplasia	
0.916	0.001	ANCA positive vasculitis	

Below the table, there are tabs for 'KEGG', 'NCI Pathway', 'Reactome', and 'Gene Ontology'. The 'KEGG' tab is selected, showing a browser window with the KEGG pathway for Caffeine metabolism. The pathway diagram shows the conversion of Xanthosine to Caffeine and its subsequent metabolism into various products like 1,7-Dimethylxanthine, Theophylline, and Theobromine. Enzymes like CYP1A2 and CYP2A6 are indicated for specific steps.

Overview

1. Historical reflections
2. PASS (Prediction of Activity Spectra for Substances)
3. PharmaExpert
4. **GUSAR (General Unrestricted Structure-Activity Relationships)**
5. Some examples of applications
6. Our web-services based on PASS, GUSAR, etc.
7. **Way2Drug.com: further progress**

QNA: Quantitative Neighborhoods of Atoms

According to the Hellman-Feynman theorem, interatomic and intermolecular forces are electrical in nature.

Feynman R. Ph. *Phys. Rev.*, 1939, 56, 340-343.

$$P_i = B_i \sum_k (\text{Exp}(-\frac{1}{2}C))_{ik} B_k$$

$$Q_i = B_i \sum_k (\text{Exp}(-\frac{1}{2}C))_{ik} B_k A_k$$

$$A = \frac{1}{2}(IP + EA)$$

$$B = (IP - EA)^{-\frac{1}{2}}$$

IP is the first ionization potential,

EA is the electron affinity.

Robert G. Parr et al. *J. Chem. Phys.*, 1978, 68(8), 3801-3807.

Gasteiger J, Marsili M. *Tetrahedron*, 1980, 36, 3219-3228.

Rappe A K and W A Goddard III. *J. Ph. Ch.*, 1991, 95, 3358-3363.

QNA descriptors' space

Phosphorus

Sulfur

Iodine

Carbon

Hydrogen

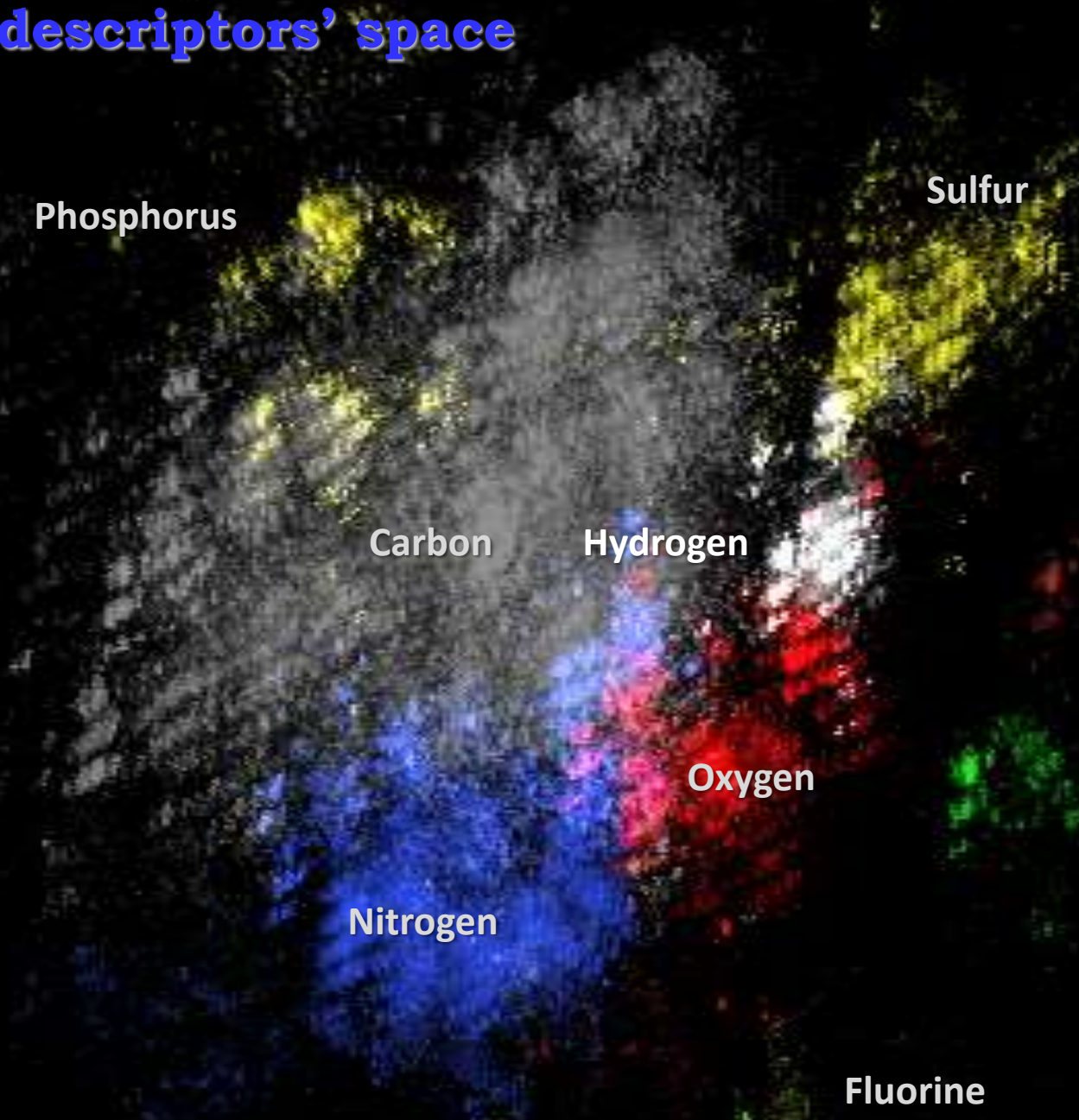
Bromine

Oxygen

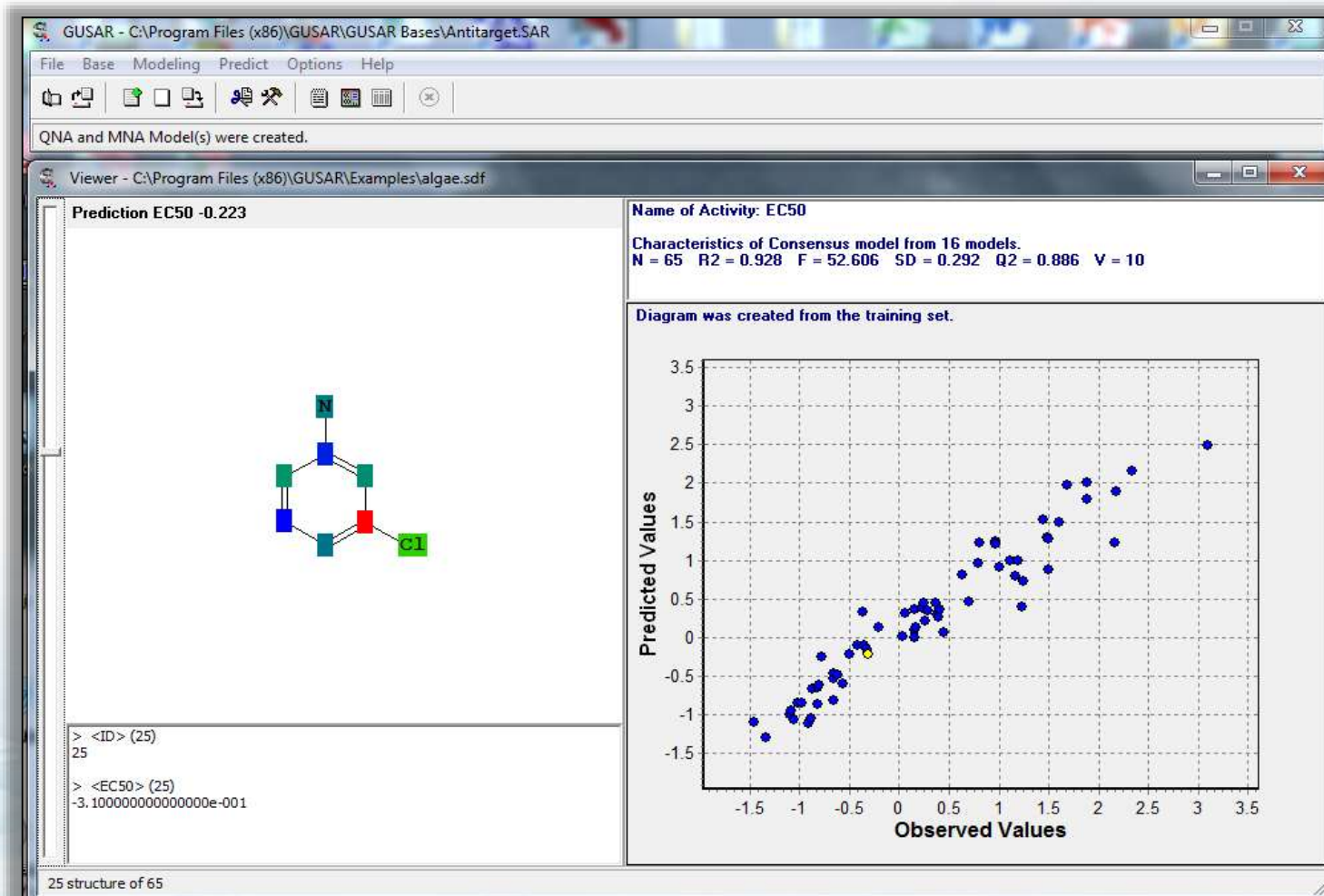
Chlorine

Nitrogen

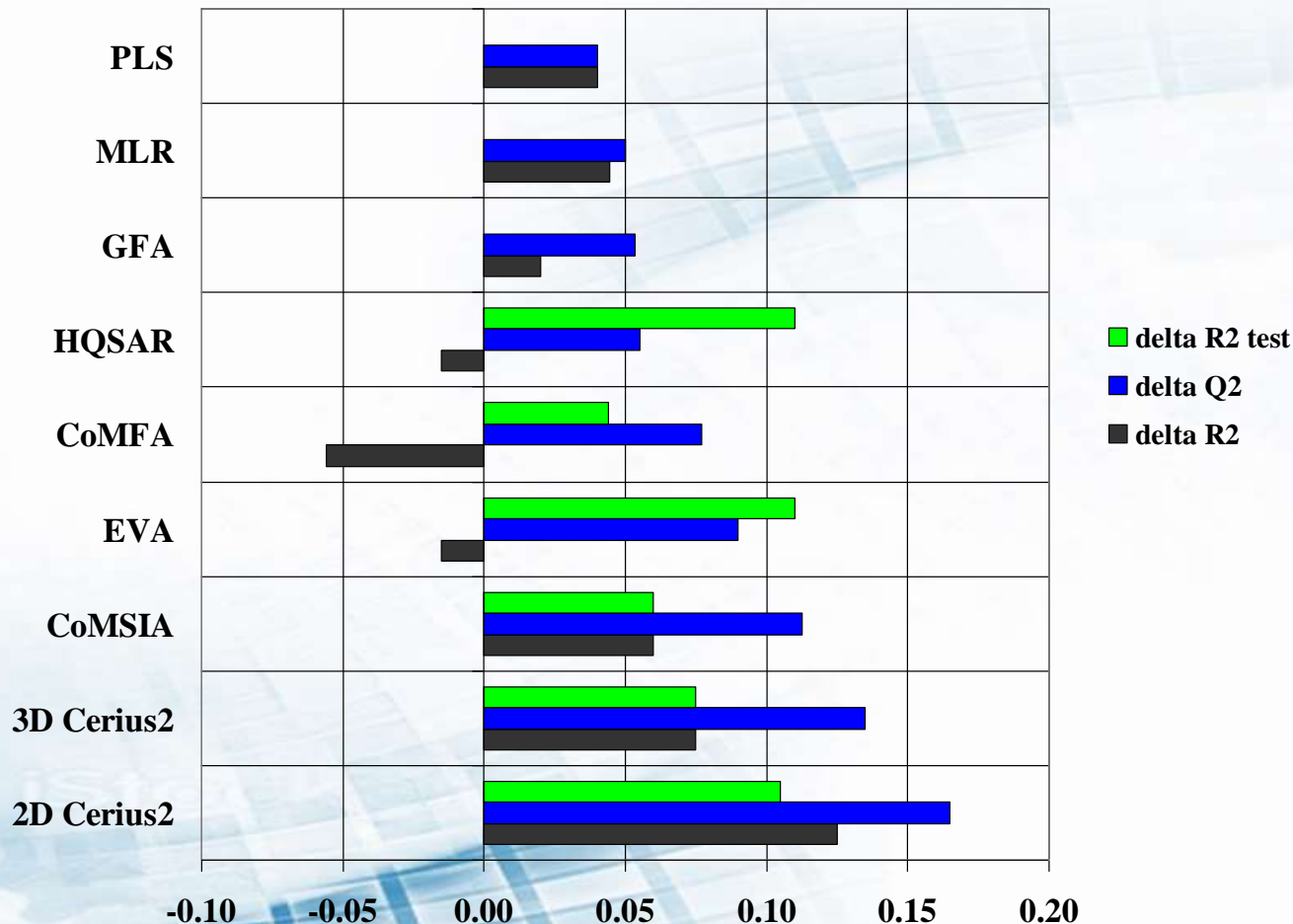
Fluorine



GUSAR: General Unrestricted Structure-Activity Relationships



GUSAR: Superiority in performance in comparison with some other (Q)SAR methods



Impacts of atoms into particular activity

For each atom in a molecule all MNA/QNA descriptors are generated. Using these descriptors for each particular activity P_a и P_i values are calculated.

