

КОМПЬЮТЕРНОЕ ПРЕДСКАЗАНИЕ БИОЛОГИЧЕСКОЙ АКТИВНОСТИ ХИМИЧЕСКИХ ВЕЩЕСТВ: ВИРТУАЛЬНАЯ ХЕМОГЕНОМИКА

В.В. Поройков, Д.А. Филимонов, Т.А. Глорнозова, А.А. Лагунин,
Д.С. Дружиловский, А.В. Степанчикова

ГУ НИИ биомедицинской химии им. В.Н. Ореховича РАМН,
e-mail: vladimir.poroikov@ibmc.msk.ru

На основе прогнозируемых с помощью компьютерной программы PASS спектров биологической активности можно выявлять наиболее вероятные молекулярные мишени, с которыми химические вещества взаимодействуют в биологических объектах, и обусловленные этими взаимодействиями биологические эффекты. Средняя точность прогноза составляет около 94 %, а разработанный метод устойчив по отношению к неполноте информации в обучающей выборке. Широкомасштабное использование данного подхода в качестве «инструмента» для виртуальной хемогеномики (<http://www.ibmc.msk.ru/PASS>) позволило в ряде случаев подтвердить компьютерные предсказания независимыми экспериментами.

Ключевые слова: спектр биологической активности, прогнозирование, PASS, хемогеномика, мишени, лиганды.

Введение

Хемогеномика определяется различными авторами (Caron *et al.*, 2001; Kubinyi, 2006; Rognan, 2007) как новая мультидисциплинарная область науки, конечной целью которой является установление полного соответствия между всеми лигандами и всеми мишенями в биологических объектах. Хемогеномика использует методы установления закономерностей на основе компьютерного анализа информации, содержащейся в химических и биологических базах данных (Rognan, 2007).

Целью создания и развития компьютерной программы PASS (Prediction of Activity Spectra for Substances) является прогнозирование всех видов биологической активности, которые химическое вещество может проявить при взаимодействии с биологическими объектами (Филимонов и др., 1995; Поройков, 1999; Poroikov, Filimonov, 2002, 2005; Филимонов, Поройков, 2006). Прогноз осуществляется на основе анализа обширной обучающей выборки, содержащей информацию о структуре извест-

ных биологически активных веществ и их взаимодействиях с биологическими объектами.

Оригинальность разработанного нами подхода к прогнозированию большого числа видов биологической активности по структурной формуле химического соединения, реализованного в компьютерной программе PASS, была отмечена в ряде публикаций (Muegge, 2003; Jonsdottir *et al.*, 2005; Chen *et al.*, 2006; Balakin *et al.*, 2006; Jenkins *et al.*, 2006; Muster *et al.*, 2008; и др.).

PASS является «инструментом» для виртуальной хемогеномики, позволяющим идентифицировать *in silico* наиболее вероятные лиганды для известных мишеней и, наоборот, наиболее вероятные мишени для известных лигандов, а также предсказывать эффекты, обусловленные взаимодействием конкретных лигандов с конкретными мишенями.

В настоящей работе представлено краткое описание PASS и приведены примеры ее практического применения для установления соответствия между лигандами, их возможными биологическими мишенями и вызываемыми этими взаимодействиями эффектами.

Методика прогнозирования спектров биологической активности

Компьютерное прогнозирование биологической активности осуществляется в программе PASS на основе анализа взаимосвязей «структура–активность» для обширной выборки известных биологически активных веществ (БАВ).

Основные компоненты прогнозирующей системы, помимо обучающей выборки, включают представление биологической активности, описание химической структуры, математический метод анализа зависимостей «структура–активность» и прогноза биологической активности для новых веществ.

