

ПРЕДИКТИВНЫЕ ТЕХНОЛОГИИ В ИССЛЕДОВАНИИ НОВЫХ ЛЕКАРСТВЕННЫХ ВЕЩЕСТВ

Пшенкина Н.Н.

ФГБОУ ВПО "Военно-медицинская академия им. С.М.Кирова" МО РФ

194044, Санкт-Петербург, ул. Академика Лебедева, 6

Тел. (812) 542-35-07 E-mail: pshenkina@mail.ru

Резюме. Использование достижений хемо- и биоинформатики в фармацевтической химии и в фармакологии позволило совершить технологический прорыв в области разработки, направленного синтеза и создания новых лекарственных веществ. Прогнозирование фармакокинетического профиля соединений-кандидатов в совокупности с технологиями высокопроизводительного скрининга позволило повысить эффективность ранних этапов разработки новых лекарственных веществ. Основу предиктивной методологии составляет количественный анализ связи «структура – свойства – активность» в различных его модификациях. В обзоре кратко рассмотрены основные принципы предиктивной методологии и примеры ее использования при разработке новых лекарственных веществ.

Ключевые слова: прогнозирование, исследование лекарств, фармакокинетика, виртуальный скрининг, связь структура-активность.

PREDICTIVE TECHNOLOGIES IN DRUG DISCOVERY

Pshenkina N.N.

Military Medical Academy named after S.M. Kirov,

Academic Lebedev str., 6, Saint Petersburg, 194044, Russia

Tel. +7 (812) 542-35-07 E-mail: pshenkina@mail.ru

Abstract. Modern drug discovery and development is the interdisciplinary process. Today, the process of drug discovery has been revolutionized due to the advances in bioinformatics. Efficient technologies such as combinatorial chemistry, high throughput screening (HTS), virtual screening, de novo design and structure-based drug design contribute greatly to drug discovery. This paper presents a review of current advances and novel approaches in experimental and computational drug discovery and design.

Key words: prediction, drug discovery, pharmacokinetics, virtual screening, QSAR

Существование зависимости между строением биологически активных веществ и их свойствами с давних пор привлекает внимание ученых. Поиск закономерностей этой взаимосвязи проводился многими исследователями в рамках методологии и концептуальных представлений, соответствующих развитию научных знаний современного периода. Первые наблюдения такого рода можно отнести к концу XIX века, однако широкомасштабные исследования связи «структура – активность» начались в середине прошлого века. В числе исследователей, заложивших фундамент этого научного направления, такие исследователи как I. Langmuir, H. Wiener, L.P. Hammett, R.W. Taft и др. [55]. Однако, признанными родоначальниками методологии оценки количественной связи «структура – активность» являются С. Hansch и Т. Fujita с их работой, опубликованной в 1964 г., в которой они предложили модель, связывающую биологическую активность с гидрофобными, электрическими и пространственными свойствами молекул [27].

В отечественной литературе для сокращенного обозначения методологии анализа количественной связи «структура – активность» иногда используется аббревиатура КССА, однако в мировой литературе принято обозначение QSAR (Quantitative Structure-Activity Relationships).

На начальных этапах становления методологии QSAR она была, в основном, востребована в аналитической химии при оценке связи между структурой и физико-химическими свойствами веществ. Привлечение к анализу молекулярных дескрипторов дало возможность более точно описать молекулу и послужило стимулом для интенсивного развития метода [40]. Современные технологии QSAR позволяют устанавливать корреляции между молекулярной структурой и различными физико-химическими свойствами соединений, что делает этот подход востребованным в фармакологических исследованиях, в частности, при разработке и создании новых лекарственных веществ.

В настоящее время для расчетов, проводимых с помощью методов компьютерной химии, используются сотни дескрипторов. Наиболее полно они описаны в энциклопедической монографии R. Todeschini, V. Consonni (2000) [58]. В целом, выделяют следующие категории молекулярных дескрипторов:

- Фрагментные дескрипторы – показывают, содержится ли данный фрагмент (подструктура) в структурной формуле анализируемого соединения [9,70].