Each atom is colored in accordance with the following:

Red	$:= 0.3 + 0.7 * P_i$	(negative impact on activity)
Green	$:= 0.3 + 0.7 * P_a$	(positive impact on activity)
Blue	$:= 1 - 0.7 * (P_i + P_a)$	(neutral impact on activity)

This can be interpreted in the following way:

If $P_a = 0$ and $P_i = 1$, then Red = 1, Green = 0.3, Blue = 0.3 – **bright red color**;

If $P_a = 1$ and $P_i = 0$, then Red = 0.3, Green = 1, и Blue = 0.3 – **bright green color**;

If $P_a = 0$ and $P_i = 0$, then Red = 0.3, Green = 0.3, Blue = 1 – **bright blue color**;

If $P_a = 0.33$ and $P_i = 0.33$, then Red = 0.53, Green = 0.53, Blue = 0.53 – **grey color**.

Case study: Sulfathiazole

Antibacterial Activity

ET_A Receptor Antagonist

PASS - C:\DATABASES\TEST-MOLECULES\sulphathiazole.sdf

File Base Predict View Options Help

C:\Program Files\PASS-ETC-AUG-2005\MNICKLAUS-AUG-2005\RunImage\PASS.SAR

C:\DATABASES\TEST-MOLECULES\sulphathiazole.sdf

Antibacterial

Activity Spectrum

Chart General Effects Mechanisms Toxicity

Dihydropteroate synthase inhibitor
Iodide peroxidase inhibitor

139 of 2005 Possible Activities at Pa > Pi

0.889	0.005	Antibesity
0.835	0.005	Para amino benzoic acid antagonist
0.736	0.006	Dihydropteroate synthase inhibitor
0.721	0.006	Antidiabetic
0.556	0.006	Antiprotozoal (Coccidial)
0.552	0.019	Prostaglandin E1 antagonist
0.509	0.026	Prostaglandin H2 antagonist
0.485	0.045	Potassium channel antagonist
0.453	0.013	Cyclooxygenase inhibitor
0.468	0.028	Antiprotozoal
0.443	0.012	Antibacterial
0.412	0.021	Diuretic inhibitor
0.408	0.024	Gingipain R inhibitor
0.421	0.053	Antinfective
0.371	0.006	Hypoglycemic
0.328	0.015	Antineoplastic (breast cancer)
0.362	0.054	Antimycobacterial
0.351	0.047	Antituberculosic
0.325	0.023	Saluretic
0.345	0.052	Myelodysplastic syndrome treatment

> cid> (2)
2

32 Substructure Descriptors; 0 new.
There are 3 known activities.
Drug-Likeness: 0.156

139 of 2005 Possible Activities
35 of 224 Possible Pharmacological Effects

2 structure of 2

PASS - C:\DATABASES\TEST-MOLECULES\sulphathiazole.sdf

File Base Predict View Options Help

C:\Program Files\PASS-ETC-AUG-2005\MNICKLAUS-AUG-2005\RunImage\PASS.SAR

C:\DATABASES\TEST-MOLECULES\sulphathiazole.sdf

Endothelin receptor antagonist

Activity Spectrum

Chart General Effects Mechanisms Toxicity

0.280	0.048	Ribonucleoside triphosphate reductase inhibitor
0.288	0.061	Channel-conductance-controlling ATPase inhibitor
0.254	0.029	Tubulin antagonist
0.269	0.061	Antiprotozoal (Trichomonas)
0.248	0.044	Thromboxane A2 antagonist
0.204	0.004	5 Hydroxytryptamine 6 antagonist
0.244	0.045	Lipoxygenase inhibitor
0.287	0.093	CyP2B2 substrate
0.246	0.060	Oligopeptidase B inhibitor
0.205	0.021	Thromboxane antagonist
0.235	0.059	Benzodiazepine inverse agonist
0.176	0.001	11-Beta-hydroxysteroid dehydrogenase 1 inhibitor
0.176	0.001	11-Beta-hydroxysteroid dehydrogenase inhibitor
0.264	0.100	Serine-phosphoethanolamine synthase inhibitor
0.241	0.083	Antithrombotic
0.235	0.079	Poly(ADP-ribose) glycohydrolase inhibitor
0.216	0.066	Corticosteroid antagonist
0.154	0.008	Thyroid hormone antagonist
0.219	0.074	Granzyme A inhibitor
0.246	0.106	Carcinogenic
0.279	0.139	Antulcerative
0.155	0.016	Beta tubulin antagonist
0.256	0.117	Carcinogenic, male mice
0.158	0.019	Endothelin receptor antagonist
0.237	0.107	(S)-3-hydroxyacid ester dehydrogenase inhibitor

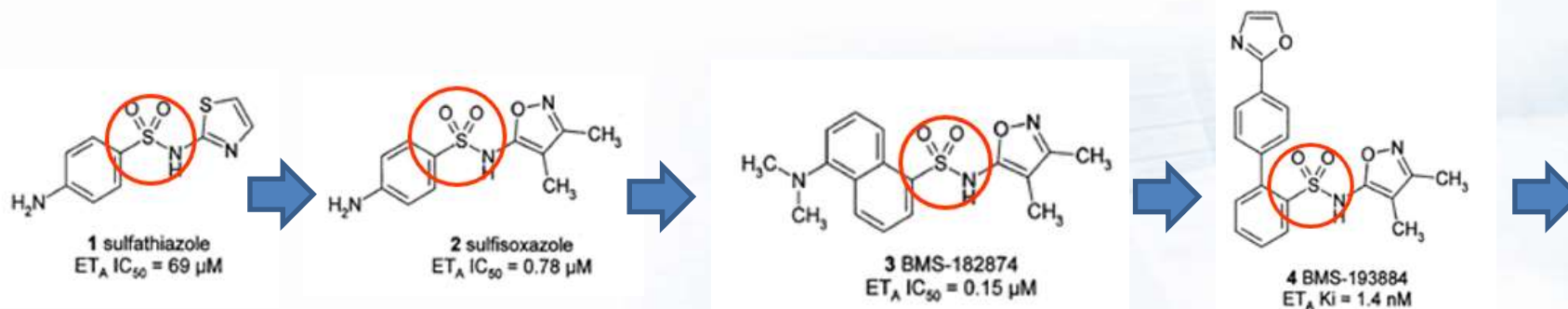
> cid> (2)
2

32 Substructure Descriptors; 0 new.
There are 3 known activities.
Drug-Likeness: 0.156

139 of 2005 Possible Activities
35 of 224 Possible Pharmacological Effects

2 structure of 2

“Critical fragment” remains unchanged



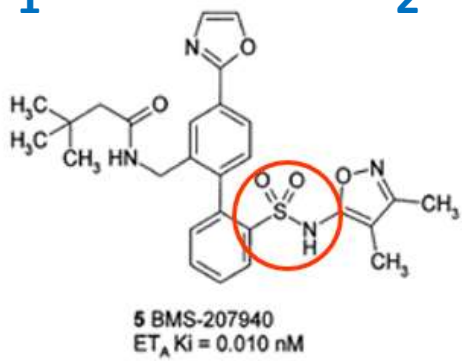
1

2

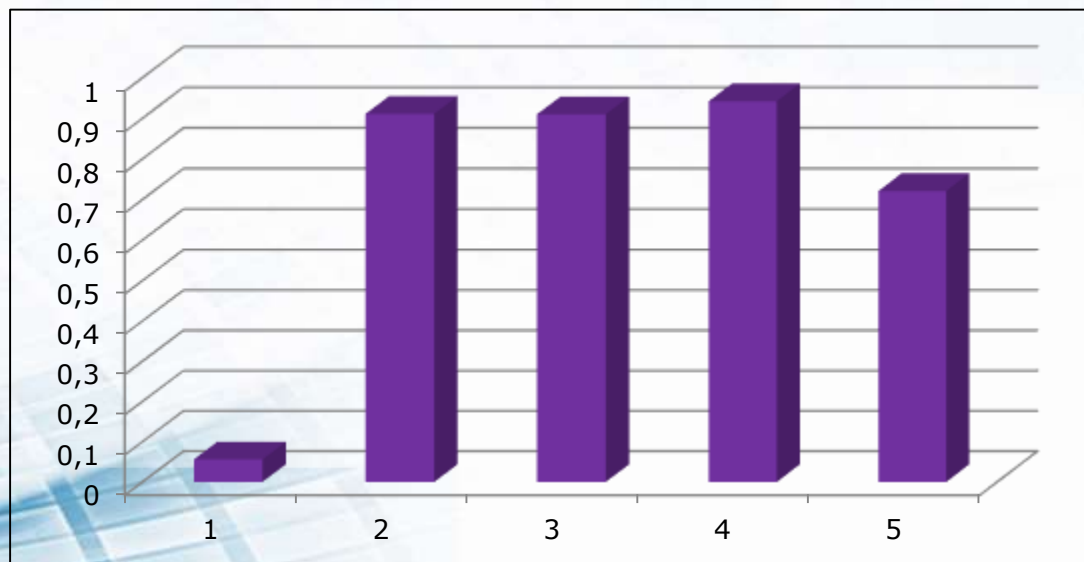
Pa

3

4



5



Compound No.

Overview

1. Historical reflections
2. PASS (Prediction of Activity Spectra for Substances)
3. PharmaExpert
4. GUSAR (General Unrestricted Structure-Activity Relationships)
5. **Some examples of applications**
6. Our web-services based on PASS, GUSAR, etc.
7. **Way2Drug.com: further progress**

GUSAR Application to the antifungal activities

Table 3. QSAR modeling of antifungal activities results.

Activity name	Number of compounds Training set/Test set	Number of models	R ² training set	Q ² training set	R ² test set	Coverage,%	RMSE test
<i>B.s.</i>	12/5	4	0.89	0.72	0.57	80	35.74
<i>F.m.</i>	12/5	21	0.89	0.77	0.80	100	28.01
<i>F.o.</i>	12/5	3	0.85	0.68	0.66	100	17
<i>R.s.</i>	12/5	20	0.91	0.79	0.72	100	27.58
<i>S.s.</i>	12/5	11	0.89	0.79	0.81	100	37.29
<i>V.i.</i>	12/5	2	0.83	0.61	0.82	100	20.37

R² – determination coefficient

Q² – determination coefficient calculated for leave-one-out cross validation procedure

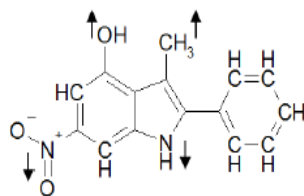


Figure 2. Atom contribution into the antifungal activity.