Поскольку поиск новых и исследование известных БАВ происходит постоянно, для повышения эффективности компьютерного прогноза необходимы регулярное пополнение и коррекция содержащейся в обучающей выборке информации. Так, в 1995 г. обучающая выборка содержала 9314 БАВ, число прогнозируемых наименований биологической активности составляло 114, а средняя точность прогноза при скользящем контроле с исключением по одному (см. ниже) для всех веществ и всех наименований активности обучающей выборки – 76 %, а в 2007 г. – 117332 БАВ, 3300 наименований активности и 94 % соответственно. Наиболее современная версия программы (бета-версия PASS 2008, июль 2008 г.) содержит 200126 БАВ в обучающей выборке и прогнозирует примерно 4000 наименований биологической активности со средней точностью около 94 %.

Биологическая активность представлена в программе PASS качественно («неактивно»/«активно»), что позволяет объединять в одной обучающей выборке информацию о БАВ, собранную из многочисленных источников. Вместе с тем при наличии соответствующих экспериментальных данных, полученных в одинаковых условиях, можно классифицировать вещества и с учетом величины активности (например «нет

активности»/«слабая активность»/«умеренная активность»/«сильная активность»).

Список прогнозируемых PASS (версия 2008 г.) биологических активностей включает около 4000 наименований¹, из них 140 наименований, характеризующих блокирование эндогенных биорегуляторов (например Acetylcholine antagonist); 71 наименование, характеризующее активирование эндогенных биорегуляторов (например Acetylcholine agonist); 5 наименований, характеризующих блокирование высвобождения эндогенных биорегуляторов (например Acetylcholine release antagonist); 9 наименований, характеризующих стимулирование высвобождения эндогенных биорегуляторов (например Acetylcholine release agonist); 9 наименований, характеризующих блокирование обратного захвата эндогенных биорегуляторов (например Adenosine uptake inhibitor); 2219 наименований, характеризующих ингибирование ферментов (например 12 Lipoxygenase inhibitor); 41 наименование, характеризующее активацию ферментов (например ATPase stimulant); 268 наименований, характеризующих блокирование рецепторов (например 5 Hydroxytryptamine 1 antagonist); 121 наименование, характеризующее активацию рецепторов (например 5 Hydroxytryptamine 1 agonist); 28 наименований, характеризующих блокирование ионных каналов (например Chloride channel antagonist); 5 наименований, характеризующих активацию ионных каналов (например Calcium channel agonist); 28 наименований, характеризующих блокирование транспортеров (например GABA transporter 1 inhibitor); 128 наименований, характеризующих вещества в качестве субстратов ферментов метаболизма² (например CYP3A4 substrate); 24 наименования, характеризующих ингибирование ферментов метаболизма (например CYP3A4 inhibitor); 24 наименования, характеризующих индукцию ферментов метаболизма (например CYP3A4 inducer); 28 наименований, характеризующих ингибирование белков, не являющихся

¹ В программе PASS используются англоязычные наименования видов биологической активности, широко распространенные в базах данных по фармакологически активным соединениям (см., например <http://integrity.prou.com/>).

² В фармакологии различают ферменты, являющиеся мишенями фармакологических веществ, воздействие на которые приводит к определенным фармакологическим эффектам, и ферменты, осуществляющие метаболизм лекарств (фармакодинамика и фармакокинетика взаимодействия вещества с организмом). Поскольку программа PASS изначально создавалась для компьютерной оценки новых фармакологических веществ, в ней используется соответствующая фармакологическая терминология, различающая ферменты-мишени и ферменты метаболизма лекарств.

ферментами (например Collagen inhibitor); 8 наименований, характеризующих ингибирование экспрессии транскрипционных факторов (например Transcription factor Rho Inhibitor); 2 наименования, характеризующих усиление экспрессии транскрипционных факторов (например Transcription factor TP53 expression enhancer); 72 наименования, характеризующих блокирование биологических процессов (например Apoptosis antagonist); 40 наименований, характеризующих стимуляцию биологических процессов (например Apoptosis agonist); 66 наименований, характеризующих воздействие веществ на инфекционные агенты (например Antileishmanial); 261 наименование, характеризующее фармакотерапевтическое действие (например Anxiolytic); 389 наименований, характеризующих побочное/токсическое действие веществ (например Carcinogen).