- Квантово-химические дескрипторы – числовые величины, получаемые в результате квантово-химических расчетов (энергия молекулярных орбиталей, частичные заряды, индексы реакционной способности и др.) [39].
- Дескрипторы молекулярных полей – величины, аппроксимирующие значения молекулярных полей путем вычисления энергии взаимодействия пробного атома, помещенного в узел решетки, с текущей молекулой. Находят широкое применение в 3D-QSAR моделях [15].
- Константы заместителей – связывают константы скорости реакции с константами равновесия для некоторых классов органических реакций [4]. Данные дескрипторы получили широкое применение, благодаря работам С. Hansch и Т. Fujita [20, 27, 37].
- Топологические индексы – числовое значение (или набор значений), описывающее структуру молекулы на базе графов. В теории графов обычно водородные атомы в расчет не принимаются, а рассматривается лишь каркас молекулы, построенный из атомов С, N, O, S и др. Топологические индексы зависят от структуры молекулы, и не отражают ее состав [3].

Среди физико-химических дескрипторов, получаемых с помощью методов компьютерной химии, в фармакологии и лекарственной химии чаще всего используются липофильность ($\log P$), молярная рефракция (MR), молекулярный вес (MW), дескрипторы водородной связи, молекулярные объемы и площади поверхностей молекул [5, 6]. Помимо физико-химических параметров при прогнозировании фармакокинетического профиля находят применение, так называемые фармакофоры. Они представляют собой набор пространственных и электронных признаков, необходимых для проявления искомого типа биологической активности. Типичными фармакофорами являются гидрофобные области, ароматические кольца, доноры и акцепторы водородной связи, анионные и катионные центры [11].

В фармакологии и лекарственной химии в последнее время активно развиваются такие разновидности QSAR метода как SAR (Structure-Activity Relationships), QSPR (Quantitative Structure-Property Relationships) и QPAR (Quantitative Property-Activity Relationships), которые исследуют, соответственно, качественную связь «структура – активность», количественную связь «структура – свойства» и «свойства – активность». Использование в исследованиях QPAR физико-химических свойств, предсказанных с помощью методов QSPR, позволяет создавать интегральные QSAR-модели, имеющие широкие предиктивные возможности.

Основная стратегия QSAR / QSPR методологии состоит в том, чтобы на примере анализа соединений с известными взаимосвязями «структура – активность (свойства)» найти такие оптимальные количественные отношения, которые затем могли бы быть использованы для предсказания свойств неисследованных и даже несуществующих молекул. Особенность QSAR-анализа заключается в том, что заранее неизвестно, сколько и какие именно дескрипторы необходимы для описания заданного свойства. Зачастую возникает соблазн использовать как можно больший набор параметров. Однако такой подход не всегда оправдан, поскольку велика вероятность выявления случайных корреляций, на основе которых можно придти к ложным заключениям.

Для построения моделей, описывающих связи типа «структура – свойства – активность», привлекаются широкие возможности современных информационных технологий. Методы компьютерного анализа можно подразделить на две основных категории: множественные линейные и нелинейные методы. Множественный линейный регрессионный анализ (Multiple Linear Regression, MLR) и метод частичных наименьших квадратов (Partial Least Squares, PLS) принадлежат к наиболее популярным линейным многомерным методам создания моделей. Находят применение метод k-ближайших соседей, кластерный, логистический, дискриминантный анализ и др. Среди нелинейных многомерных методов чаще других используются искусственные нейронные сети (Artificial Neural Networks, ANN).