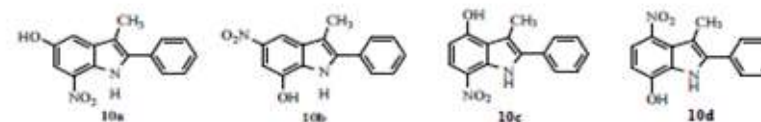
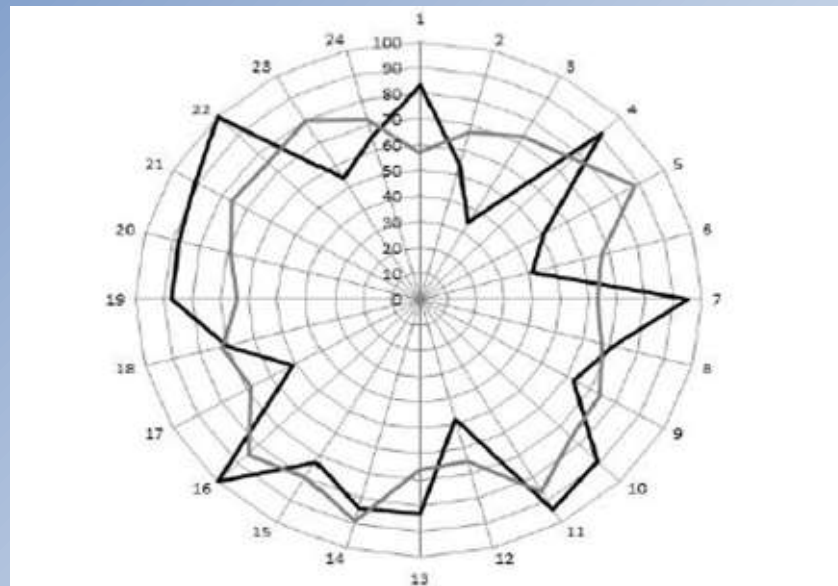


Figure 3. Comparison of the experimental (black line) and predicted (grey line) antifungal activities for compounds 10a (1-6), 10b (7-12), 10c (13-18), 10d (19-24).

Drug repositioning

DRUG REPOSITIONING: IDENTIFYING AND DEVELOPING NEW USES FOR EXISTING DRUGS

Ted T. Ashburn and Karl B. Thor

Biopharmaceutical companies attempting to increase productivity through novel discovery

Drug News Perspect 22(2), March 2009

MEETING REPORT

DRUG REPOSITIONING SUMMIT: FINDING NEW ROUTES TO SUCCESS

Highlights from the Cambridge Healthtech Institute's Third Annual Drug Repositioning Summit, held October 6-7, 2008, in Boston

Drug News Perspect 22(2), March 2009

INSIGHT

THE VALUE OF DRUG REPOSITIONING IN THE CURRENT PHARMACEUTICAL MARKET

Repurposing a primary drug discovery broadly research pharmaceutical companies

by Edward L. Tobinick

covery efforts efficiently has become of pri-

SUMMARY

REVIEWS



Reviews - MEDICINE REVIEW

In silico repositioning of approved drugs for rare and neglected diseases

Sean Ekins^{1,2,3,4},
Matthew D. Krasa

Drug Discovery Today • Volume 16, Numbers 7/8 • April 2011

SAR and QSAR in Environmental Research
2001, Vol. 12, pp. 327-344
Reprints available directly from the publisher
Photocopying permitted by license only

© 2001 OPA (Overseas Publishers Association) N.V.
Published by license under
the Gordon and Breach Science Publishers imprint,
a member of the Taylor & Francis Group.

OPEN ACCESS Freely available online

PLOS COMPUTATIONAL BIOLOGY

Drug Discovery Using Chemical Systems Biology: Repositioning the Safe Medicine Comtan to Treat Multi-Drug and Extensively Drug Resistant Tuberculosis

Sarah L. Kinnings^{1,2,3}, Nina Liu^{2,3}, Nancy Buchmeier^{3,3}, Peter J. Tonge², Lei Xie^{4*}, Philip E. Bourne^{4,5*}

1 Department of Biology, University of York, York, United Kingdom, 2 Institute of Chemical Biology & Drug Discovery, Department of Chemistry, Stony Brook University, Stony Brook, New York, United States of America, 3 Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California, United States of America, 4 San Diego Supercomputer Center, University of California San Diego, La Jolla, California, United States of America, 5 Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California, United States of America

TOP 200 MEDICINES: CAN NEW ACTIONS BE DISCOVERED THROUGH COMPUTER-AIDED PREDICTION?*

V. POROIKOV^{a,b,†}, D. AKIMOV^b, E. SHABELNIKOVA^b
and D. FILIMONOV^a

^aInstitute of Biomedical Chemistry of the Russian Academy of Medical Sciences, 10, Pogodinskaya Street, Moscow, 119832, Russia; ^bMedical and Biological Faculty of the Russian State Medical University, 1, Ostrovityanova Street, Moscow, 117869, Russia

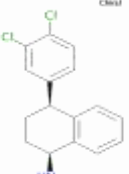
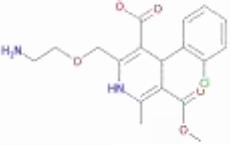
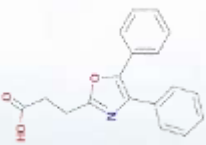
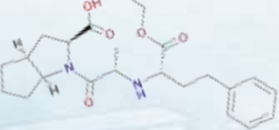
(Received 30 June 2000; In final form 31 March 2001)

Computer-aided prediction of the biological activity spectra by the program PASS was applied to a set of 130 pharmaceuticals from the list of the Top 200 medicines. The known pharmacological effects were found in the predicted activity spectra in 93.2% of cases. Additionally,

Drug repositioning based on PASS prediction

In 2001 we published predictions of new effects for 8 medicines from the list of Top200 Drugs [1].

Which predictions are confirmed?
(informational search, September 2014)

				Ref.
	Sertraline	Cocain dependency treatment	+	[2]
	Amlodipine	Antineoplastic enhancer (moderate BCRP/ABCG2 inhibitor)	+	[3]
	Oxaprozin	Interleukin 1 antagonist (Inhibitor of production of Interleukin 1 β)	+	[4]
	Ramipril	Antiarthritic	+	[5]

1. Poroikov V. et al. *SAR and QSAR Environ. Res.*, 2001, 12: 327-344.

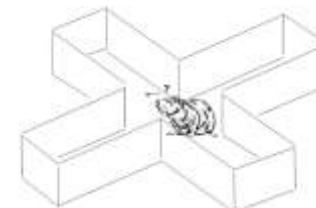
2. Mancino M.J. et al. *J. Clin. Psychopharmacol.*, 2014, 34: 234-239.

3. Takara K. et al. *Mol. Med. Rep.*, 2012, 5: 603-609.

4. Rainsford K.D. et al. *Inflammopharmacology*, 2002, 10: 85-239.

5. Shi Q. et al. *Arthritis Res. Ther.*, 2012, 14: R223.

Nootropic effect in some antihypertensive drugs?



Name	Pa (Nootropic effect), %
Captopril	44,6
Enalapril	65,5
Lisinopril	61,8
Perindopril	60,9
Quinapril	65,1
Ramipril	63,3
Monopril	30,9
Piracetam	81,7
Amlodipin	-
Hydrochlorothiazide	-

Perindopril in dose of **1 mg/kg**, and **quinapril** and **monopril** in doses of **10 mg/kg** improved the patrolling behavior in the maze, like **piracetam** and **meclofenoxate** (in doses of **300** and **120 mg/kg**, respectively).

BMJ Open 2013;3:e002881 doi:10.1136/bmjopen-2013-002881

Geriatric medicine

Effects of centrally acting ACE inhibitors on the rate of cognitive decline in dementia

Yang Gao^{1,2}, Rónán O'Caolmh¹, Liam Healy¹, David M Kerins^{3,4}, Joseph Eustace⁵, Gordon Guyatt⁶, David Sammon², D William Molloy^{1,7}

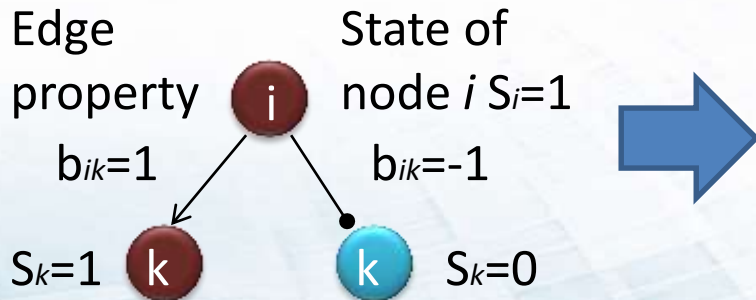
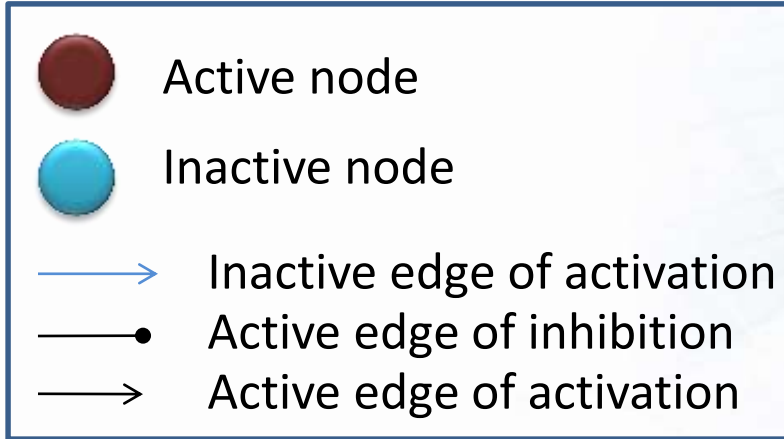
• Author Affiliations

Correspondence to

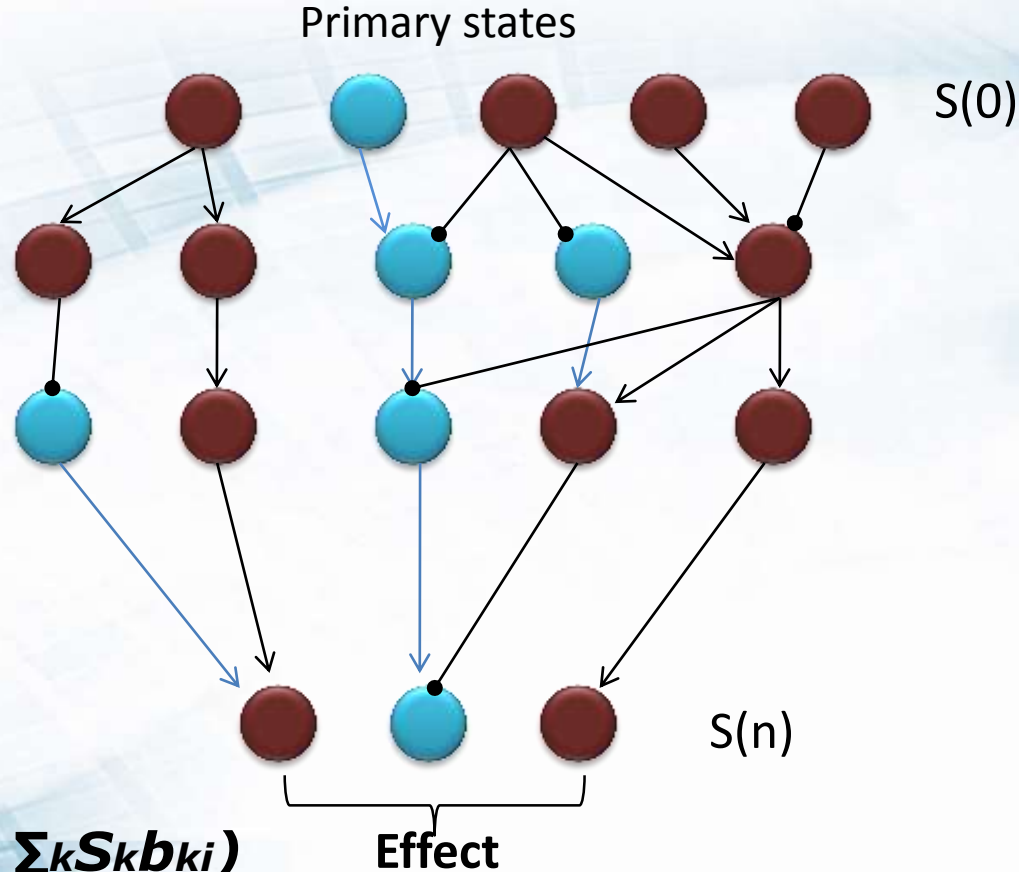
Professor D William Molloy; w.molloy@ucc.ie