Для описания структуры органических соединений в PASS выбрана в качестве основы структурная формула. Это соответствует задаче прогноза спектров активности новых, планируемых к синтезу (сконструированных *in silico*), органических соединений. Стереохимические особенности молекул не учитываются, несмотря на очевидную зависимость от них биологической активности. Это обусловлено невозможностью обеспечить полноту такой информации для достаточно большой выборки структур, а также тем, что единой универсальной стереохимической номенклатуры нет, и наиболее адекватным способом представления стереохимической информации является на самом деле пространственная структура молекул. На основе структурной формулы вещества PASS генерирует так называемые MNA-дескрипторы (Multilevel Neighborhoods of Atoms – многоуровневые атомные окрестности), которые строятся рекурсивно для каждого атома молекулы следующим образом:

– MNA-дескриптор 0-го уровня – метка A самого атома;

– MNA-дескриптор любого следующего уровня – условное обозначение структурного фрагмента $A(D_1 D_2 \dots D_j \dots)$, где D_i – MNA-дескриптор предыдущего уровня для i -го непосредственного соседа данного атома с меткой A . Дескрипторы соседей $D_1 D_2 \dots D_j \dots$ записываются в каком-либо однозначном порядке, например лексикографическом.

В настоящее время в PASS используется бесповторное множество MNA-дескрипторов 1-го и 2-го уровней; при этом в дескрипторах 2-го уровня используется индикатор « \rightarrow » для обозначения атомов в цепях.

Детальное описание MNA-дескрипторов приведено в публикации Filimonov *et al.* (1999), а методология их разработки описана в статье Д.А. Филимонова и В.В. Поройкова (2006).

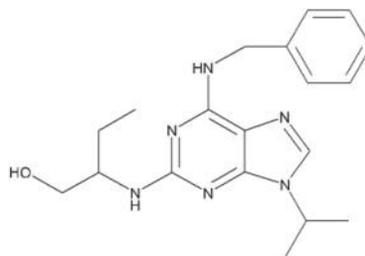
Используемый в PASS математический подход детально описан в публикациях (Филимонов, Поройков, 2006; Filimonov, Poroiikov, 2008).

Результат прогноза спектра биологической активности представлен в PASS в виде упорядоченного списка названий соответствующих активностей и вероятностей P_a «быть активным» и P_i «быть неактивным» для прогнозируемого соединения. Упорядочение выполнено по убыванию разности $P_a - P_i$, соответственно более вероятные виды активности находятся в начале спрогнозированного спектра. Прогноз спектра активности может анализироваться в соответствии с потребностями пользователя, но по умолчанию в него включаются активности, для которых $P_a > P_i$.

На рис. 1 в качестве примера приведены структурная формула и часть прогнозируемого PASS (версия 2007) спектра биологической активности для препарата Seliciclib (R-Roscovitine, Cyclacel Ltd.).

Как видно из рис. 1, известное для данного препарата ингибирование CDK1 прогнозируется с вероятностью, превышающей 50 % (выделено полужирным шрифтом). В качестве дополнительных фармакотерапевтических эффектов прогнозируется Systemic lupus erythematosus treatment, Multiple sclerosis treatment и Myocardial ischemia treatment, а в качестве действия на дополнительную мишень – Hydrolase inhibitor.

В процессе обучения программы реализуется процедура скользящего контроля с исключением по одному. При этом каждое вещество обучающей выборки «исключается» из нее со всей сопутствующей информацией о биологической активности, и прогноз для него осуществляется так же, как для нового соединения. Процедура повторяется для всех веществ обучающей выборки, в результате чего мы получаем оценки точности прогноза,



```

> <PASS_RESULT_COUNT>
6 of 3300 Possible Activities at Pa > 0.500

> <PASS_ACTIVITY_SPECTRUM>
0.785 0.002 Cyclin-dependent kinase inhibitor
0.677 0.051 Systemic lupus erythematosus treatment
0.625 0.001 Cyclin-dependent kinase 1 inhibitor
0.555 0.065 Multiple sclerosis treatment
0.600 0.122 Hydrolase inhibitor
0.591 0.148 Myocardial ischemia treatment

```

Рис. 1. Пример прогноза спектра биологической активности для препарата Seliciclib.

