PREDICTION OF ACTIVITY SPECTRA FOR SUBSTANCES: TWENTY YEARS OF THE DEVELOPMENT

Vladimir Poroikov, Prof. Dr.

Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., Pogodinskaya Str. 10/8, Moscow, 119121, Russia
http://www.ibmc.msk.ru
The Mission:
Better Medicines through Global Education and Research

http://www.iuphar.org/
Disease

Target(s)
> 2,000 targets (Thomson Reuters Integrity)

Ligand(s)
~ 14,400 codes (ICD-10)
~60 M samples (ChemNavigator)
~166 B virtual (JCIM, 2012, p.56)
~10^{120} possible (JCICS, 2003, p.374)
Changing of paradigm: from «target-centric» approach to the analysis of signal transduction regulatory networks & pathways

XX century
Disease $\rightarrow$ Target $\rightarrow$ Drug

XXI century
Multitargeted drugs

Due to biological activity, chemical compound may be used as a medicine for treatment of certain disease. Due to biological activity, chemical compound may cause adverse or toxic effects in human.
How to estimate biological activity spectra of pharmacological agents?

Clinical trials:
- Risks for a human health.

Pre-clinical studies:
- Time-consuming and expensive.

Computational prediction:
- Decreasing of risks, time and cost.
- Can be done for virtual structures.
Outline

1. Computational approaches to biological activity prediction.

2. NetFlowEx (targets identification)


5. GUSAR: General Unrestricted Structure-Activity Relationships.

6. Some applications.

7. Web-services.

8. Summary.
Outline

1. Computational approaches to biological activity prediction.

2. NetFlowEx (targets identification)


5. GUSAR: General Unrestricted Structure-Activity Relationships.

6. Some applications.


8. Summary.
Approaches to discover new pharmaceutical agents

- 3D structure of the target and structural formulae of ligands are known.
  - Structure-based design

- 3D structure of the target is unknown but structural formulae of ligands are known.
  - Ligand-based design

- 3D structure of the target is known but structural formulae of ligands are unknown.
  - De-novo design

- 3D structure of the target and structural formulae of ligands are unknown.
  - Combinatorial chemistry and HTS

Target-based approaches to prediction of biological activity

**Prerequisites:**

✓ Data about 3D structure of target macromolecule (X-ray, NMR, Modeling).
✓ Data about 3D structure of active site (binding site).
✓ Docking and estimation of binding energy (scoring function).
✓ Active site mapping and *de novo* design.
Limitations of Target-Based Drug Design Methods

✓ 3D structure of the target is necessary.

✓ 3D structure in crystal vs. 3D structure in solution.

✓ Approximation of energy binding estimates.

✓ Approximation of 3D conformation of flexible ligands.
The simplest Ligand-Based Drug Design methods are based on a similarity principle: “Me-too-drugs” design.

**FIGURE 6.1** Angiotensin AT1 receptor antagonists derived from losartan. Despite their structural similarity of the structures, it can be assumed that the corresponding discoveries were made independently. The first year under parentheses is the basic patent year, the second one is the year of the first launch.
Quantitative estimation of similarity

Tanimoto Coefficient of similarity for Molecules A and B:

\[ S_{ab} = \frac{c}{a + b - c} \]

Where:
- \( a \) = bits set to 1 in A,
- \( b \) = bits set to 1 in B,
- \( c \) = number of 1 bits common to both

Range is 0 to 1.

Value of 1 does not mean the molecules are identical.
Similarity search in ChemNavigator library

**Simple Sketcher (JME) Query Entry**

**Please select** database collections to search.

**Search Type**
Find compounds:

- [ ] > 80 % similar to this one
- [ ] Containing this as a substructure

**Search Options**
- Include Sample Duplicates
- Allow substitution at all H atoms
- Fill valences with H atoms

No more than 20 hits

New List name: MyHits

Search

**Acetylsalicylate**

http://www.chemnavigator.com
The most similar compounds for Acetylsalicylate

1. **CNC-315394159**
   - **Structure ID:** 180927362
   - **Collection:** Archived Compounds
   - **Matched by:** Similarity: 99%
   - **Major MW:** 183.18
   - **cLogP:** 1.38
   - **Min Purity:** 90
   - **Shipping Window:** 30 Days
   - **TC=99%**

2. **CNC-315419843**
   - **Structure ID:** 79937939
   - **Collection:** Archived Compounds
   - **Matched by:** Similarity: 87%
   - **Major MW:** 194.19
   - **cLogP:** 1.86
   - **Min Purity:** 90
   - **Shipping Window:** 30 Days
   - **TC=87%**