Published 22 July 2013

Dichotomic modelling of regulatory networks behavior



$$F_i(S_1, S_2, \dots, S_n) = \Theta(a_i + \sum_k S_k b_{ki})$$



Input Data for Breast Cancer Modeling

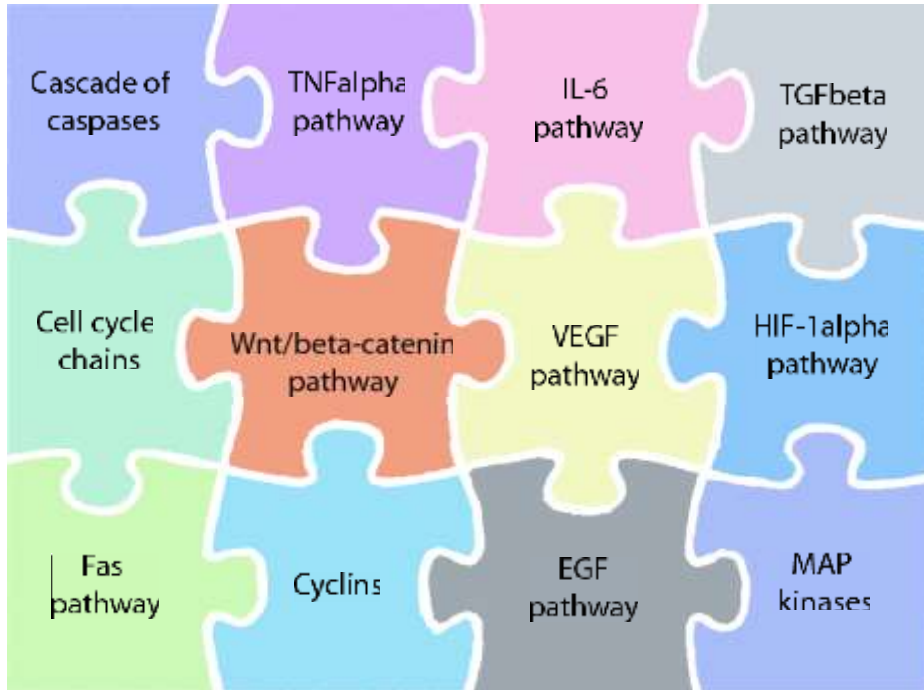
Regulatory network
TRANSPATH® database

Fragment: 2336 edges and 1405 nodes

**Microarray data for
breast cancer**

Cyclonet database

<http://cyclonet.biouml.org>



- HER2/neu-positive breast carcinomas.
- Ductal carcinoma.
- Invasive ductal carcinoma and/or a nodal metastasis.
- Generalized breast cancer.

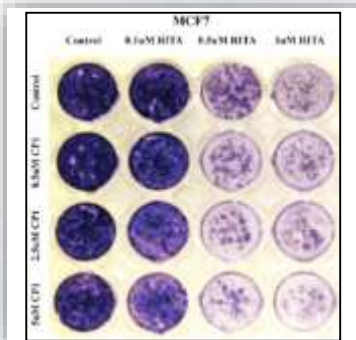
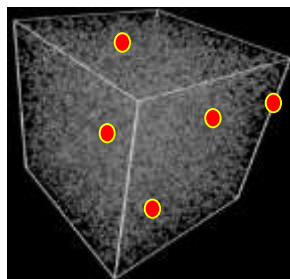
Some Double and Triple Targets' Combinations Identified for Breast Cancer

No	Compounds	Action 1	Action 2	Action 3
1	4	Bcl2 antagonist	CDK-2 inhibitor	
2	10	Bcl2 antagonist	Myc inhibitor	
3	10	Bcl2 antagonist	Phosphatidylinositol 3-kinase beta inhibitor	
4	3	CDK-2 inhibitor	Myc inhibitor	
5	7	HIF-1 alpha inhibitor	Myc inhibitor	
6	10	HIF-1 alpha inhibitor	Phosphatidylinositol 3-kinase beta inhibitor	
7	10	Myc inhibitor	Phosphatidylinositol 3-kinase inhibitor	
8	10	Bcl2 antagonist	Myc inhibitor	Phosphatidylinositol 3-kinase beta inhibitor

European project «From analysis of gene regulatory networks to drug» (Net2Drug)



Participants: 9 teams
from 8 countries



2 active compounds
(BC, melanoma)
Synergism with RITA.

ChemNavigator database
(~24,000,000 structures of organic
compounds)

Virtual screening of potential
multitarget anticancer substances
(PASS, GUSAR)

11 compounds tested in cellular
assays

Further progress:

Activity confirmed in experiments
on mouse xenograft models

ALab – resident of «Skolkovo» (2012)

Grant of «Skolkovo» (2013)

More active analogs (2014)

Mechanism(s) of action (2015)

InterBioScreen library of natural compounds

File Edit Options Object Database Search List Window Help

ibs2014oct_nc.db/ibs

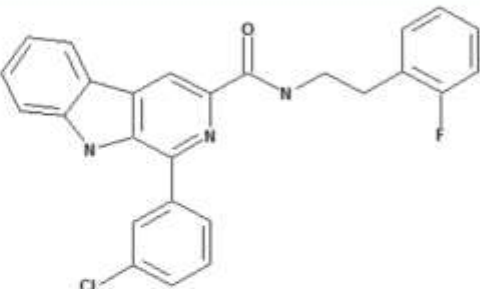
<Root> 25876 of 55580
Search Domain: All

Forms Query Browse Update

IBS
INTERBIOSCREEN

Tel: +7 49652 40091 E-mail: screen@ibscreen.chg.ru
Fax: +7 49652 40092 Web: http://www.ibscreen.com

Natural Compounds



ID:	STOCK1N-51602	Salt:		MW:	443.9122
Formula:	C ₂₆ H ₁₉ ClFN ₃ O	Comment:		Index:	DNC
H-Bond Donors:	2	H-Bond Acceptors:	4	Rotatable Bonds:	6
				Rings:	5
				TPSA:	57.7800
				CLogP:	7.4390

IUPAC Name:

1-(3-chlorophenyl)-N-(2-fluorophenethyl)-9H-pyrido[3,4-b]indole-3-carboxamide

<http://www.ibscreen.com/>

Predictions of RA associated activities for IBS library

PharmaExpert

Prediction & Interpretation - C:\Users\vvp\Desktop\RIGA-JAN-2016\ACTIVITY-SELECTION\ibs2015mar_nc (PASS2014)-RA-Pa-more-20.SDF: 1/55742

STOCK1N-00004 STOCK1N-00005 STOCK1N-00007 STOCK1N-00010 STOCK1N-00013 STOCK1N-00022 STOCK1N-00029 STOCK1N-00028 STOCK1N-00030

Pa	Pi	Activity
0.372	0.040	Angiogenesis inhibitor
0.320	0.077	Apoptosis agonist
0.217	0.061	Interleukin 10 agonist
0.264	0.190	Antiinflammatory

Effect: 3 Mechanism: 2

Interleukin 10 agonist 0.217 0.061

Statistics

No	Pa>Pi	Pa>30%	Pa>50%	Pa>70%	Types of Activity
1	31530	20306	7185	1439	Apoptosis agonist
2	31215	22575	9446	2873	Antiinflammatory
3	22759	14214	5660	2249	Immunosuppressant
4	22644	9403	2397	851	Cell adhesion molecule inhibitor
5	20171	13162	5044	2250	Transcription factor NF kappa B inhibitor
6	18179	9622	2198	463	Hypoxia inducible factor 1 alpha inhibitor
7	16357	9474	3122	482	Free radical scavenger
8	13831	5136	526	0	Transcription factor STAT3 inhibitor
9	12824	7083	1678	171	Immunomodulator
10	12813	3555	323	8	Nitric oxide scavenger
11	11262	740	20	3	T cell inhibitor
12	10850	4173	862	193	Antioxidant
13	8937	1757	18	1	Phospholipase A2 inhibitor
14	8743	5062	1716	374	Autoimmune disorders treatment
15	8207	3387	660	121	Non-steroidal antiinflammatory agent
16	5219	1328	295	142	Interleukin 6 antagonist
17	4681	1507	359	89	Interleukin antagonist
18	4402	2019	513	69	Rheumatoid arthritis treatment

Number of selected compounds:

<id> STOCK1N-00002; 53 Substructure descriptors, 0 new; 4 Possible activities.

Search for multitargeted agents for RA treatment

PharmaExpert

File Tools View Help

No Patient

Prediction & Interpretation

STOCKIN-0000, H O

STOCKIN-00004

STOCKIN-00006

Save TXT Save SD Clipboard Exclude

Pa Pi Activity

Pa	Pi	Activity
0.372	0.040	Angiogenesis inhibitor
0.320	0.077	Apoptosis agonist
0.217	0.061	Interleukin 10 agonist
0.264	0.190	Antiinflammatory

Pa Pi <id>

Number of selected compounds:

<id> STOCKIN-00002: 53 Substructure descriptors, 0 new; 4 Possible activities.

Multitargeted actions

Effects

Rheumatoid arthritis treatment

Number of targets: 4

Run Load Save

- 5-Lipoxygenase inhibitor
- Adenosine deaminase inhibitor
- Bisphosphonate
- Calcineurin inhibitor
- Cathepsin K inhibitor
- CC chemokine 1 receptor antagonist
- CC chemokine 2 receptor antagonist
- Chemokine receptor antagonist
- Collagenase inhibitor
- CXC chemokine 4 receptor antagonist
- Cyclooxygenase 2 inhibitor
- Cyclooxygenase inhibitor
- Dihydropyridine dehydrogenase inhibitor
- Free radical scavenger
- Gelatinase inhibitor
- Integrin antagonist
- Interferon gamma antagonist
- Interleukin 1 antagonist
- Interleukin 1 beta converting enzyme inhibitor
- Interleukin 10 agonist
- Interleukin 1b antagonist
- Interleukin 5 antagonist
- Interleukin 9 antagonist

No	Pa	Number	Activity type	Activity type	Activity type	Activity type
574	0.252	3	Integrin antagonist	Interleukin 5 antagonist	VLA-4 antagonist	
575	0.776	1	Integrin antagonist	Phospholipase A2 inhibitor	VLA-4 antagonist	
576	0.307	1	Integrin antagonist	Thiol protease inhibitor	VLA-4 antagonist	
577	0.252	3	Integrin antagonist	Transcription factor NF kappa B inhibitor	VLA-4 antagonist	
578	0.244	1	Interleukin 10 agonist	Interleukin 6 antagonist	VLA-4 antagonist	
579	0.307	1	Interleukin 10 agonist	Thiol protease inhibitor	VLA-4 antagonist	
580	0.244	1	Interleukin 10 agonist	Transcription factor NF kappa B inhibitor	VLA-4 antagonist	
581	0.252	3	Interleukin 6 antagonist	Transcription factor NF kappa B inhibitor	VLA-4 antagonist	
582	0.296	2	Free radical scavenger	Integrin antagonist	VLA-4 antagonist	Transcription factor NF kappa
583	0.296	2	Free radical scavenger	Interleukin 6 antagonist	VLA-4 antagonist	Transcription factor NF kappa
584	0.265	1	Integrin antagonist	Interleukin 10 agonist	VLA-4 antagonist	Transcription factor NF kappa
585	0.296	3	Integrin antagonist	Interleukin 6 antagonist	VLA-4 antagonist	Transcription factor NF kappa
586	0.295	1	Interleukin 10 agonist	Interleukin 6 antagonist	VLA-4 antagonist	