усредненные по всей обучающей выборке и по подвыборкам веществ, обладающих конкретными видами биологической активности. Как уже говорилось выше, средняя точность прогноза составляет около 94 %. Минимальная точность прогноза составляет около 70 %, например для Antineoplastic (Renal cancer); максимальная – 100 % (для ряда активностей, включая 3С-Like protease (Human coronavirus) inhibitor, Angiotensin AT1 A receptor antagonist, Antidote (organophosphates) и др.

В специальном компьютерном эксперименте было показано, что алгоритм прогноза PASS является устойчивым по отношению к неполноте информации о структурах и спектрах биологической активности химических соединений в обучающей выборке. Сокращение реально известной информации о структурах или активности химических соединений в обучающей выборке на 50 % лишь незначительно уменьшало точность прогноза в перекрестном контроле (Poroikov *et al.*, 2000).

Величины P_a и P_i интерпретируются как оценки меры принадлежности вещества к классам активных и неактивных соединений либо как оценки ошибок первого и второго рода. Чем больше для конкретной активности величина P_a и чем меньше величина P_i , тем больше шанс обнаружить данную активность в эксперименте.

В дальнейшем мы будем рассматривать ситуацию, когда значение P_a достаточно высоко и значительно превосходит P_i . Если при анализе прогноза спектра активностей для исследования выбираются те виды активности, для которых $P_a > 90\%$, то мы рискуем пропустить около 90 % действительно активных соединений, но вероятность ложноположительных прогнозов при этом ничтожно мала; для $P_a > 80\%$ пропустим уже только 80 % активных соединений, но и вероятность ложноположительных прогнозов будет выше, наконец, для $P_a > P_i$ вероятности ошибок первого и второго рода равны. Этот порог используется «по умолчанию» для отсекаания вероятных видов активности при прогнозе.

Если прогноз спектра активности обширен, то структура молекулы данного вещества довольно проста и не содержит каких-либо особенностей, обеспечивающих высокую селективность его биологического действия.

Возможности и примеры использования программы PASS

Как уже говорилось выше, PASS позволяет предсказывать *in silico*: 1) вероятные мишени для известных лигандов; 2) вероятные лиганды для известных мишеней; 3) вероятные фарма-

котерапевтические эффекты, а также побочное и токсическое действие веществ.

С целью расширения возможностей использования PASS научным сообществом мы реализовали веб-сервис PASS INet (<http://www.ibmc.msk.ru/PASS>), позволяющий осуществить прогноз спектра биологической активности веществ через Интернет (Lagunin, *et al.*, 2000; Садым и др., 2002; Sadym *et al.*, 2003). Используя стандартный Интернет-браузер, пользователь может направить на сайт <http://www.ibmc.msk.ru/PASS/> структурную формулу вещества, представленную в виде MOL-файла, и автоматически получить на дисплее своего компьютера прогноз наиболее вероятных для данной молекулы наименований биологической активности. В 2007 г. число зарегистрированных пользователей составило около 3000, а число структур молекул, направленных на прогноз спектра биологической активности, более 70000.

На основе прогнозируемых PASS INet спектров биологической активности синтезировано и протестировано около 20 новых молекул, принадлежащих к различным химическим классам; результаты прогноза подтверждены в эксперименте. В частности, показано наличие ГАМК-эргической и анксиолитической активности у производных хиназолинона, ноотропной активности у циклических нитронов, ЦНС-модулирующей активности у производных азетидин-2-она, противовоспалительной активности у производных 1-ациламиноалкил-3,4-диалкоксибензола, антибактериальной активности у замещенных амидов и гидразидов дикарбоновых кислот, антилейшманиозной активности у 9-амино-2-метоксиакридинов и др. (Geronikaki *et al.*, 2008).