3. **CNC-310472181**
   - **Structure ID:** 69154554
   - **Collection:** Virtual Custom Chemistry
   - **Matched by:** Similarity: 86%
   - **Major MW:** 300.31
   - **cLogP:** 4.15
   - **Min Purity:** 92
   - **Shipping Window:** 60 Days
   - **TC=86%**

4. **CNC-310472239**
   - **Structure ID:** 69154582
   - **Collection:** Virtual Custom Chemistry
   - **Matched by:** Similarity: 84%
   - **Major MW:** 270.28
   - **cLogP:** 3.72
   - **Min Purity:** 92
   - **Shipping Window:** 60 Days
   - **TC=86%**

5. **CNC-308281405**
   - **Structure ID:** 36090778
   - **Collection:** Aldrich Market Select Screening Compounds
   - **Matched by:** Similarity: 84%
   - **Major MW:** 242.27
   - **cLogP:** 4.12
   - **Min Purity:** 92
   - **Shipping Window:** 14 Days
   - **TC=84%**

6. **CNC-310472240**
   - **Structure ID:** 69154583
   - **Collection:** Virtual Custom Chemistry
   - **Matched by:** Similarity: 84%
   - **Major MW:** 270.28
   - **cLogP:** 3.72
   - **Min Purity:** 92
   - **Shipping Window:** 60 Days
   - **TC=84%**
...“there is only a 30% chance that a compound that is > 0.85 (Tanimoto) similar to an active is itself active”.

“The best material model for a cat is another [cat], or preferably the same cat".

Ligand-based approaches to prediction of biological activity

Prerequisites:

Set of ligands with known biological activity (training set).

Methods:

(Quantitative) Structure-Activity Relationships (Q)SAR, Pharmacophore models.
Outline

1. Computational approaches to biological activity prediction.

2. NetFlowEx (targets identification)


5. GUSAR: General Unrestricted Structure-Activity Relationships.

6. Some applications.


8. Summary.
Dichotomous modeling of regulatory networks by NetFlowEx program for identification of antitumor targets

Primary states

\[ F_i (S_1, S_2, \ldots, S_n) = \Theta(ai + \sum_k S_kb_{ik}) \]

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.

Simulation of tumor cell division

Generalized breast cancer

Steps of trajectory

Inactive
Active

Proteins regulating cell cycle and apoptosis

Proteins regulating cell cycle and apoptosis

Cell cycle complexes

Cell cycle regulatory proteins

Apoptotic proteins

## Some Double and Triple Targets’ Combinations Identified For Breast Cancer

<table>
<thead>
<tr>
<th>No</th>
<th>Target 1</th>
<th>Target 2</th>
<th>Target 3</th>
<th>Number of compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bcl2 antagonist</td>
<td>Cyclin-dependent kinase 2 inhibitor</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Bcl2 antagonist</td>
<td>Myc inhibitor</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Bcl2 antagonist</td>
<td>Phosphatidylinositol 3-kinase beta inhibitor</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Cyclin-dependent kinase 2 inhibitor</td>
<td>Myc inhibitor</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Hypoxia inducible factor 1 alpha inhibitor</td>
<td>Myc inhibitor</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>Hypoxia inducible factor 1 alpha inhibitor</td>
<td>Phosphatidylinositol 3-kinase beta inhibitor</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>Myc inhibitor</td>
<td>Phosphatidylinositol 3-kinase inhibitor</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>Bcl2 antagonist</td>
<td>Myc inhibitor</td>
<td>Phosphatidylinositol 3-kinase beta inhibitor</td>
<td>10</td>
</tr>
</tbody>
</table>
Outline

1. Computational approaches to biological activity prediction.

2. NetFlowEx (targets identification)


5. GUSAR: General Unrestricted Structure-Activity Relationships.

6. Some applications.


8. Summary.
Prediction of Activity Spectra for Substances
The key persons in PASS development

Dmitry Filimonov

Tatyana Gloriozova

Vladimir Poroikov

Alexey Lagunin
## PASS 2012 Characteristics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training Set</strong></td>
<td>313,345 drugs, drug-candidates, pharmacological and toxic substances comprise the training set</td>
</tr>
<tr>
<td><strong>Biological Activity</strong></td>
<td>6400 biological activities can be predicted (Active vs. Inactive)</td>
</tr>
<tr>
<td><strong>Chemical Structure</strong></td>
<td>Multilevel Neighborhoods of Atoms (MNA) descriptors [1, 2]</td>
</tr>
<tr>
<td><strong>Mathematical Algorithm</strong></td>
<td>Bayesian approach was selected by comparison of many different methods [2]</td>
</tr>
<tr>
<td><strong>Validation</strong></td>
<td>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]</td>
</tr>
</tbody>
</table>

Components of biological activity spectra predicted by PASS

- **Main pharmacological effects**
  (antihypertensive, hepatoprotective, anti-inflammatory etc.);

- **Mechanisms of action**
  (5-HT1A 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.).
Multilevel Neighborhoods of Atoms (MNA) Descriptors

MNA/0:  C

MNA/1:  C(CN-H)

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

Algorithm for 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 \( D_1, ..., D_m \) are independent, then:

\[
P(S|A) = P(D_1, ..., D_m|A) = \prod_i P(D_i|A)
\]

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

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

\[
P(A) = \frac{\sum_i \sum_k g_k(D_i)w_k(A)}{\sum_i \sum_k g_k(D_i)}
\]
How PASS predicts biological activity spectrum?