Применение вычислительных методов позволяет произвести расчеты физико-химических и других свойств практически для неограниченного числа соединений, для которых известна лишь структурная формула. Это могут быть соединения, уже имеющие практическое применение или находящиеся в стадии экспериментального синтеза и разработки. Полученные результаты составляют содержание различных баз данных и библиотек, многие из которых общедоступны в сети Internet [21, 65, 68] и могут использоваться фармакологами и химиками при оптимизации существующих и разработке новых лекарственных средств. Одной из наиболее обстоятельных является фармакологическая база данных DrugBank, созданная в 2006 г. специалистами Университета Альберта (University of Alberta, Edmonton, Canada) [17]. DrugBank представляет собой электронный ресурс, содержащий всеобъемлющую информацию о более чем 1500 соединений, имеющих медицинское применение или находящихся в стадии экспериментальных исследований. Индивидуальная карта, составленная для каждого из соединений, имеет около 100 полей данных, содержащих сведения о физико-химических параметрах, фармакологическом действии, фармакокинетике, лекарственных

взаимодействиях, о генетических и молекулярно-биологических свойствах мишеней и другую информацию.

Помимо баз данных в настоящее время существует множество компьютерных программ, позволяющих исследователям самостоятельно выполнять вычисления физико-химических и других параметров в режиме on-line, имея лишь сведения о структуре интересующего соединения. Многие из таких приложений свободно доступны в сети Internet [29, 30, 62].

Накоплению данных о свойствах и активности лекарственных веществ в значительной мере способствовало внедрение исследовательских технологий высокопроизводительного скрининга (High Throughput Screening, HTS) [24, 45]. HTS-технологии основаны на использовании нетрудоемких и в то же время информативных аналитических методов, адаптированных для одновременного выполнения большого количества измерений. Методы скрининга разработаны для малых объемов образцов, что позволяет проводить анализ с использованием специальных планшетов, содержащих от 96 до 1536 лунок, что позволяет исследовать до 100 000 образцов в день [48]. Вместе с тем, на современном этапе стремление к увеличению пропускной способности скрининга явно снижается. По прогнозам специалистов, в дальнейших HTS-исследованиях больше внимания будет уделяться повышению качества и точности характеристики соединений-кандидатов и мишеней их действия [49].

В современных HTS-технологиях используются методы анализа флуоресценции, аффинной хроматографии, поверхностного плазмонного резонанса, анализа микроокружения ДНК и др. Полагают, что методы, основанные на анализе флуоресценции, будут одними из наиболее востребованных в HTS-технологиях, поскольку сочетают в себе высокую информативность, пригодность для решения широкого круга задач, высокую чувствительность и восприимчивость к автоматизации анализа [28, 38]. Другой современной технологией, находящей широкое применение в высокопроизводительном скрининге лекарств, являются методы спектроскопии, основанные на исследовании ядерно-магнитного резонанса (ЯМР). ЯМР-спектроскопия позволяет непосредственно получить информацию о локализации связывающих центров в молекуле белка и о средстве лигандов к этим центрам. Полученные результаты позволяют актуализировать информацию молекулярных библиотек и могут быть использованы в дальнейших HTS-исследованиях. Аффинная хроматография позволяет получать ценную информацию о важнейших фармакологических процессах как абсорбция, распределение, экскреция, активация рецепторов. Возможности метода позволяют измерять параметры

комплексообразования в режиме реального времени, а также получать другие характеристики взаимодействия малых молекул с макромолекулярными мишенями. Широкие возможности открываются при интеграции HTS-технологий с данными геномики и протеомики, позволяющие проводить более целенаправленный синтез и эффективный скрининг препаратов, специфических к целевым макромолекулярным структурам [14, 19, 54].

Применение технологий высокопроизводительного скрининга позволяет в короткие сроки и в одинаковых экспериментальных условиях получать значительный объем информации об исследуемых соединениях, что позволяет в дальнейшем проводить адекватное сравнение и анализ полученных данных. В настоящее время, благодаря интенсивному использованию методов хемоинформатики и HTS технологий, накоплен огромный массив данных, характеризующих различные свойства лекарственных соединений [45, 73], которые требуют своего анализа и осмысления. Не менее значительными представляются дальнейшие перспективы применения методов хемоинформатики и анализа связи «структура-активность» в области совершенствования существующих лекарственных веществ, расширения областей их применения, разработке лекарств *de novo*. Детальный анализ этих направлений выходит за рамки настоящего исследования. Его можно найти в последних обзорах R. Guha и W.A. Warr (2011) [25, 63].