Обширный список прогнозируемых PASS наименований биологической активности позволяет вести поиск как монотаргетных лигандов, действующих на отдельные мишени, так и мультитаргетных веществ, действующих на две или более мишеней. Виртуальный скрининг может проводиться в базах данных (БД) коммерчески доступных образцов химических веществ, поставляемых рядом фирм (см., например, <http://www.chembridge.com>, <http://www.ibscreen.com>, <http://www.asinex.com>).

Рассмотрим в качестве примера поиск моно- и мультитаргетных веществ среди веществ из

БД ChemBridge DIVERSet (версия 2006 г.). С помощью PASS (версия 2007 г.) был выполнен прогноз спектров биологической активности для 39856 из 40080 молекул, содержащихся в этой БД (224 структуры являются двухкомпонентными, и для них прогноз не был получен). Время прогноза на стандартном персональном компьютере – около 15 минут, т. е. около 0,02 секунды/молекулу.

Результаты прогноза были введены в компьютерную программу PharmaExpert (Poroikov *et al.*, 2005), которая дает возможность поиска веществ с заданным профилем биологической активности. Мы выполнили поиск молекул, прогнозируемых как ингибиторы циклин-зависимой киназы 1, с помощью программы PharmaExpert. Оказалось, что на уровне $P_a > P_i$ такая активность прогнозируется для 165 веществ; при $P_a > 10\%$ – для 164 веществ; при $P_a > 20\%$ – для 8 веществ; при $P_a > 30\%$ – для 3 веществ; при $P_a > 40\%$ – не прогнозируется ни для одной молекулы. Таким образом, вероятность выявления ингибиторов циклин-зависимой киназы 1 среди веществ из БД ChemBridge DIVERSet не велика. Аналогичный поиск может быть проведен в других БД, например в ChemNavigator (<http://www.chemnavigator.com>), содержащей информацию о более чем 25 млн уникальных структур.

Далее мы проанализировали возможность выявления в этой БД мультитаргетных противоопухолевых веществ. Оказалось, что число веществ, прогнозируемых с $P_a > 70\%$ как действующих одновременно на 2 мишени, составляет 395 молекул; на 3 – 334 молекулы; на 4 – 248 молекул; на 5 – 93 молекулы; на 6 – 18 молекул; на 7 – 1 молекулу. Таким образом, на основе компьютерного прогноза спектров биологической активности можно отбирать потенциальные мультитаргетные противоопухолевые препараты. Ранее такого рода поиск был нами проведен для антигипертензивных (Lagunin *et al.*, 2003) и противовоспалительных веществ (Geronikaki *et al.*, 2008) с дуальными механизмами действия, и результаты прогноза были подтверждены экспериментально (Lagunin *et al.*, 2003; Geronikaki *et al.*, 2008).

На основе прогноза спектров биологической активности с помощью PASS были выявлены вещества с противоязвенным, гепатопротекторным, противоопухолевым, местноанестезирующим,

противовоспалительным, антиаритмическим, антиоксидантным, анксиолитическим, ноотропным, антибактериальным, антилейшманиозным и рядом других эффектов (Филимонов, Поройков, 2006; Geronikaki *et al.*, 2008). Нами также были показаны возможности применения PASS для прогноза удлинения QT-интервала (Filz *et al.*, 2007) и канцерогенности (Lagunin *et al.*, 2005).

Приведенные примеры убедительно демонстрируют, что с использованием PASS можно успешно предсказывать как основные фармако-терапевтические, так и побочные эффекты уже на стадии конструирования молекул *in silico* до проведения реального химического синтеза и биологического тестирования.