1. Structural formula of new compound
2. Estimating of probability for each particular biological activity
3. Predicted biological activity spectrum

Pa – probability to be active, Pi – probability to be inactive
Example of prediction for Clopidogrel

6 of 463 Possible Pharmacological Effects at Pa > 0.500

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
Some publications, where PASS algorithm is described


18977 compounds with 124 activities were selected from MDDR.

The set of compounds was 50 times divided at random into two equal subsets.

The first subset was used as the training set, the second one as the evaluation subset and vice versa (100 experiments).

20, 40, 60, 80% of information (activity/structure data) were excluded from the training set.

Average accuracy of prediction (IAP) was calculated for each type of activity.
Robustness of Biological Activity Spectra Predicting by Computer Program PASS for Noncongeneric Sets of Chemical Compounds

V. V. Poroikov, * † D. A. Filimonov, † Yu. V. Borodina, † A. A. Lagunin, † and A. Kos ‡

Institute of Biomedical Chemistry RAMS, Pogodinskaya Str., 10, Moscow 119832, Russia, and AKos Consulting & Solutions GmbH, Rossligasse 2, CH-4125 Riehen, Switzerland

Received March 1, 2000

The computer system PASS provides simultaneous prediction of several hundreds of biological activity types for any drug-like compound. The prediction is based on the analysis of structure–activity relationships of the training set including more than 30000 known biologically active compounds. In this paper we investigate the influence on the accuracy of predicting the types of activity with PASS by (a) reduction of the number of structures in the training set and (b) reduction of the number of known activities in the training set. The compounds from the MDDR database are used to create heterogeneous training and evaluation sets. We demonstrate that predictions are robust despite the exclusion of up to 60% of information.

INTRODUCTION

Traditional QSAR and 3D molecular modeling are successful at predicting the biological activities for chemical structures, provided they work with small number of types of activity and usually stay in the same chemical series. 1–5 Similarity searching 6,7 and clustering methods 7,8 can be used to separate compounds into structural groups 9 and for the prediction of biological activities and compound selection. 10

Table 1. Some Predicted Biological Activities for Cavinton

<table>
<thead>
<tr>
<th>no.</th>
<th>Pa</th>
<th>Pi</th>
<th>activity</th>
<th>expt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.929</td>
<td>0.004</td>
<td>peripheral vasodilator</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.900</td>
<td>0.000</td>
<td>multiple sclerosis treatment</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.855</td>
<td>0.005</td>
<td>vasodilator</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.844</td>
<td>0.003</td>
<td>abortion inducer</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>0.812</td>
<td>0.001</td>
<td>antineoplastic enhancer</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>0.760</td>
<td>0.006</td>
<td>coronary vasodilator</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>0.732</td>
<td>0.007</td>
<td>spasmogenic</td>
<td></td>
</tr>
</tbody>
</table>
PASS  Professional Training Procedure: New Base & Add

PASS

File  Base  Predict  View  Options  Help

- New Base
- Open Base...
- Save Base
- Save Base As...

Passage:

SAR Base Information

- Substances: 100  Modified
- Descriptors: 1092  Modified
- Activity Types: 54  Modified
- Selected Activity Types: 0
- Average IEP
- Prediction

Disabled
PASS Professional Training Procedure: Training
### PASS Professional Training Procedure: Selection