Прогнозирование параметров фармакокинетики

В контексте данной работы представляет интерес связь физико-химических и структурных свойств молекул с их основными фармакокинетическими параметрами, такими как абсорбция, распределение, метаболизм и экскреция. В зарубежной литературе этот комплекс параметров фармакокинетики обозначается аббревиатурой ADME (Absorption – Distribution – Metabolism – Excretion). В ряде случаев в число показателей включается оценка токсичности, тогда совокупность характеристик фармакокинетики обозначается как ADME/T.

Ряд физико-химических свойств молекул лекарственных веществ, такие как водорастворимость, коэффициенты распределения ($\log P$ и $\log D$), константа ионизации (pK_a), площадь топологической полярной поверхности ($TPSA$) и т.д. играют важную роль в фармакокинетических процессах. Сегодня прогнозирование фармакокинетического профиля (ADME/T профилирование) новых химических соединений на ранних этапах их разработки стало обычной практикой для крупных фармацевтических компаний, таких как Hoffmann-La Roche, Aventis, Novartis, Pfizer и многих других [18, 49, 59]. Вещества с

нежелательными физико-химическими свойствами исключаются на ранних этапах скрининга, что значительно сокращает объем финансовых вложений, трудозатраты, минимизирует трудоемкие эксперименты на животных, способствует гуманизации процесса разработки в целом, а главное – значительно повышает его эффективность.

Абсорбция и биодоступность

Известно, что большинство лекарств, поступающих перорально, преодолевает кишечный эпителий посредством пассивного транспорта, который, в свою очередь, в значительной степени зависит от физико-химических свойств препарата.

Одна из плодотворных попыток установить зависимость между структурными и физико-химическими параметрами лекарственных веществ и их активностью при приеме внутрь была предпринята С.А. Lipinski [47]. Проанализировав большое количество лекарственных веществ и кандидатов в лекарства, С.А. Lipinski сформулировал эмпирическое правило, которое получило известность как «Правило 5-ти» (Lipinski's Rule-of-Five), поскольку установленные численные границы параметров кратны пяти.

Правило Lipinski гласит, что лекарственные вещества, имеющие хорошую биодоступность при приеме внутрь, должны обладать следующими структурными и физико-химическими характеристиками:

- число доноров водорода в молекуле (число атомов азота или кислорода в соединении с одним или несколькими атомами водорода) должно быть не более 5;
- число акцепторов водорода (число атомов азота или кислорода) должно быть не более 10;
- молекулярная масса не должна превышать 500 D;
- логарифм распределения вещества в системе октанол/вода должен быть не более 5.

Число отклонений от этого правила получило название числа Lipinski, при значении которого, равном 2 и более, прогнозируется плохая биодоступность препарата при пероральном приеме.

Помимо правила Lipinski были предложены и другие правила, описывающие зависимость параметров фармакокинетики от физико-химических и структурных свойств молекул. Так, D.F. Veber et al. (2002) в результате исследования биодоступности 1100 соединений, рассматриваемых в качестве кандидатов на роль лекарств, пришли к выводу, что молекулярная масса не является параметром, строго определяющим биодоступность. По мнению авторов, гораздо более значимыми являются гибкость молекулы, измеряемая числом вращающихся связей, и площадь полярной поверхности молекулы, зависящая от

числа доноров и акцепторов водорода [60]. Согласно D.F. Veber et al., вещества будут иметь высокую биодоступность при приеме внутрь, если их структура удовлетворяет следующим требованиям:

- число вращающихся связей в молекуле – не более 10;
- площадь топологической полярной поверхности – не более $140 \text{ \AA}^2$;
- сумма доноров и акцепторов водорода не превышает 12.