Pa Pi > Pi > 5-Lipoxygenase

20/24 11/07/2016

One promising hit for experimental validation

PharmaExpert

Prediction & Interpretation - C:\Users\wvp\Desktop\RIGA-JAN-2016\ACTIVITY-SELECTION\ibs2015mar_nc (PASS2014)-RA-Pa-more-20.SDF, 9703/55742

STOCK1N-27990 STOCK1N-27991 STOCK1N-27995 STOCK1N-27996 STOCK1N-27997 STOCK1N-27999 STOCK1N-28005 STOCK1N-28006 STOCK1N-28007 STOCK1N-28008

Save TXT Save SD Clipboard Exclude

Pa	Pi	Activity
0.849	0.003	Integrin antagonist
0.778	0.002	VLA-4 antagonist
0.640	0.011	Autoimmune disorders treatment
0.460	0.068	Antiinflammatory
0.374	0.018	Free radical scavenger
0.361	0.050	Immunomodulator
0.316	0.107	Immunosuppressant
0.208	0.046	Phospholipase A2 inhibitor

Effect: 4 Mechanism: 6 Antitarget: 1

KEGG | NCI Pathways | Reacto | Therapeutic effects

0.208 0.046 Phospholipase A2 inhibic
alpha-Linolenic acid
Arachidonic acid met
Ether lipid metabolism
Fat digestion and ab
Glycerophospholipid
Linoleic acid metaboli
Pancreatic secretion
Vascular smooth mus

UniProt ID	Gene name(s)	Species

Pa Pi Activity Predicted value descending

5Lipoxygenase inhibitor New Descriptors: 0

Number of selected compounds: 1

<id> STOCK1N-28007; 24 Substructure descriptors; 0 new; 3 Possible activities.

PASS + PharmaExpert

The search for new compounds with multiple mechanisms of action

J. Med. Chem., 2003, 46(15), 3326-3332

J. Med. Chem. 2008, 51(6), 1601-1609

Drug repositioning

Pharmaceut. Chem. J., 2011, 45 (10), 605-611

The search for new compounds with specific therapeutic effect(s) or/and interaction with specific target(s)

J. Med. Chem., 2004, 47(11), 2870-2876

Bioorg. Med. Chem., 2004, 12(24), 6559-6568

Eur. J. Med. Chem., 2009, 44 (2), 473-481

Assessment of drug-drug interactions and between natural compounds - components of medicinal plants

Curr. Pharm. Des. 2010, 16(15), 1703-1717

Med. Chem. Res. 2011, 20(9), 1509-1514

Cardiovascul. Therap. Prof., 2008, 7(5), 100-104

More info: way2drug.com/suc_stor.php

Success Stories

Here you will find a brief overview of publications describing applications of our software. This overview provides you a guide how people use our tools, to achieve better results in their projects aimed at drug design & discovery or chemical safety assessment. Among the over 250 publications, there are many success stories about virtual screening, drug repurposing, revealing the hidden potential of natural products, chemical safety assessment, etc.

If you have some experience with the utilization of our software, please, [tell us your story](#). Sharing of your experience, will support newcomers in their first steps to apply computer-aided drug discovery methods in practice, and will help us to improve our web-services.

- ◆ [General papers with references on our computational tools](#)
- ◆ [Virtual Screening](#)
- ◆ [Drug repurposing](#)
- ◆ [Drug safety & risk assessment](#)
- ◆ [Evaluation of hidden potential of natural products](#)
- ◆ [Analysis of fragments' contribution to the activity](#)
- ◆ [Some other PASS predictions confirmed by the experiments](#)
- ◆ [List of references](#)
- ◆ [Some papers cited us](#)

General papers with references on our computational tools

"One of the first approaches in the field of in silico pharmacology was PASS (Prediction of Activity Spectra for Substances), which applies a set of 2D descriptors to compounds that are then correlated with a set of bioactivities."

Vegner L. et al. [J. Med. Chem.](#) 2013, 56, 8377. DOI: 10.1021/jm400813y

"Several ligand-based methods apply data mining methods in order to identify unknown drug-target interactions. One of the first initiatives in this field was PASS developed by Poroikov et al. (SAR & QSAR Environ. Res., 2007, 18: 101). It can predict the biological activity profile of a compound based on the analysis of structure-activity relationships for more than 250 000 biologically active substances."

Peragovics A. et al. [J. Chem. Inform. Model.](#) 2013, 53, 103. DOI: 10.1021/ci3004489

"Thorough studies have revealed pronounced differences between natural and synthetic compounds in terms of their structural and physicochemical properties, which renders the inference of targets for natural products from well-characterized drug-like compounds conceptually difficult. In fact, only a few select applications have been described." (One of the two mentioned publications is Laouni A., Filimonov D., Poroikov V. Multi-targeted natural products

Overview

1. Historical reflections
2. PASS (Prediction of Activity Spectra for Substances)
3. PharmaExpert
4. GUSAR (General Unrestricted Structure-Activity Relationships)
5. Some examples of applications
6. **Our web-services based on PASS, GUSAR, etc.**
7. Way2Drug.com: further progress

Way2Drug web platform

We have proposed the local correspondence concept, which is based on the fact that most biological activities of organic compounds are the result of molecular recognition, which in turn depends on the correspondence between the particular atoms of the ligand and the target.

.....

Using this concept, we have developed a consistent system of atom-centered neighborhoods of atoms descriptors including MNA, QNA, and LMNA, and have implemented them in several SAR/QSAR/QSPR modeling approaches.



PASSOnline

Pa	P	Activity
0.913	0.002	Phenolyleto
0.834	0.002	Indoleacetic acid O-acetyltransferase inhibitor
0.834	0.002	Antipyrone
0.830	0.003	Antibiotic
0.828	0.002	Propyl aminoglycoside inhibitor
0.822	0.004	Prostaglandin H2 synthase inhibitor
0.819	0.004	Alkylglyoxalaminolase inhibitor
0.814	0.004	Chitinase inhibitor
0.811	0.003	Chitinase inhibitor



Predicts about 4000 biological activity types of organic compounds by their structural formulas, including pharmacological effects, mechanisms of action, toxicity and side effects, interaction with metabolic enzymes, effects on gene expression, etc.



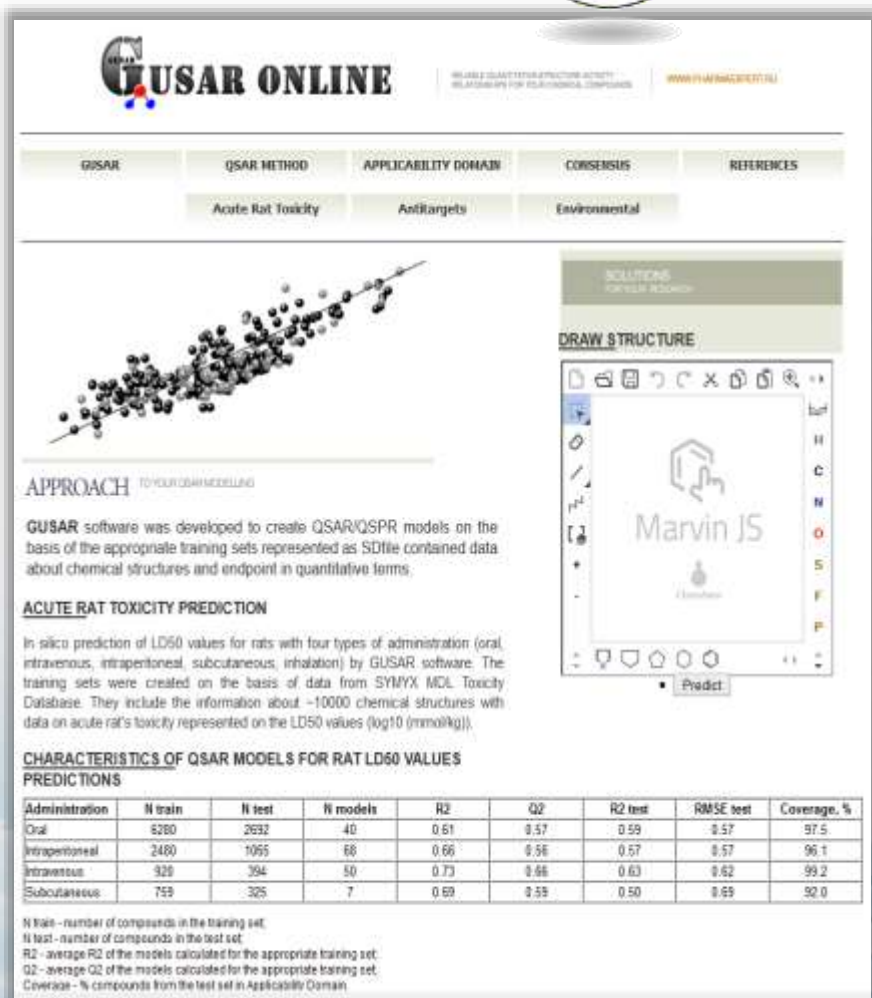
Training set with more than 313,000 known biologically active substances, belonging to different chemical classes.



Constantly working to improve the quality of prediction, updating the training set, and making changes in calculation methods.

average training set LOO CV: 0.95

GUSAR Online



GUSAR ONLINE

IN SILICO QSAR/QSPR PREDICTION SOFTWARE
RELATIONSHIP FOR ECOLOGICAL CHEMISTS

WWW.WAY2DRUG.COM

GUSAR	QSAR METHOD	APPLICABILITY DOMAIN	CONSENSUS	REFERENCES
	Acute Rat Toxicity	Antitargets	Environmental	

APPROACH TO QSAR MODELING

GUSAR software was developed to create QSAR/QSPR models on the basis of the appropriate training sets represented as SDF file contained data about chemical structures and endpoint in quantitative terms.

ACUTE RAT TOXICITY PREDICTION

In-silico prediction of LD50 values for rats with four types of administration (oral, intravenous, intraperitoneal, subcutaneous; inhalation) by GUSAR software. The training sets were created on the basis of data from SYMFX MCL Toxicity Database. They include the information about ~10000 chemical structures with data on acute rat's toxicity represented on the LD50 values (log10 (mmol/kg)).

CHARACTERISTICS OF QSAR MODELS FOR RAT LD50 VALUES PREDICTIONS

Administration	N train	N test	N models	R2	Q2	R2 test	RMSE test	Coverage, %
Oral	6280	2632	40	0.61	0.57	0.59	0.57	97.5
Intraperitoneal	2480	1055	68	0.66	0.56	0.57	0.57	96.1
Intravenous	929	394	50	0.73	0.66	0.63	0.62	99.2
Subcutaneous	759	325	7	0.69	0.59	0.50	0.69	92.0

N train - number of compounds in the training set.
N test - number of compounds in the test set.
R2 - average R2 of the models calculated for the appropriate training set.
Q2 - average Q2 of the models calculated for the appropriate training set.
Coverage - % compounds from the test set in Applicability Domain.

GUSAR online presents: consensus prediction, applicability domain assessment, internal and external models validation and clearly interpretations of obtaining results.

Developed to create QSAR/QSPR models on the basis of the appropriate training sets represented as SDF file format contained data about chemical structures and different endpoints in quantitative terms.

- Prediction of acute rat toxicity;
- Prediction of antitarget interaction profiles for chemical compounds;
- Prediction of ecotoxicity for chemical compounds.