#### Select Activity Types to be Predicted

<table>
<thead>
<tr>
<th>Predictable Activity Type</th>
<th>Group</th>
<th>Number</th>
<th>IEP, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleotide metabolism regulator</td>
<td>M</td>
<td>4</td>
<td>10.417</td>
</tr>
<tr>
<td>Antiulcer, non-allergic</td>
<td>E</td>
<td>3</td>
<td>3.966</td>
</tr>
<tr>
<td>Psychotropic</td>
<td>E</td>
<td>11</td>
<td>9.806</td>
</tr>
<tr>
<td>HERG channel blocker</td>
<td>MA</td>
<td>5</td>
<td>8.211</td>
</tr>
<tr>
<td>Potassium channel (Voltage-sensitive) blocker</td>
<td>M</td>
<td>5</td>
<td>8.211</td>
</tr>
<tr>
<td>Antibacterial</td>
<td>E</td>
<td>15</td>
<td>8.157</td>
</tr>
<tr>
<td>Atherosclerosis treatment</td>
<td>E</td>
<td>5</td>
<td>7.573</td>
</tr>
<tr>
<td>QT interval prolongation</td>
<td>T</td>
<td>3</td>
<td>7.560</td>
</tr>
<tr>
<td>Diuretic</td>
<td>E</td>
<td>10</td>
<td>7.558</td>
</tr>
<tr>
<td>Antiallergic</td>
<td>E</td>
<td>18</td>
<td>7.523</td>
</tr>
<tr>
<td>Unused Activity Type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potassium channel blocker</td>
<td>M</td>
<td>6</td>
<td>16.135</td>
</tr>
<tr>
<td>Histamine H1 receptor antagonist</td>
<td>MA</td>
<td>3</td>
<td>15.464</td>
</tr>
<tr>
<td>Histamine antagonist</td>
<td>M</td>
<td>3</td>
<td>15.464</td>
</tr>
<tr>
<td>Antihistaminic</td>
<td>E</td>
<td>3</td>
<td>15.464</td>
</tr>
<tr>
<td>Cardiotoxic</td>
<td>T</td>
<td>4</td>
<td>13.802</td>
</tr>
<tr>
<td>Antibiotic</td>
<td>E</td>
<td>6</td>
<td>13.298</td>
</tr>
<tr>
<td>Antinflammatory</td>
<td>E</td>
<td>25</td>
<td>13.130</td>
</tr>
<tr>
<td>Antipyretic</td>
<td>E</td>
<td>4</td>
<td>13.021</td>
</tr>
<tr>
<td>Spasmolytic</td>
<td>E</td>
<td>8</td>
<td>11.413</td>
</tr>
</tbody>
</table>

Cyclic AMP phosphodiesterase inhibitor          | M     | 4      | 10.417 |

Selected Activity Types: 32 of 54, Av. IEP, %: 4.875
PASS  Professional Training Procedure: Ready for Prediction
Outline

1. Computational approaches to biological activity prediction.
2. NetFlowEx (targets identification)
5. GUSAR: General Unrestricted Structure-Activity Relationships.
6. Some applications.
8. Summary.
PharmaExpert: Tool for analysis of PASS prediction results
Search for “HIV reverse transcriptase inhibitors AND HIV integrase (strand transfer) inhibitors”
Search for “HIV reverse transcriptase inhibitors AND HIV integrase (strand transfer) inhibitors AND non-mutagenic”
Search for multitargeted compounds using PharmaExpert

**Antihypertensive agents, ACE and NEP inhibitors**

<table>
<thead>
<tr>
<th>No</th>
<th>ID</th>
<th>Structure</th>
<th>Prediction</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>I_{50} ACE (M)</td>
<td>I_{50} NEP (M)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>purified ACE</td>
<td>blood serum</td>
</tr>
<tr>
<td>I</td>
<td>397482</td>
<td><img src="image1" alt="Structure" /></td>
<td>0.655 0.002 NEP inhibitor</td>
<td>1 10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.605 0.003 ACE inhibitor</td>
<td>-</td>
</tr>
<tr>
<td>II</td>
<td>219863</td>
<td><img src="image2" alt="Structure" /></td>
<td>0.649 0.005 NEP inhibitor</td>
<td>1 10^{-3}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.773 0.003 ACE inhibitor</td>
<td>-</td>
</tr>
<tr>
<td>III</td>
<td>119565</td>
<td><img src="image3" alt="Structure" /></td>
<td>0.684 0.003 NEP inhibitor</td>
<td>3 10^{-2}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.569 0.004 ACE inhibitor</td>
<td>-</td>
</tr>
<tr>
<td>IV</td>
<td>119562</td>
<td><img src="image4" alt="Structure" /></td>
<td>0.618 0.004 NEP inhibitor</td>
<td>5 10^{-3}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.312 0.005 ACE inhibitor</td>
<td>-</td>
</tr>
<tr>
<td>Lisinopril</td>
<td><img src="image5" alt="Structure" /></td>
<td>0.155 0.002 NEP inhibitor</td>
<td>3 10^{-4}</td>
<td>2 10^{-4}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.607 0.004 ACE inhibitor</td>
<td>3 10^{-4}</td>
</tr>
</tbody>
</table>