Анализ обширной базы данных, в которую входили лекарственные вещества различных фармакологических групп, проведенный A.K. Ghose et al., показал, что свойства лекарственных веществ, в целом, укладываются в пределы, обозначенные правилами Lipinski, Veber и др. Так, более 80% лекарств, зарегистрированных в США, имели значения $\log P$ от $-0,4$ до $5,6$ (в среднем $2,52$), молекулярную массу MW – от 160 до 480 (в среднем 357), коэффициент молярной рефракции MR – от 40 до 130 (в среднем 97), общее количество атомов в молекуле – от 20 до 70 (в среднем 48) [23].

В дальнейшем перечень структурных и физико-химических параметров веществ-кандидатов в лекарства подвергался многократным уточнениям и дополнениям [41, 46, 52, 53, 71]. Высказывались также нарекания в излишнем упрощении представлений, на которых основывается правило Lipinski, в приближенности прогнозов, в необходимости рассматривать не только пероральный, но и парентеральный путь введения лекарств и т.д.. Однако до сегодняшнего дня правило Lipinski эффективно используется на ранних этапах разработки лекарств при отборе кандидатов из числа биологически активных соединений, и считается, своего рода, золотым стандартом. Кроме того, внедрение «Правил 5-ти» в фармакологический скрининг индуцировало принципиальный сдвиг в парадигме мышления медицинских химиков. Обобщение, сформулированное С.А. Lipinski, позволило сделать важный шаг в направлении понимания различий между биологически активными молекулами и лекарствами, а также подойти к разработке лекарств с учетом фармакокинетики, начиная с ранних этапов их создания [41].

Работы, связанные с развитием *in silico* моделей гастроинтестинальной абсорбции, рассмотрены в ряде обзоров [35, 51, 64]. В недавнем исследовании A. Yan et al. были разработаны QSAR-модели на основе набора данных для 552 соединений, для которых известны экспериментальные параметры абсорбции [67]. Установлено, что для прогнозирования важны такие дескрипторы как $\log P$, $\log S$, MW (молекулярная масса) и $TPSA$. В другом подобном исследовании для классификации соединений с высокой или низкой биодоступностью ($FA > 30 \%$ или $FA \leq 30 \%$) использовали метод опорных векторов (support vector machines, SVM) [32]. Обучающая выборка состояла из 480

структурно различных соединений, валидация модели проводилась на основе набора из 98 лекарственно-подобных соединений. В общей сложности использовали 10 дескрипторов, включая $\log P$, $TPSA$, $\log D_{6.5}$, $\log S$, число доноров и акцепторов водорода.

В последнее время все большее применение находит моделирование кишечной абсорбции *in vitro*. В качестве таких модельных систем чаще всего используют монослой культуры клеток Caco-2 и параллельные искусственные мембраны. Примечательно, что данные, полученные в модельных экспериментах *in vitro*, находят применение при разработке *in silico* моделей, прогнозирующих кишечную абсорбцию [22].

В недавнем исследовании K. Nakao et al. предприняли попытку создать *in silico* модели для предсказания проницаемости искусственных параллельных мембран (parallel artificial membrane permeability assay - PAMPA), используя дескрипторы $\log P$, pK_a и $TPSA$ [51]. В работе использована выборка из 71 соединения, было создано 4 QSAR уравнения. По заявлению авторов, эти уравнения могут использоваться для надежного предсказания проницаемости методами *in silico*.

Популярной моделью кишечного эпителия человека является монослой культуры клеток Caco-2. Эта модель находит применение в QSAR исследованиях для прогнозирования кишечной абсорбции. T.J. Hou et al. в одной из работ в серии исследований, посвященных QSAR анализу и прогнозированию различных параметров ADME, изучали корреляцию между проницаемостью клеток Caco-2 и свойствами малых молекул [36]. Установлено, что клеточная проницаемость зависит, в первую очередь, зависит от коэффициента распределения $\log D$ при pH 7.4, от площади молекулярной поверхности, несущей высокий заряд, и от радиуса молекулярного вращения, определяющего объем молекулы.

Сравнительный анализ различных методов, дескрипторов и других связанных параметров, вовлеченных в *in silico* прогнозирование кишечной абсорбции и биодоступности можно найти в работах [31, 44].