DIGEP-Pred

PHARMAEXPERT
PREDICTIVE SERVICES

DIGEP-Pred: Prediction of drug-induced changes of gene expression profile

Home | Training Sets | Products/Services | Interpretation | Contacts

SWILES | Use files | Matrix input | Prediction by mRNA based data | Protein based data | CoC2cc(C=CC)=OCC(F=C) | Make prediction

Pa>0.5 | Save All | Save Data

mRNA based prediction result

Pa	Pi	DownRegulation	Pa	Pi	UpRegulation
0.8950.004		HIF1A	0.8950.007		VASP
0.9340.004		CDC2L2	0.8440.004		EGCG1
0.9240.003		ESF2	0.9120.002		BRSP2
0.9120.004		TNFRSF25B	0.8280.003		CCL11
0.8800.005		KRAS	0.8920.007		IL6
0.8830.001		NLE10	0.8940.001		SDC1
0.8280.008		CYP1B1	0.8900.004		TGFR2
0.7220.009		CKORF48	0.8980.028		PTGDR1
0.6780.038		TCR8	0.8340.007		CASP2

Protein based prediction result

Pa	Pi	DownRegulation	Pa	Pi	UpRegulation
0.9290.003		MMP2	0.9530.005		HMGB1
0.9280.004		CASP8	0.8980.002		IGG01
0.9200.003		FLT1	0.7820.019		CCL11
0.8280.004		MMP2	0.7500.008		SMN2
0.8050.004		IL13	0.8820.003		HSPH4
0.7490.005		TNFP2	0.7010.041		IL6
0.7540.011		CTSLB1	0.6140.066		FLAU
0.6340.002		MMP14	0.6200.109		CD81
0.5750.080		MMP7	0.5880.134		HRT1

Click to the name of gene to see the relationships of genes with diseases, side effects and biological pathways in Comparative Toxicogenomics Database.

Curated chemical-gene interactions data in the training sets were retrieved from the Comparative Toxicogenomics Database (CTD), Mount Desert Island Biological Laboratory, Salisbury Cove, Maine.
<http://ctdbase.org> [October, 2012]

Lagunin A., Ivanov S., Rudik A., Filimonov D., Porokov V. DIGEP-Pred: web-service for in silico prediction of drug-induced gene expression profiles based on structural formula. *Bioinformatics*, 2013, 29: 2062-2063. [\[abstract PDF\]](#)

Gene expression profiles are used to solve various problems in pharmaceutical research, such as the repositioning of drugs, overcoming resistance, estimating toxicity and drug-drug interactions.

Training sets:
mRNA-based - 1385 compounds for 952 genes (500 up- and 475 downregulations);
Protein-based - 1451 compounds for 129 genes (85 up- and 51 downregulations).

Results of prediction are linked to CTD (Comparative Toxicogenomics Database) for the purpose of their interpretation.

mRNA-based training set LOO CV: 0.853



CLC-Pred

Way2Drug PREDICTIVE SERVICES
 Understanding Chemical-Biological Interactions

CLC-Pred: *in silico* prediction of cytotoxicity for tumor and normal cell-lines

Home Training Set Products/Services Interpretation Contacts

SMILES: Pa0.5 v Save Log Save Note

Tumor cell-line prediction result

Pa	PI	Cell-line	Cell-line full name	Tissue	Tumor type
0.598	0.153	523T	Epithelial bladder carcinoma cells	Bladder	Carcinoma
0.538	0.155	CEM19	Leukemic T-cells	Blood	Leukemia
0.577	0.145	HCC95	Non-small cell lung carcinoma cells	Lung	Carcinoma
0.25	0.25	OVCAR-5	Ovarian adenocarcinoma cells	Ovary	Adenocarcinoma
0.25	0.25	CMS-1	Kidney carcinoma cells	Kidney	Carcinoma
0.87	0.87	MCF-10	Gastric adenocarcinoma cells	Stomach	Adenocarcinoma

Normal cell-line prediction result

PI	Cell-line	Cell-line full name	Tissue
0.011	HFF	Fibroblast fibroblasts	Fibroblast
0.032	IMR90	Fibroblast	Fibroblast
0.005	FHsC	Fibroblast	Fibroblast

Chemical structure: CC(=O)OC1=CC=C(C=C1)C(=O)O

Make prediction

The data on cytotoxicity of chemicals for tumor and normal cell-lines in the training sets were retrieved from ChEMBLdb (version 17) (<https://www.ebi.ac.uk/chembl/>) [August, 2015]. EMBL-EBI

Web-service for *in silico* prediction of cytotoxicity to the tumor and non-tumor cell-lines based on structural formula of chemical compound.

Training sets on the basis of DB ChEMBLdb (ver.17) were collected from 76804 chemical compounds, which reflected the current level of knowledge of the cytotoxicity of chemical compounds in relation to the 44 tumor and 48 non-tumor cell-lines.

In this case, the spectrum of biological activity is the assessment of cytotoxicity in relation to different cell lines.

Training set LOO CV: 0.96



Meta-Pred

Way2Drug PREDICTIVE SERVICES
Understanding Chemical-Biological Interactions

SMP: Prediction of substrate/metabolite specificity

Substrate based prediction result

Pa	Pi	Enzyme
0.963	0.005	UGT1A8
0.875	0.007	UGT1A8
0.752	0.011	UGT1A1
0.673	0.009	UGT1A10
0.334	0.148	UGT2B7
0.259	0.112	UGT1A7
0.263	0.173	UGT1A6
0.294	0.212	UGT1A3
0.228	0.183	UGT2B15
0.224	0.245	UGT2B4

Metabolite based prediction result

Pa	Pi	Enzyme
0.152	0.069	UGT1A8
0.149	0.069	UGT1A1
0.134	0.063	UGT1A10

Make prediction

JME Editor: courtesy of Peter Ertl, Novartis
The creation of web-service was supported by Russian Scientific Foundation grant 14-15-00449

Prediction of interaction with 18 cytochrome P450 and UGT isoforms: CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4, UGT1A10, UGT1A1, UGT2B7, UGT1A7, UGT2B15, UGT1A8, UGT1A4, UGT2B17, UGT2B10, UGT1A3, UGT1A9, UGT1A6, UGT2B4.

Substrate training set – 3411 compounds.

Metabolite-based training set – 2104 compounds.

Training set LOO CV: 0.934



SOMP

Prediction of sites of metabolism for drug-like compounds for (five major human) cytochrome P450s: CYP1A2, CYP2C9, CYP2C19, CYP2D6 and CYP3A4. Also in the training set were included the sites of glucoronidation, catalyzed by UGT.

Atom number	Rank	Cofactor
4	1	0.308
5	2	0.093
10	3	-0.992
8	4	-0.997
2	5	-0.991
9	6	-0.924
13	7	-0.898
1	8	-0.852
7	9	-0.717
11	10	-0.527
3	11	-0.452
12	12	-0.382
6	13	-0.293

Enzyme	Substrate amount	LOO CV
CYP3A4	960	0.89
CYP2D6	588	0.92
CYP2C9	446	0.92
CYP2C19	388	0.93
CYP1A2	573	0.92
UGT	592	0.98

>500 K predictions

91 country

12 949 users



- India
- Russia
- Ukraine
- Mexico
- China
- United States
- Egypt
- Kazakhstan
- Brazil
- Other



About 300 papers published citing our web-services
(>50% with the experimental confirmation; the other 50% -
just with the prediction results without experimental testing)

Available online at www.sciencedirect.com

European Journal of Medicinal Chemistry 45 (2010) 1015–1024

Original article

Synthesis, properties, and perspectives of *gem*-diphosphono substituted-thiazoles

Available online at www.sciencedirect.com

Bioorganic & Medicinal Chemistry Letters 15 (2005) 2145–2148

Bioorganic Medicinal Chemistry Letters

Quinazolines revisited: search for novel anxiolytic and GABAergic agents

ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Jan. 2003, p. 174–180
0066-4804/03/5008-00+0 DOI: 10.1128/AAC.47.1.174-180.2003
Copyright © 2003, American Society for Microbiology. All Rights Reserved.

Vol. 47, No. 1

In Vitro Activities of 7-Substituted 9-Chloro and 9-Amino-2-Methoxyacridines and Their Bis- and Tetra-Acridine Complexes against *Leishmania infantum*

Carole Di Giorgio,^{1*} Florence Delmas,¹ Nathalie Filloux,² Maxime Robin,² Laetitia Seferian,² Nadine Azas,¹ Monique Gasquet,¹ Muriel Costa,¹ Pierre Timon-David,¹ and Jean-Pierre Galzy²

Laboratoire de Parasitologie, Hygiène et Zoologie, Faculté de Pharmacie, Marseille Cedex 05,¹ and Laboratoire de Valorisation de la Chimie Fine, Université d'Aix-Marseille III, Site de Saint Jérôme, Marseilles,² France

Bioorganic & Medicinal Chemistry 20 (2012) 2930–2939

620, Russian Journal of Bioorganic Chemistry, 2013, Vol. 39, No. 2, pp. 202–210. © Pleiades Publishing, Ltd., 2013.
Russian Text © O.B. Kazakova, I.E. Smirnova, H. Do Tkhi Tkhu, Tkhanh Tra Nguen, G.N. Apryshko, O.S. Zhukova, N.I. Medvedev, A.F. Ismagilova, K. Yu. Suponitsky, D.V. Kazakov, F.E. Safarov, G.A. Tolstikov, 2013, published in Bioorganicheskaya Khimiya, 2013.

Synthesis, Structure, and Pharmacological Activity of (7*D*, 8*C*) Epoxy (12*D*, 17*D*) tricyclic Abiotic Compounds

UDC 547.67

V.I. Zvarych, R.Ya. Musyanovych, V.G. Chervetsov, O.Z. Komarovska-Porokhnyavets, M.V. Stasevych, V.P. Novik, Lviv Polytechnic National University, Department of Technology of Biologically Active Substances, Pharmacy and Biotechnology

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Identification of novel isocytosine derivatives as xanthine oxidase inhibitors from a set of virtual screening hits

SYNTHESIS OF NEW DERIVATIVES OF 2-ACYLISOTHIOCYANATE OF 1-NITRO-9,10-ANTHRAQUINONE WITH ANTIMICROBIAL ACTIVITY

УДК 378.147:547

Комбинаторная химия в высшей школе: десятилетний опыт научных, учебных и организационных проектов

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry 45 (2010) 2606–2612

European Journal of Medicinal Chemistry

УДК 378.147:547

Комбинаторная химия в высшей школе: десятилетний опыт научных, учебных и организационных проектов

Overview

1. Historical reflections
2. PASS (Prediction of Activity Spectra for Substances)
3. PharmaExpert
4. GUSAR (General Unrestricted Structure-Activity Relationships)
5. Some examples of applications
6. Our web-services based on PASS, GUSAR, etc.
7. **Way2Drug.com: further progress**

Integration of all web-services (I)

Our Services



PASSOnline
biological potential of your
compounds ([more](#))



Gusar Online
create QSAR/QSPR models
([more](#))



SMP
prediction substrate/metabolite
specificity ([more](#))



DIGEP Pred
drug-induced changes of gene
expression ([more](#))



Meta-Pred
in silico prediction of sites of
metabolism ([more](#))

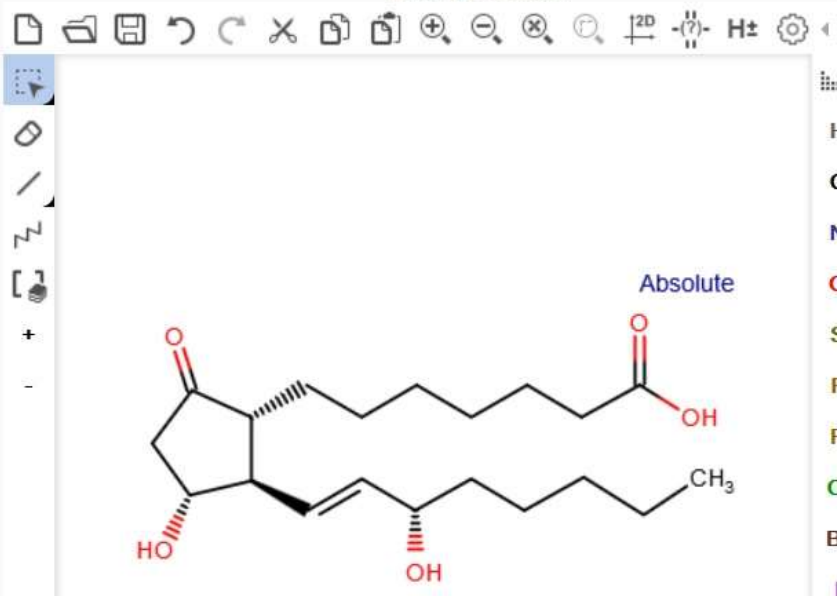


CLC-Pred
in silico prediction of cytotoxicity


Predict activities/properties of your compound


To receive results, please, enter:


Draw a structure:





Integration of all web-services (II)


 **DIGEP Pred**
drug-induced changes of gene expression ([more](#))


 **Meta-Pred**
in silico prediction of sites of metabolism ([more](#))

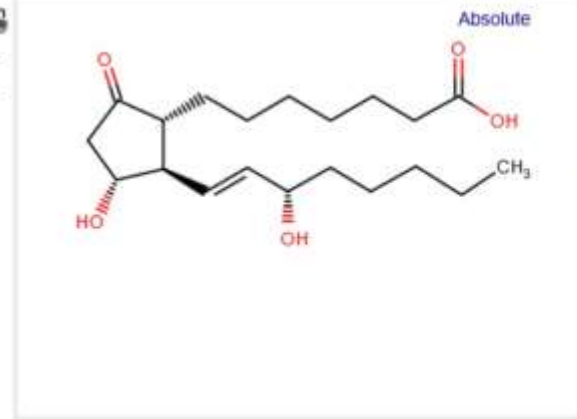
 **CLC-Pred**
in silico prediction of cytotoxicity for tumor and normal cell-lines ([more](#))

 **B.B.B. Prediction**
calculate the permeability of blood-brain barrier ([more](#))

 **SAR Creator**
obtain SDF files with structure and data information for sets of compounds ([more](#))

 **Aurveda**
mechanisms of action and pharmacological effects in Ayurveda ([more](#))

 **SPrOS**
is developed to analyze the amino acid sequences related to the same protein family ([more](#))



Absolute

O

S

F

P

Cl

Br

I

+

A

Choose activities/properties which you want to predict:

Select/unselect all

PASS_Online-All_Activities

Pass_Online-Effects

Pass_Online-Mechanism

Pass_Online-Metabolism

Pass_Online-Transport

PASS_Online-Adverse_Effects&Toxicity

SOMP

GUSAR-antitarget

GUSAR-Acute_Rat_Toxicity

GUSAR-Environmental_Toxicity

DIGEP_Pred-mRNA-Level

DIGEP_Pred-Protein-Level

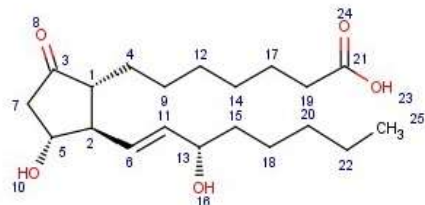
LogBB

And finally:

• [Click here to predict](#)

Integration of all web-services (III)

PASS_Online-All_Activities Web Server prediction results



Pa	Pi	Activity
0,980	0,001	Vasoprotector
0,980	0,001	Antisecretoric
0,979	0,001	Prostaglandin-E2 9-reductase inhibitor
0,976	0,002	Vasodilator, peripheral
0,975	0,001	GST A substrate
0,972	0,002	Mucomembranous protector
0,948	0,001	UGT2B1 substrate
0,944	0,002	Glutathione S-transferase substrate
0,937	0,001	CYP4A11 substrate
0,939	0,003	Vasodilator
0,934	0,002	Antithrombotic
0,933	0,001	Morphine 6-dehydrogenase inhibitor
0,927	0,001	Leukotriene-B4 20-monooxygenase inhibitor
0,923	0,000	CYP4F substrate
0,920	0,002	Antiulcerative
0,921	0,004	CYP2E1 substrate
0,914	0,001	Cytoprotectant
0,915	0,004	CYP2E substrate
0,913	0,001	GST M substrate
0,909	0,001	Gastrin inhibitor
0,912	0,006	Mucositis treatment
0,901	0,000	Oxytocic
0,901	0,002	CYP4A substrate
0,907	0,005	Anticanceratic

...

Computer-aided analysis of hidden potential in traditional Indian medicine Ayurveda



INDIAN-RUSSIAN JOINT RESEARCH PROJECT
COMPUTER-AIDED STUDY OF HIDDEN POTENTIAL IN TRADITIONAL INDIAN MEDICINE AND ITS PHARMACOLOGICAL VALIDATION

Natural compounds is that what you do need.

Ayurveda
Ayurvedic medicine is a system of traditional medicine native to India and a form of alternative medicine.

What is the purpose of this project?
The purpose of this project is to analyze the mechanisms of action and pharmacological effects of individual components and combinations of the 50 medicinal plants used in Ayurveda, based on computer prediction of biological activity spectra of individual compounds using the program PASS, and to assess their drug-drug interactions using PharmaExpert. The information will be used to identify the hidden potential of traditional Indian medicine, and to validate some computer-aided predictions in biological assays.

Natural compounds are used in folk medicine for thousands of years, now occupying over 30% of the world pharmaceutical market. They have a high chemical diversity in comparison with substances obtained by synthetic, but only a small part of their pharmacological potential is used by medicine. The vast amount of empirical data on the pharmacological properties of natural compounds accumulated in traditional Indian medicine (TAM) Ayurveda.

SUPPORTED BY: RFBR, IBMC, Department of Science & Technology

Email: info@ibmc.rsu.lv © 2011 Project

Natural products are used in folk medicine since many thousands year. They represent a significant, though often underappreciated resource for the development of new medicines.

Contents:

- (1) 50 medicinal plants;
- (2) structural formulae of 1906 phytochemicals;
- (3) biological activity of 288 phytochemicals.

Criteria:

- (1) Ayurvedic /traditional medicinal use;
- (2) adequately explored for phytochemical analysis;
- (3) unexplored for pleiotropic pharmacological studies.

New Features in PharmaExpert adopted to analysis of phytoconstituents of medicinal plants

Prediction & Interpretation - C:\Base\RF\India 2011-2012\11.12\Plants\Aloe vera (Xanthorrhoeaceae) SDF 1/31

Isosakuranin
Aloesinol (2'R)-form, 7-Me ether, 2'-O-Aloesinol (2'R)-form, 1'-Hydroxy, 7-Me Aloesinol (2'R)-form, 1'-Hydroxy, 7-Me Aloesinol (2'R)-form, 1'-Hydroxy, 7-Me Aloesinol (2'S)-form
Aloesinol (2'S)-form, 7-Me ether
Aloesinol (2'S)

Save TXT Save SD Clipboard Exclude

Pa	Pi	Types of Activities	Pa/Pi descending
0.838	0.005	Cytotoxic	
0.825	0.004	UGT1A9 substrate	
0.820	0.003	Heat shock protein 27 antagonist	
0.841	0.029	Membrane integrity agonist	
0.803	0.004	Cell adhesion molecule inhibitor	
0.825	0.029	CD11 glycosyl glycerolphosphatase inhibitor	
0.795	0.018	Glucuronate 2-dehydrogenase (acceptor) inhibitor	
0.796	0.021	Hepatotoxic	
0.791	0.032	Toxic	
0.762	0.005	Antimutagenic	
0.795	0.005	Hepatoprotectant	
0.778	0.039	Diathesis	
0.777	0.039	Ulcer, esophageal	
0.729	0.004	Free radical scavenger	
0.764	0.052	PhosphodiacylM-methyltransferase N-methyltransferase inhibitor	
0.734	0.033	Sugar-phosphatase inhibitor	
0.709	0.009	Anticarcinogenic	
0.702	0.018	Vasoprotective	
0.712	0.033	Berberate-CoA ligase inhibitor	
0.701	0.026	Fibrinolytic	
0.708	0.091	Transferrase inhibitor	
0.665	0.012	Antitumor	
0.706	0.059	Hemolytic	
0.699	0.013	UGT1A substrate	
0.700	0.062	Allergic reaction	
0.694	0.018	Inflammation	
0.677	0.051	Hemolytic	
0.668	0.033	Phosphatase inhibitor	
0.625	0.008	UGT1A1 substrate	

Substance that binds to interleukin 2 receptor and stimulates its function or stimulates formation or release of interleukin 2.

Antitarget: 7 Metabolism: 31 Gene Expression: 2 Toxicity: 59

Effect: 75	Mechanism: 133	KEGG	NCI Pathway	Reactions	Therapeutic effects
Caspase 8 stimulant 0.441 0.004	Levorotatory inhibitor 0.439 0.002	0.424 0.026	DNA polymerase I inhibitor	AP1 transcription factor network NF-1-alpha transcription factor network Notch-mediated HES/HEY network	Anabolic Antigenic Antigenic, non-opioid Anticancer Antianginal Antiallergic Antibacterial Antidiabetic Antidiabetic (type 1) Antidiabetic (type 2) Antifibrinolytic
Hypoxia inducible factor 1 alpha inhibitor 0.437 0.004	Immunosuppressant 0.436 0.007	0.401 0.062	Interleukin 2 agonist	AP1 transcription factor network IL12 signaling mediated by STAT4 IL2 signaling events mediated by STAT5 SH2 signaling	Side effects Cardiotoxic Neurotoxic
Histamine release stimulant 0.434 0.073	Glucan endo-1,3-beta-D-glucosidase inhibitor 0.432 0.089	0.380 0.246	Smooth muscle myosin light chain kinase inhibitor	Regulation of telomerase Downstream signaling in natural vs CD8+ T cells IL2 signaling events mediated by PI3K IL27-mediated signaling events Calcium signaling in the CD4+ TCR pathway Calcium-regulated NFAT-dependent transcription in lymphocyte IL12-mediated signaling events IL2-mediated signaling events Glucocorticoid receptor regulatory network	
DNA polymerase I inhibitor 0.424 0.026	Ribulose-phosphate 3-epimerase inhibitor 0.423 0.123	0.353 0.001	Caspase 3 stimulant	Calcium-regulated NFAT-dependent transcription in lymphocyte GTPNH-mediated signaling Posttranslational regulation of adherens junction stability and dynamics Caspase Cascade in Apoptosis FAS (CD95) signaling pathway Role of Calcium-dependent NFAT signaling in lymphocytes shp1 Med. Negative effector of Fas and TNF-alpha LFA receptor mediated events Syndecan 2-mediated signaling events	
Hydrolase inhibitor 0.430 0.163	DNA polymerase II inhibitor 0.419 0.021				
Adenosine regulator 0.418 0.021	Lactate dehydrogenase stimulant 0.401 0.136				
Leukine 0.416 0.005	Interleukin 2 agonist 0.401 0.056				
Glycosyltransferase inhibitor 0.416 0.004	Antidiabetic 0.563 0.010				
Maltose transporting ATPase inhibitor 0.411 0.013	Antineoplastic (lymphoma) 0.513 0.100				
Herceptin 0.409 0.018	Antineoplastic 0.323 0.121				
Glutathione disulfide reductase inhibitor 0.409 0.009	Antifolate (Hepatitis B) 0.315 0.020				
Lipid peroxidase inhibitor 0.407 0.018	Major histocompatibility complex class II beta N-acetylglucosaminidase inhibitor 0.397 0.058				
Lactate dehydrogenase stimulant 0.401 0.136	Transcription factor stimulant 0.387 0.058				
Transcription factor NF-1 inhibitor 0.387 0.058	Transcription factor NF-1 inhibitor 0.387 0.058				

UniProt ID: P04911, P02658
Gene name(s): IL2
Species: Mus musculus, Homo sapiens

Number of selected compounds: 48
Substructure descriptors: 1 new, 286 Possible activities.