Заключение

Применение PASS позволяет решать некоторые задачи хемогеномики *in silico*, однако обязательным условием этого является предварительное наличие информации о некотором числе лигандов, которые могут составить обучающую выборку. В случае новых мишеней, для которых такого рода лиганды еще не известны, сначала должен быть проведен их поиск экспериментальными методами.

Представляется весьма перспективным применение PASS для поиска мультитагетных лигандов путем виртуального скрининга баз данных коммерчески доступных образцов химических соединений, поскольку такие препараты могут обладать значительными преимуществами в сравнении с монотаргетными лекарствами (Wermuth, 2004).

Совместное использование программ PASS и PharmaExpert дает возможность анализа межлекарственных взаимодействий, что весьма актуально как при разработке комбинированных синтетических препаратов, так и при оценке действия природных субстанций, полученных на основе лекарственных растений.

Благодарности

Работа выполнена при частичной финансовой поддержке грантов РФФИ (05-07-90123, 06-03-08077, 06-03-39015), CRDF (RC1-2064), ИНТАС (03-55-5218), МНТЦ (574, 3197, 3777), FP6 (LSHB-CT-2007-037590), FP7 (200787).

Литература

- Поройков В.В. Компьютерное предсказание биологической активности веществ: пределы возможного // Химия в России. 1999. № 2. С. 8–12.
- Филимонов Д.А., Поройков В.В. Прогноз спектров биологической активности органических соединений // Рос. химич. журнал. 2006. Т. 50. № 2. С. 66–75.
- Филимонов Д.А., Поройков В.В., Караичева Е.И. и др. Компьютерное прогнозирование спектра биологической активности химических соединений по их структурной формуле: система PASS // Эксперим. клинич. фармакология. 1995. Т. 58. № 2. С. 56–62.
- Садым А.В., Лагунин А.А., Филимонов Д.А. и др. Интернет-система прогноза спектра биологической активности химических соединений // Хим.-фарм. журнал. 2002. Т. 36. № 10. С. 21–26.
- Balakin K.V., Tkachenko S.E., Kiselyov A.S. *et al.* Focused chemistry from annotated libraries // Drug Discov. Today: Technologies. 2006. V. 3. № 4. P. 397–403.
- Caron P.R., Mullican M.D., Mashal R.D. *et al.* Chemogenomic approach to drug discovery // Cur. Opin. Chem. Biol. 2001. V. 5. № 4. P. 464–470.
- Chen X., Liang Y., Xu J. Toward automated biochemotype annotation for large compound libraries // Mol. Diversity. 2006. V. 10. № 3. P. 495–509.
- Geronikaki A., Druzhilovsky D., Zakharov A. *et al.* Computer-aided predictions for medicinal chemistry via Internet // SAR and QSAR in Environ. Res. 2008. V. 19. № 1/2. P. 27–38.
- Geronikaki A.A., Lagunin A.A., Hadjipavlou-Litina D.I. *et al.* Computer-aided discovery of anti-inflammatory thiazolidinones with dual cyclooxygenase/lipoxygenase inhibition // J. Med. Chem. 2008. V. 51. № 6. P. 1601–1609.
- Filimonov D.A., Poroikov V.V. Probabilistic approach in activity prediction // Chemoinformatics Approaches to Virtual Screening / Ed. A. Varnek, A. Tropsha. Cambridge: RSC Publishing, 2008.
- Filimonov D., Poroikov V., Borodina Yu. *et al.* Chemical similarity assessment through multilevel neighbourhoods of atoms: definition and comparison with the other descriptors // J. Chem. Inf. Comput. Sci. 1999. V. 39. № 4. P. 666–670.
- Filz O., Lagunin A., Filimonov D. *et al.* Computer-aided prediction of QT-prolongation // SAR and QSAR Environ. Res. 2007. V. 19. № 1/2. P. 81–90.
- Jenkins J.L., Bender A., Davies J.W. *In silico* target fishing: predicting biological targets from chemical structure // Drug Discov. Today: Technologies. 2006. V. 3. № 4. P. 413–421.