Биодоступность – это фракция перорально поступившего препарата, которая в активной форме достигает системного кровотока (F). Биодоступность является одной из важнейших характеристик фармакокинетики, поскольку она напрямую связана с разработкой схем дозирования и приема лекарств.

В работе T. Hou et al. были использованы сведения о биодоступности 768 соединений, почерпнутые из открытой базы данных [10]. Авторами были выполнены QSPR исследования и проведено сравнение данных, полученных расчетным путем, с результатами оценки биодоступности, полученными ранее в опытах на крысах [33]. В

результате критического анализа существующих моделей и сопоставления прогнозов с экспериментальными данными T. Hou et al. пришли к заключению, что ни одно существующее правило не обеспечивает эффективного прогноза биодоступности, поскольку не учитывает влияние метаболических процессов.

В целом, НИА представляет сложное явление, очень зависящее от физико-химических свойств лекарственных соединений. Расширение баз данных и унификация методов измерения гастроинтестинальной абсорбции будут способствовать разработке универсальных прогнозирующих моделей *in silico*. Кроме того, модели, прогнозирующие пероральную абсорбцию, можно рассматривать как начальный шаг к созданию более сложных моделей, позволяющих прогнозировать биодоступность, в частности, при приеме внутрь.

Проницаемость гематоэнцефалического барьера

Проницаемость гематоэнцефалического барьера (ГЭБ) – решающий фактор, определяющий способность лекарственных соединений проникать в ткани головного мозга и оказывать центральное действие. Не случайно, оценка способности лекарств проходить через ГЭБ необходима на ранних стадиях их разработки. Лекарства, оказывающие терапевтическое действие на ЦНС, должны легко преодолевать ГЭБ, тогда как препараты, от которых не ожидают центрального действия, не должны проходить через ГЭБ во избежание развития побочных эффектов. Как известно, липофильные вещества легко проникают через ГЭБ путем диффузии, в то время как для транспорта низколипофильных соединений необходимо участие специальных транспортеров [1]. Несмотря на очевидную зависимость проницаемости от липофильности, моделирование ГЭБ - проницаемости представляет собой достаточно сложную задачу в связи с недостаточностью и чрезвычайной вариабельностью необходимых экспериментальных данных.

Обычно в качестве параметра, характеризующего проницаемость ГЭБ для вещества, используют $\log BB$, представляющий собой отношение равновесных концентраций лекарства в головном мозге и в крови. Как правило, при разработке прогнозирующих моделей используются литературные данные о концентрациях лекарственных веществ в крови и в головном мозге *in vivo*, а физико-химические параметры получают расчетными методами *in silico*. В работе T.J. Hou и X.J. Xu из Пекинского университета использованы данные о проницаемости ГЭБ *in vivo* для 115 органических соединений, на основе анализа которых была разработана QSPR модель,

прогнозирующая значения $\log BB$ [34]. В качестве дескрипторов в модель вошли расчетные значения $\log P$, площадь полярных областей с высоким зарядом (HCPSA) и избыточный молекулярный вес, превышающий 360 (MW_{360}). Установлено, что обширные полярные поверхности оказывают отрицательное влияние на $\log BB$, но при этом частичный заряд на атомах, окружающих полярные поверхности, должен быть выше 0.10|e|. Кроме того, плотная упаковка эндотелиальных мембран ЦНС создает препятствие для прохождения гидрофильных молекул с молекулярным весом свыше 360.

Y.H. Zhao et al., используя выборку из 1593 молекул и 19 простых молекулярных дескрипторов, также подтвердили зависимость $\log BB$ от площади полярной поверхности молекул. Вместе с тем, они показали, что существенную роль играет наличие функциональных групп, способных образовывать водородные связи, которое, в свою очередь, определяется числом доноров и акцепторов водорода. Разработанная модель позволила классифицировать вещества по способности проникать через ГЭБ с точностью до 90% и прогнозировать проницаемость для тестовой выборки с точностью до 95% [72].