**Antiinflammatory agents, COX-1, COX-2, LOX inhibitors**

anti-inflammatory (CPE)\(^a\) activity and COX/LOX\(^b\) inhibitory activity

<table>
<thead>
<tr>
<th>compd</th>
<th>CPE%</th>
<th>COX-1</th>
<th>COX-2</th>
<th>LOX</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>57.3 ± 3.4</td>
<td>62.0</td>
<td>0.0</td>
<td>44.0</td>
</tr>
<tr>
<td>2</td>
<td>72.7 ± 6.8</td>
<td>25.0</td>
<td>6.2</td>
<td>51.0</td>
</tr>
<tr>
<td>3</td>
<td>51.1 ± 4.2</td>
<td>8.0</td>
<td>2.5</td>
<td>22.4</td>
</tr>
<tr>
<td>4</td>
<td>66.1 ± 1.2</td>
<td>60.0</td>
<td>4.5</td>
<td>12.5</td>
</tr>
<tr>
<td>5</td>
<td>69.4 ± 2.3</td>
<td>25.0</td>
<td>12.1</td>
<td>76.0</td>
</tr>
<tr>
<td>6</td>
<td>54.2 ± 2.4</td>
<td>31.0</td>
<td>6.2</td>
<td>25.0</td>
</tr>
<tr>
<td>7</td>
<td>44.5 ± 1.8</td>
<td>90.0</td>
<td>30.4</td>
<td>12.0</td>
</tr>
<tr>
<td>8</td>
<td>62.0 ± 2.5</td>
<td>50.0</td>
<td>2.1</td>
<td>44.2</td>
</tr>
<tr>
<td>9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

---

**Computer-Aided Selection of Potential Antihypertensive Compounds with Dual Mechanism of Action**

Alexey A. Lagunin, Oleg A. Gomazkov, Dmitrii A. Filimonov, Tatyana A. Gureeva, Elvira A. Dilakyan, Elena V. Kugaevskaya, Yulia E. Eliseeva, Nina I. Soldyeyeva, and Vladimir V. Porenikov.

Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, Pogodinskaya Street, 10, Moscow 119121, Russia.

Received November 6, 2002.

The prediction of biological activity spectra for substances as an approach for searching compounds with complex mechanisms of action was studied. New compounds with dual mechanisms of antihypertensive action were found by this approach. Biological activity spectra for substances were predicted on the basis of structural formulas by the computer program PASS. Thirty molecular mechanisms of action of compounds from the MDDR 99.2 database, which cause the antihypertensive effect and can be predicted by PASS, have been identified. The analysis of predictions for compounds with 15 dual antihypertensive mechanisms of action from the MDDR 99.2 database confirmed high accuracy of prediction. This approach was applied to databases of commercially available compounds (Ashrix and ChemBridge) and allowed us to select four substances that are potential inhibitors of angiotensin converting enzyme (ACE) and of neutral endopeptidase (NEP). At a later time, all these compounds were found to be the inhibitors of both ACE and NEP. The most potent compounds had I_{50} values of 10^{-7}–10^{-9} M for ACE and 10^{-3}–10^{-5} M for NEP. New combinations of dual mechanisms of action were found before found for antihypertensive compounds were predicted.

---

**Computer-Aided Discovery of Anti-Inflammatory Thiadiazolidines with Dual Cyclooxygenase/Lipoxygenase Inhibition**


Department of Pharmaceutical Chemistry, School of Pharmacy, Aristotle University, Thessaloniki, 54124, Greece, Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, Pogodinskaya Street, 10, Moscow, 119121, Russia, and Medicinal Chemistry Division, Central Drug Research Institute, Chatter Manzil Palace, Lucknow-226 001, India.

Received July 24, 2007.

New anti-inflammatory agents possessing dual cyclooxygenase/lipoxygenase (COX/LX) inhibition were discovered by computer-aided prediction of biological activity for 573 virtually designed chemical compounds. Prediction of biological activity was performed by PASS, and prediction results were analyzed with PharmaExpert software. Nine 2-(thiazole-2-ylaminio)-5-phenylidene-4-thiazolidinone derivatives differing by the phenyl group substitution were selected for synthesis and experimental testing as potential COX/LX inhibitors. Eight tested compounds exhibited anti-inflammatory activity in the carrageenin-induced paw edema. It was shown that seven tested compounds (77.8%) were COX inhibitors, seven compounds were COX inhibitors (77.8%), and six tested compounds (66.7%) were dual COX/LX inhibitors. Analysis of lipophilicity of the compounds showed a negative correlation with inhibition of edema formation. The binding modes of the most active compounds of this series (2-thiazole-2-ylaminio)-5-(m-chlorophenylidene)-4-thiazolidinone for COX-1 and COX-2, and 2-(thiazole-2-ylaminio)-5-(m-nitrophenylidene)-4-thiazolidinone for 15-LOX) were proposed on the basis of docking studies.
The search for new compounds with specific therapeutic effect(s) or/and interaction with specific target(s)