More information could be found in our joint publications

Med Chem Res (2011) 20:1509–1514
DOI 10.1007/s00044-010-9398-y

ORIGINAL RESEARCH

PASS-assisted exploration of new therapeutic potential of natural products

Rajesh Kumar Goel · Damanpreet Singh · Alexey Lagunin · Vladimir Poroikov

Received: 15 March 2010 / Accepted: 22 July 2010 / Published online: 6 August 2010
© Springer Science+Business Media, LLC 2010

Abstract The use of drug substances derived from plants, fungi, bacteria, and marine organisms are “Mother Nature Gift” for diseases of mankind. Many of these are discovered serendipitously and have a long tradition in medicine. Till date, the use of natural products, their semisynthetic and synthetic derivatives have been mostly confined to their ethnic use. But it has been well known that each

Keywords Ayurveda · Biological activity spectrum · Herbal drugs · Natural products · PASS

Introduction Natural products (NPs) are used in folk medicine since

Chemo- and Bioinformatics resources and *in silico* approaches for drug discovery from Plants used in Traditional Indian Medicine: A Critical Review.

Lagunin A.A.¹, Goel R.K.², Gawande D.Y.², Pahwa P.², Glorizova T.A.¹, Dmitriev A.V.¹, Ivanov S.A.¹, Rudik A.V.¹, Konova V.I.¹, Pogodin P.V.^{1,3}, Druzhilovsky D.S.¹, Poroikov V.V.^{1,3}

¹ Orekhovich Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., 119121, Pogodinskaya Str. 10/8, Moscow, Russia. alexey.lagunin@ibmc.msk.ru

² Department of Pharmaceutical Sciences and Drug Research, Punjabi University, 147002, India goelrkpup@gmail.com

³ Russian National Research Medical University, department of Biochemistry of Biological Faculty, 117997, Ostrovitianov str. 1, Moscow, Russia.

European Journal of Pharmacology 704 (2013) 33–40

Contents lists available at ScienceDirect

European Journal of Pharmacology

journal homepage: www.elsevier.com/locate/ejphar

Behavioural pharmacology

Ameliorative effect of Curcumin on seizure severity, depression like behavior, learning and memory deficit in post-pentylenetetrazole-kindled mice

Kailash M. Choudhary^a, Awanish Mishra^a, Vladimir V. Poroikov^b, Rajesh Kumar Goel^{a*}

^a Department of Pharmaceutical Sciences and Drug Research, Punjab University, Patiala-147002, Punjab, India
^b Institute of Biomedical Chemistry, Russian Acad. Med. Sci., Moscow, Russia

ARTICLE INFO

Article history:
Received 20 August 2012
Received in revised form 1 February 2013
Accepted 7 February 2013
Available online 24 February 2013

Keywords:
Curcumin
PTZ kindling
epileptic comorbidity
Depression
Learning and memory deficit

Компьютерная оценка скрытого потенциала фитоконпонентов лекарственных растений из традиционной индийской медицины Аюрведа

Лагунин А.А.¹, Дружилевский Д.С.¹, Рудик А.В.¹, Филимонов Д.А.¹, Gawande D.², Suresh K.², Goel R.², Пороиков В.В.¹

¹ Федеральное государственное бюджетное учреждение «Научно-исследовательский институт имени В.Н.Ореховича» Российской академии медицинских наук 119121 Москва, ул. Погодинская, 10/7, тел.: 7 (499) 245-09-20, факс: 7 (499) 245-08-57, e-mail: vladimir.poroikov@ibmc.msk.ru
² Department of Pharmaceutical Sciences and Drug Research, Punjabi University, Patiala-147002, India, tel.: 91 (175) 304-62-54, fax: 91 (175)228-30-73, E-mail: goelrkpup@gmail.com

Реферат
С целью изучения скрытого потенциала традиционной индийской медицины Аюрведа создан веб-

NPR

REVIEW



Cite this: Nat. Prod. Rep., 2014, 31, 1585

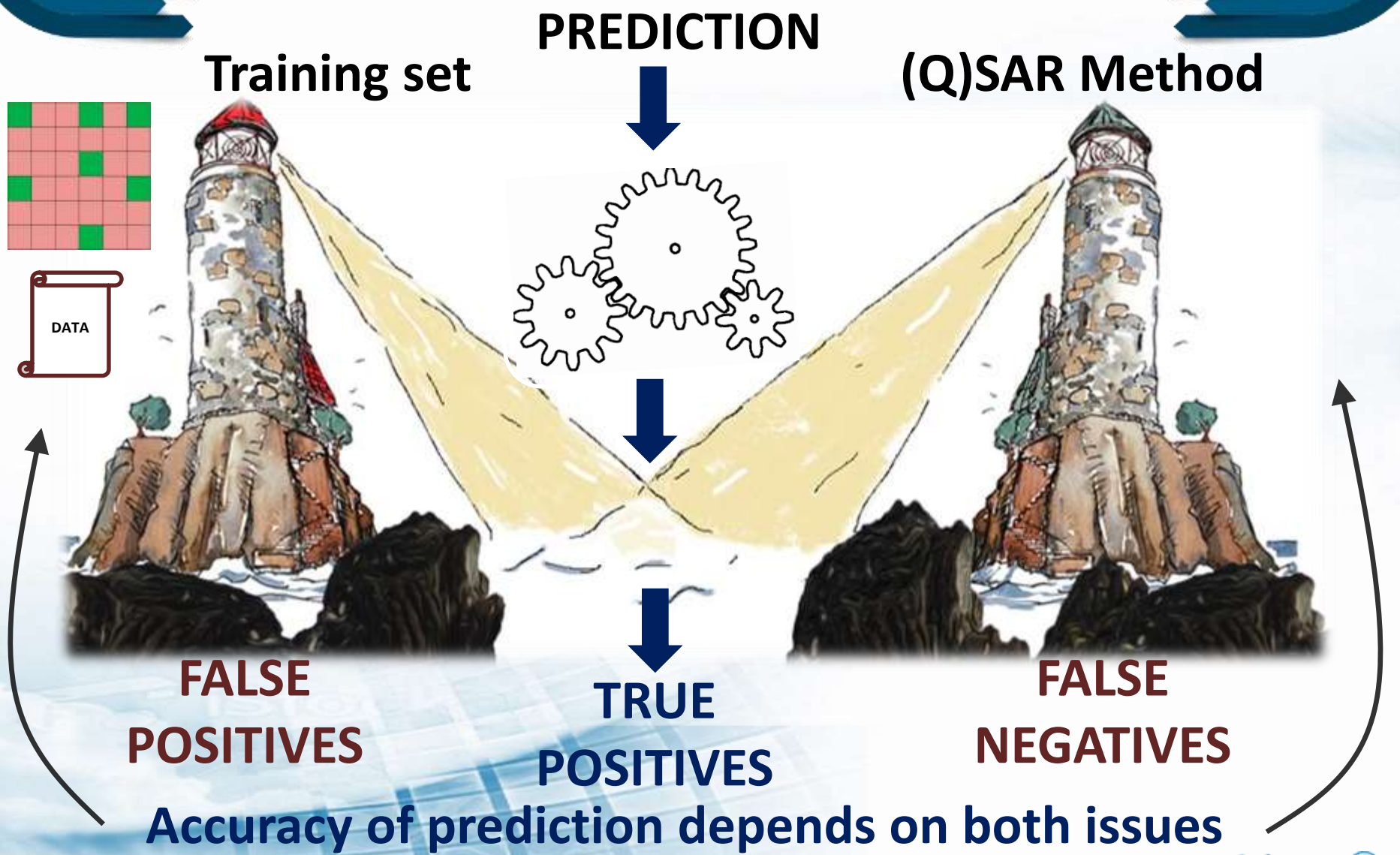
Chemo- and bioinformatics resources for *in silico* drug discovery from medicinal plants beyond their traditional use: a critical review†

Alexey A. Lagunin,^{a,b,c} Rajesh K. Goel^{*b}, Dinesh Y. Gawande,^b Priyanka Pahwa,^b Tatyana A. Glorizova,^a Alexander V. Dmitriev,^a Sergey M. Ivanov,^a Anastassia V. Rudik,^a Varvara I. Konova,^a Pavel V. Pogodin,^{a,c} Dmitry S. Druzhilovsky^a and Vladimir V. Poroikov^{a,b,c}

Covering: up to 2014



(http://ayurveda.pharmexpert.ru). В данных введена информация о 50 индийской медицине и входящих в их состав 288 фитоконпонентов. биологической активности 946 видами биологической активности и о выборку компьютерной программы цифрированной обучающей выборки и м по одному и кросс-валидация запас, что значения средней ошибки от значений, полученных при 5,395%, соответственно), что специализированной версии программы базис PASS получен прогноз спектров 0 лекарственных растений ТИМ. С н анализ результатов прогноза для ТИМ; для ряда растений проведено ми экстрактов из лекарственных фармакологически не эффекты



Study on quality of data in publicly and commercially available databases

JOURNAL OF
**CHEMICAL INFORMATION
AND MODELING**

Article

pubs.acs.org/jcim

QSAR Modeling Using Large-Scale Databases: Case Study for HIV-1 Reverse Transcriptase Inhibitors

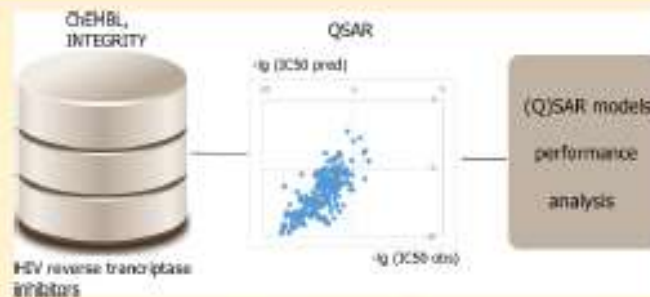
Olga A. Tarasova,^{*,†} Aleksandra F. Urusova,[†] Dmitry A. Filimonov,[†] Marc C. Nicklaus,[‡] Alexey V. Zakharov,[‡] and Vladimir V. Poroikov[†]

[†]Institute of Biochemical Chemistry, 10-8, Pogodinskaya St., 119121, Moscow, Russia

[‡]CADD Group, Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, DHHS, NCI-Frederick, 376 Boyles St., Frederick, Maryland 21702, United States

Supporting Information

ABSTRACT: Large-scale databases are important sources of training sets for various QSAR modeling approaches. Generally, these databases contain information extracted from different sources. This variety of sources can produce inconsistency in the data, defined as sometimes widely diverging activity results for the same compound against the same target. Because such inconsistency can reduce the accuracy of predictive models built from these data, we are addressing the question of how best to use data from publicly and commercially accessible databases to create accurate and predictive QSAR models. We investigate the suitability of commercially and publicly available databases to QSAR modeling of antiviral activity (HIV-1 reverse transcriptase (RT) inhibition). We present several methods for the creation of modeling (i.e., training and test) sets from two, either commercially or freely available, databases: Thomson Reuters Integrity and ChEMBL. We found that the typical predictivities of QSAR models obtained using these different modeling set compilation methods differ significantly from each other. The best results were obtained using training sets compiled for compounds tested using only one method and material (i.e., a specific time



of QSAR models obtained using these different modeling set compilation methods differ significantly from each other. The best results were obtained using training sets compiled for compounds tested using only one method and material (i.e., a specific time

Summary

- Based on long-term projects in chemoinformatics (local correspondence concept, PASS, PharmaExpert, GUSAR, etc.), we have developed several web-services useful in computer-aided drug discovery.
- These web-services are widely used by many researchers from over 90 countries; more than 300 papers published with citations of our web-resources.
- Further development of these resources requires integration, curation of the information, improvement of functionality, etc.
- Active cooperation with the Way2Drug users will significantly enhance our initiatives.

Acknowledgements to the key persons and to the financial support

Tatyana Glorizova, M.Sc.



Alexey Lagunin, Dr. Sci.



Dmitry Filimonov, Ph.D.



Dmitry Druzhilovskiy, Ph.D.



Alexey Zakharov, Ph.D.



And to many other colleagues who participate(d) in our projects



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

РОССИЙСКИЙ НАУЧНЫЙ ФОНД
ПОДДЕРЖКА И РАЗВИТИЕ



Sixth Framework Programme
2002 - 2006

RESEARCH & INNOVATION
FP7





Way2Drug | PREDICTIVE
SERVICES
Understanding Chemical-Biological Interactions

Thank you for your kind attention!

We are open for collaboration; please, address your suggestions to:
vladimir.poroikov@ibmc.msk.ru; vvp1951@yandex.ru