- Jonsdottir S.O., Jorgensen F.S., Brunak S. Prediction methods and databases within chemoinformatics: emphasis on drugs and drug candidates // *Bioinformatics*. 2005. V. 21. № 10. P. 2145–2160.
- Kubinyi H. Chemogenomics in drug discovery // *Ernst Schering Res. Found. Workshop*. 2006. V. 58. P. 1–19.
- Lagunin A.A., Dearden J., Filimonov D.A. *et al.* Computer-aided rodent carcinogenicity prediction // *Mutat. Res.* 2005. V. 586. № 2. P. 138–146.
- Lagunin A.A., Gomazkov O.A., Filimonov D.A. *et al.* Computer-aided selection of potential antihypertensive compounds with dual mechanisms of action // *J. Med. Chem.* 2003. V. 46. № 15. P. 3326–3332.
- Lagunin A., Stepanchikova A., Filimonov D. *et al.* PASS: prediction of activity spectra for biologically active substances // *Bioinformatics*. 2000. V. 16. № 8. P. 747–748.
- Muegge I. Selection criteria for drug-like compounds // *Med. Res. Rev.* 2003. V. 23. № 3. P. 302–321.
- Muster W., Breidenbach A., Fisher H. *et al.* Computational toxicology in drug development // *Drug Discov. Today*. 2008. V. 13. № 7. P. 303–310.
- Poroikov V.V., Filimonov D.A. How to acquire new biological activities in old compounds by computer prediction // *J. Comput. Aid. Mol. Des.* 2002. V. 16. № 11. P. 819–824.
- Poroikov V., Filimonov D. PASS: Prediction of Biological Activity Spectra for Substances // *Predictive Toxicology / Ed. Christoph Helma*. London: Taylor and Francis, 2005. P. 459–478.
- Poroikov V.V., Filimonov D.A., Borodina Yu.V. *et al.* Robustness of biological activity spectra predicting by computer program PASS for non-congeneric sets of chemical compounds // *J. Chem. Inform. Comput. Sci.* 2000. V. 40. № 6. P. 1349–1355.
- Poroikov V., Lagunin A., Filimonov D. PharmaExpert: diseases, targets and ligands – three in one // *QSAR and Molecular Modelling in Rational Design of Bioactive Molecules / Ed. Esin Aki Sener, Ismail Yalcin*. Ankara (Turkey): CADD & D Society, 2005. P. 514–515.
- Rognan D. Chemogenomic approach to rational drug design // *British J. Pharmacol.* 2007. V. 152. № 1. P. 38–52.
- Sadym A., Lagunin A., Filimonov D. *et al.* Prediction of biological activity spectra via Internet // *SAR and QSAR Environ. Res.* 2003. V. 14. № 5/6. P. 339–347.
- Wermuth C.G. Multitargeted drugs: the end of the «one-target-one-disease» philosophy? // *Drug Discov. Today*. 2004. V. 9. № 19. P. 826–827.

COMPUTER-AIDED PREDICTION OF BIOLOGICAL ACTIVITY SPECTRA FOR SUBSTANCES: VIRTUAL CHEMOGENOMICS

V.V. Poroikov, D.A. Filimonov, T.A. Glorizova, A.A. Lagunin,
D.S. Druzhilovsky, A.V. Stepanchikova

The Orekhovich Institute of Biomedical Chemistry of Russian Academy of Medical Sciences,
Moscow, Russia, e-mail: vladimir.poroikov@ibmc.msk.ru

Summary

On the basis of predicted with computer program PASS biological activity spectra it is possible to reveal the most probable molecular targets, with which chemical compounds interact in biological objects, and biological effects caused by these interactions. Average accuracy of prediction is about 94 %, and the method is robust in relation to the incompleteness of information in the training set. Wide utilization of this approach as a «tool» for virtual chemogenomics (<http://www.ibmc.msk.ru/PASS>) provides the opportunities for independent experimental confirmation of predictions in a dozen of cases.