Drug repositioning

The search for new compounds with multiple mechanisms of action

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

PASS & PharmaExpert

J. Med. Chem., 2004, 47(11), 2870-2876
Pharmaceut. Chem. J., 2011, 45 (10), 605-611

Curr. Pharm. Des. 2010, 16(15), 1703-1717
Cardiovascul. Therap. Prof., 2008, 7(5), 100-104

Outline

1. Computational approaches to biological activity prediction.

2. NetFlowEx (targets identification)


5. GUSAR: General Unrestricted Structure-Activity Relationships.

6. Some applications.


8. Summary.
GUSAR: General Unrestricted Structure-Activity Relationships

Synthesis, Antifungal Activity and QSAR study of 2-Arylhydroxynitroindoles

Galina V. Kokurkina a , Mikhail D. Dutov b, *, Svyatoslav A. Shevelev a, Sergey V. Popkov b, Alexey V. Zakharov c and Vladimir V. Poroikov c

a Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation
b Mendeleev University of Chemical Technology of Russia, 125047 Moscow, Russian Federation
c Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, 119121 Moscow, Russian Federation

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted
Available online

ABSTRACT

A series of 2-arylhydroxynitroindoles were prepared and tested for antifungal activity in vitro. The preliminary bioassays indicated that some compounds are comparable to the commercial fungicide (triadimefon). To further explore the structure-activity relationships, the data set of the seventeen structures and their quantitative values of antifungal activities were used for QSAR modeling. Based on the obtained QSAR models four new chemical compounds were designed, synthesized and tested in fungicidal assays. Reasonable correspondence between the experimental and predicted values of antifungal activity was observed.

2011 Elsevier Ltd. All rights reserved.

**QSAR Modelling of antifungal activities**

Table 3. QSAR modeling of antifungal activities results.

<table>
<thead>
<tr>
<th>Activity name</th>
<th>Number of compounds</th>
<th>Number of models</th>
<th>R² training set</th>
<th>Q² training set</th>
<th>R² test set</th>
<th>Coverage</th>
<th>RMSE test</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.s.</td>
<td>12/5</td>
<td>4</td>
<td>0.89</td>
<td>0.72</td>
<td>0.57</td>
<td>80</td>
<td>35.74</td>
</tr>
<tr>
<td>F.m.</td>
<td>12/5</td>
<td>21</td>
<td>0.89</td>
<td>0.77</td>
<td>0.80</td>
<td>100</td>
<td>28.01</td>
</tr>
<tr>
<td>F.o.</td>
<td>12/5</td>
<td>3</td>
<td>0.85</td>
<td>0.68</td>
<td>0.66</td>
<td>100</td>
<td>17</td>
</tr>
<tr>
<td>R.s.</td>
<td>12/5</td>
<td>20</td>
<td>0.91</td>
<td>0.79</td>
<td>0.72</td>
<td>100</td>
<td>27.58</td>
</tr>
<tr>
<td>S.s.</td>
<td>12/5</td>
<td>11</td>
<td>0.89</td>
<td>0.79</td>
<td>0.81</td>
<td>100</td>
<td>37.29</td>
</tr>
<tr>
<td>V.t.</td>
<td>12/5</td>
<td>2</td>
<td>0.83</td>
<td>0.61</td>
<td>0.82</td>
<td>100</td>
<td>20.37</td>
</tr>
</tbody>
</table>

R² – determination coefficient

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

Figure 1. QSAR modeling workflow

Figure 2. Atom contribution into the antifungal activity.
Comparison of computational predictions with the experiment

**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). 1-6, 7-12, 13-18 and 19-24 are activities against B.s., F.m., F.o., R.s., S.s. and V.i., respectively. Average RMSE values calculated for each activity vary from 12 to 25; for each compound – from 12 to 28. All values are given in percent of inhibition at 30 µg mL⁻¹ concentration of the compound.
Outline

1. Computational approaches to biological activity prediction.
2. NetFlowEx (targets identification)
5. GUSAR: General Unrestricted Structure-Activity Relationships.
6. Some applications.
8. Summary.
Drug repositioning: new indications for known drugs

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

Sarah L. Kinnings1,2, Nina Liu2, Nancy Buchmeier3, Peter J. Tonge2, Lei Xie4, Philip E. Bourne4

1Department of Biology, University of York, York, United Kingdom. 2Institute of Chemical Biology & Drug Discovery, Department of Chemistry, Storrs Brook University, New York, United States of America. 3Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California, United States of America. 4San Diego Supercomputer Center, University of California San Diego, La Jolla, California, United States of America.