S. Vilar et al. классифицировали 307 соединений, для которых были известны параметры проницаемости через ГЭБ в условиях *in vivo* [61]. Используя набор физико-химических дескрипторов, они получили 2 уравнения дискриминантных функций, позволяющие надежно классифицировать соединения. Так, вещества, имеющие значения $\log BB > 0.3$, легко проникали через ГЭБ, тогда как соединения с $\log BB < -1$ плохо распределялись в головном мозге. Высокие значения $\log BB$ определялись высокими параметрами липофильности $\log P$, низкими величинами полярности $TPSA$, а также малым суммарным количеством кислых и основных атомов в молекуле.

Печеночный метаболизм

Метаболическая устойчивость – один важнейших аспектов фармакокинетики, поскольку связан с метаболическим клиренсом соединений. Кроме того, сведения о путях метаболизма вновь создаваемых лекарственных веществ крайне важны для выявления потенциальных канцерогенных и токсических свойств кандидатов в лекарства. В связи с выраженным полиморфизмом ферментов микросомального окисления печени прогнозирование метаболической судьбы лекарственных веществ и ксенобиотиков в целом представляет собой очень непростую задачу. В геноме человека закодировано около 60 изоформ цитохрома P-450 (CYP), однако, более 90% лекарственных веществ метаболизируются при участии пяти изоферментов: CYP 3A4, 2D6, 1A2, 2C9 и 2C19 [66, 74].

В последнее время разработано множество моделей, построенных на разных принципах и позволяющих прогнозировать печеночный метаболизм лекарственных веществ [2, 7, 8, 12, 16]. Так, S. Haji-Memonian et al. на основе QSAR анализа набора данных из 24 соединений создали модель для прогнозирования субстратной специфичности изоформы CYP2D6 [26]. P. Crivori и I. Proggesi приводят исследования, в которых для разработки QSMR-модели (quantitative structure metabolism relationship) использовались экспериментальные значения константы Михаэлиса K_m [16]. В ряде исследований в качестве прогнозируемого параметра использовали доступные в литературе константы ингибирования (IC_{50}) изоформ CYP 3A4, 2C9 и 2D6 [43, 56, 69]. Недостатком ранних моделей, предназначенных для прогнозирования метаболического профиля, считается то, что при разработке использовались малочисленные выборки, а также рассматривалось взаимодействие с ограниченным числом изоферментов [50, 42].

L. Terfloth et al. при разработке QSAR-модели для прогнозирования субстратной специфичности в отношении изоформ CYP 3A4, 2C9 и 2D6 использовали обучающую выборку из 379 лекарств и лекарственных аналогов, набор из 303 дескрипторов и 3 метода анализа данных (множественная логистическая регрессия, деревья решений, метод опорных векторов) [57]. Валидация модели выявила правильность прогнозов для внешней выборки в 83% случаев.

Предлагается также компьютерная модель для прогнозирования микросомального печеночного клиренса у крыс, разработанная на основе огромного массива ($n = 27\,697$) экспериментальных данных [13]. Модель дает прогноз с точностью до 82%, и, как полагают, будет востребована при ADME/T профилировании на ранних этапах разработки новых лекарственных соединений.

Представляет интерес недавняя работа N.K. Mishra et al., в которой для прогнозирования метаболического профиля использовали выборку из 216 препаратов, представленных в базе данных DrugBank [17, 50]. Важным аспектом дизайна исследования является то, что отбирались соединения, для которых известно, что они метаболизируются при участии лишь одной из изоформ CYP. Такой подход позволил достаточно точно охарактеризовать зависимость субстратной специфичности от свойств лигандов. Отбор дескрипторов для разработки модели производился из числа более чем 2000 параметров, которые всесторонне характеризовали исследуемые соединения. По результатам оценки релевантности был отобран 41 параметр, на базе которого разработано 5 уравнений, по одному для каждой из изоформ CYP. В заключение авторы отмечают, что модели, разработанные на базе дескрипторов, рассчитанных с помощью

коммерческих программ, имеют хорошие прогностические характеристики. Вместе с тем, отмечается необходимость совершенствования модели путем пополнения обучающей выборки новыми экспериментальными данными с целью повышения ее репрезентативности. Авторы считают, что для эффективного продвижения предиктивной технологии необходимо создание общедоступных web-серверов, позволяющих специалистам эффективно использовать информационные ресурсы.