Abstract

The rise of multi-drug resistant (MDR) and extensively drug resistant (XDR) tuberculosis around the world, including industrialized nations, poses a great threat to human health and defines a need to develop new, effective and inexpensive anti-tubercular agents. Previously, we developed a chemical systems biology approach to identify off-targets of major pharmaceuticals on a proteome-wide scale. In this paper, we further demonstrate the value of this approach through the discovery that existing commercially available drugs, prescribed for the treatment of Parkinson’s disease, have the potential to treat MDR and XDR tuberculosis. These drugs, entacapone and tolcapone, are predicted to bind to the enzyme InhA and

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

V. Poroikov1, A. Akimovb, E. Shabelnikova1 and D. Filimonov1

1Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences, 10, Pogodinskaya Street, Moscow, 119832, Russia. 2Medical and Biological Faculty of the Russian State Medical University, 1, Ostrovitsyana Street, Moscow, 117997, Russia.

(Received 30 June 2011; In final form 31 March 2011)

Computer-aided prediction of the biological activity spectra by the program PASS was applied to a set of 130 pharmaceuticals from the set of the Top 200 medicines. The known pharmacological effects were found in the predicted activity spectra in 93.5% of cases. Additionally, the probability of some supplementary effects was also predicted to be significant, including engagement inhibition, bone formation stimulation, possible use in cognition disorders treatment, multiple sclerosis treatment, etc. These predictions, if confirmed experimentally, may become a cause for a new application of pharmaceuticals from the Top 200 list. Most of known side and toxic effects were also predicted by PASS. PASS predictions at earlier RAID stages may thus provide a basis for finding new “Lead-4” among already licensed drugs and may help direct more attention to those particular effects of pharmaceuticals in clinical use which become apparent only in a small part of the population and require additional precautions.

Keywords: Biological activity spectra; Top 200 medicines; Side effect; Toxicity; Computer-aided predictions; PASS
Prediction of nootropic effect for some antihypertensive drugs

PASS predictions

Mice: behavioral reactions in cross-maze

- Perindopril in dose of 1 mg/kg, and quinapril and monopril in doses of 10 mg/kg improved the patrolling behavior in the maze. This effect is similar to the effects of standard nootropic drugs piracetam and meclofenoxate (in doses of 300 and 120 mg/kg, respectively).

- The observed nootropic effect of these ACE inhibitors is likely to be unrelated to their antihypertensive effect, since the nootropic action took place only at relatively low doses of perindopril, quinapril and monopril and was not observed with further increase in dose.

Pharmacological targets for breast cancer therapy

Breast Cancer Targetscape

Pharmacological targets for breast cancer therapy

Thomson Reuters Integrity
Participants: 9 teams from 8 countries

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

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) and Grant of «Skolkovo» (2013)
In Silico Fragment-Based Drug Design Using PASS approach

Effects | Mechanisms | Toxicity | Antitargets | Metabolism | Gene Exp.
---|---|---|---|---|---
0.723 | 0.006 | Antithrombotic

0.712 | 0.004 | Platelet aggregation inhibitor
0.618 | 0.019 | Antiangiinal
0.533 | 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

Clopidogrel
In silico Fragment-Based Drug Design using PASS approach (COX-1/2 & LOX inhibitors)

<table>
<thead>
<tr>
<th></th>
<th>COX-1</th>
<th></th>
<th>COX-2</th>
<th></th>
<th>LOX</th>
<th></th>
<th>COX-1</th>
<th></th>
<th>COX-2</th>
<th></th>
<th>LOX</th>
<th></th>
<th>COX-1</th>
<th></th>
<th>COX-2</th>
<th></th>
<th>LOX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>89</td>
<td></td>
<td>100</td>
<td></td>
<td>69</td>
<td></td>
<td>77</td>
<td></td>
<td>96</td>
<td></td>
<td>100</td>
<td></td>
<td>93</td>
<td></td>
<td>95</td>
<td></td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td></td>
<td>94</td>
<td></td>
<td>79</td>
<td></td>
<td>65</td>
<td></td>
<td>80</td>
<td></td>
<td>97</td>
<td></td>
<td>84</td>
<td></td>
<td>87</td>
<td></td>
<td>13</td>
</tr>
</tbody>
</table>

Scheme 1. Structure of the Tested Compounds

Test Results:

8 – Active

7 – Inhibitors of one Enzyme

6 – Inhibitors of both Enzymes
NIH/RFBR project “Virtual screening and biological testing of anti-HIV microbicides”

The purpose: Using virtual screening techniques for multiple targets simultaneously that are of relevance in the early stages of exposure to the virus after sexual intercourse, with the goal to identify commercially available compounds that show cell-protective/microbicide activity in a subsequent cell-based assay.

Collaborators:

Marc Nicklaus, Computer-Aided Drug Design (CADD) Group, Chemical Biology Laboratory, Center for Cancer Research, NCI/NIH.

Stephen Hughes, Ph.D., Director, HIV Drug Resistance Program; Chief, Retroviral Replication Laboratory; and Head, Vector Design and Replication Section, Center for Cancer Research, NCI/NIH.

Virtual Screening Approaches Used in the Project

**Ligand-Based:**
- SAR (PASS)
- QSAR (GUSAR)

**Target-Based:**
- Docking (Glide)

ChemNavigator iResearch Library
Market Select (5,755,574 str.)

50-100 compounds for testing in cell-based assays
From 45 compounds tested in cellular HIV infectivity assay, 16 (~35%) were active with EC$_{50}$-WT = 0.3 – 16.2 µM.

For Tenofovir, which is studied in phase III clinical trials, as microbicide, EC$_{50}$-WT = 10 - 50 µM (in different cellular HIV infectivity assays).
Outline

1. Computational approaches to biological activity prediction.

2. NetFlowEx (targets identification)


5. GUSAR: General Unrestricted Structure-Activity Relationships.

6. Some applications.


8. Summary.
Trust but verify! - I will check how good PASS prediction is for my compounds at: www.way2drug.com/passonline

I trust in PASS prediction!

I do not trust in PASS prediction!

Thanks to the courtesy of Prof. Sergey Vasilevsky
Total users: 8890
Total prediction: 305730
Total countries: 91

Country Predictions
- Russia: 136935
- India: 59564
- China: 14495
- Poland: 13950
- Kazakhstan: 11733
- Ukraine: 10948
- USA: 9552
- Italy: 8718
- ...
Comparison of web-services for biological activity prediction

Web-services:

ChemSpider – http://chemspider.com
CPI-DRAR – http://cpi.bio-x.cn/drar/
PASS Online – http://www.way2drug.com/passonline
SuperPred – http://bioinformatics.charite.de/superpred/
SEA – http://sea.bkslab.org/

Evaluation set (23 diverse compounds, 46 activities):

2011 FDA drug approvals

The US FDA approved 30 new therapeutics last year, including 11 first-in-class agents.

Criteria:

Sensitivity = TP/Na; Computational time

Sensitivity and computational time of different web-services for biological activity prediction
Over fifty publications with independent confirmation of PASS online predictions


---
GUSAR-Based Web-Service: www.way2drug.com/gusar


Site is under construction

We expect an official launch in:

218 Days, 8 Hours, 24 Minutes, and 18 Seconds

But, you can use direct links to our services:

PASS  BBB  DSEP

or during 15 seconds you will be redirected on PASSOnline service

www.way2drug.com
Summary

✓ Computer-aided approaches is useful for finding of hits and their optimization to lead compounds.

✓ PASS predictions allow to identify the most relevant biological screens for testing of particular chemical compounds.

✓ PharmaExpert provides the means for selection of chemical compounds with desirable biological activity spectra (incl. multitargeted actions).

✓ GUSAR can be used as an universal tool for solving various QSAR/QSPR problems.

✓ Predictive web-services are freely available at web-site: http://www.way2drug.com

Contacts: vladimir.poroikov@ibmc.msk.ru (Prof. Vladimir Poroikov)
Acknowledgements

IBMC of RAMS, Russia:
Dmitry Filimonov, Ph.D.
Alexey Lagunin, Ph.D.
Tatyana Gloriozova
Anastasia Rudik, Ph.D.
Olga Tarasova (Filz), Ph.D.
Sergey Ivanov, Ph.D. Student
Pavel Pogodin, Student

NCI/NIH, US:
Marc Nicklaus, Ph.D.
Alexey Zakharov, Ph.D.
Laura Guasch, Ph.D.
Stephen Hughes, Ph.D.
Steven Smith, Ph.D.

Aristotelian University of Thessaloniki, Greece:
Athina Geronikaki, PhD

Karolinska Institute, Sweden:
Galina Selivanova, Ph.D.

GeneXplain GMbH, Germany:
Alexander Kel, Ph.D.

Funding: RFBR (03-07-90282, 05-07-90123, 06-03-08077), CRDF (RC1-2064), INTAS (00-0711, 03-55-5218), ISTC (3197, 3777), FP6 (LSHB-CT-2007-037590), FP7 (200787), RFBR/CRDF grant # 12-04-91445-NIH_A/RUB1-31081-MO-12.
20th EuroQSAR
Understanding Chemical-Biological Interactions

20-th European Symposium on Quantitative Structure-Activity Relationships

Saint-Petersburg, Russia
August 31 – September 4, 2014