Учитывая сложную природу взаимодействий CYPs и лекарственных веществ, а также чрезвычайное структурное многообразие ксенобиотиков, взаимодействующих с ферментами биотрансформации в печени, разработка прогнозирующих *in silico* моделей является чрезвычайно сложной задачей. В связи с этим, любые попытки представляют собой шаг вперед в этом направлении. Вместе с тем, для повышения эффективности предпринимаемых усилий необходимо тесное взаимодействие научного сообщества в рамках информационного поля и, прежде всего, широкий обмен информацией и технологиями.

Литература

1. Альберт А. Избирательная токсичность. Физико-химические основы терапии. В 2-х томах / Пер. с англ. М.А. Думпис, М.Б. Ганиной под ред. проф. В.А. Филова / М.: Медицина. 1989. - 400 с.
2. Бородина Ю.В., Садым А.В., Рудик А.В., Филимонов Д.А., Поройков В.В. Компьютерный прогноз биотрансформации ксенобиотиков. //Аллергия, астма и клиническая иммунология. – 2003 – Т. 7, № 8. – С.85-89.
3. Кинг Р. Химические приложения топологии и теории графов. – М.: Мир. – 1987. – 560 с.
4. Пальм В.А. Основы количественной теории органических реакций. — 2-е, пер. и доп.. — Л.: Химия, 1977. — 360 с.
5. Раевский О.А. Дескрипторы водородной связи в компьютерном молекулярном дизайне. Рос. хим. ж. (Ж. Рос. хим. об-ва им. Д.И.Менделеева) (2006). L (2): 97-107.
6. Раевский О.А. Дескрипторы молекулярной структуры в компьютерном дизайне биологически активных веществ // Успехи химии. – 1999. – Т. 68, № 6. – С. 555-575.
7. Харчевникова Н.В., Дмитриев А.В., Бородина Ю.В., Дьячков П.Н. Квантовохимическая модель для прогноза положения гидроксирования ароматических соединений под воздействием цитохрома Р-450. // Биомедицинская химия. – 2005. – Т.51, №3. – С. 341-355.

8. Arimoto R. Computational models for predicting interactions with cytochrome p450 enzyme // *Curr. Top. Med. Chem.* – 2006. – Vol.6, №15. – P.1609-1618.
9. Baskin I., Varnek A. Fragment Descriptors in SAR/QSAR/QSPR Studies, Molecular Similarity Analysis and in Virtual Screening. In: *Chemoinformatic Approaches to Virtual Screening*, A. Varnek, A. Tropsha, eds., RCS Publishing, 2008, P. 1-43
10. Bioinformatics and CADD. URL: <http://modem.ucsd.edu/adme>.
11. Bonachera F., Parent B., Barbosa F., Froloff N., Horvath D. Fuzzy Tricentric Pharmacophore Fingerprints. 1. Topological Fuzzy Pharmacophore Triplets and Adapted Molecular Similarity Scoring Schemes // *J. Chem. Inf. Model.* – 2006. – Vol. 46, № 6. – P. 2457-2477.
12. Borodina Yu., Rudik A., Filimonov D., Kharchevnikova N., Dmitriev A., Blinova V., Poroikov V. New statistical approach to predicting aromatic hydroxylation sites. Comparison with model-based approaches // *J. Chem. Inf. Comput. Sci.* – 2004. – Vol.44, №6. – P.1998-2009.
13. Chang C., Duignan D.B., Johnson K.D., Lee P.H., Cowan G.S., Gifford E.M., Stankovic C.J., Lepsy C.S., Stoner C.L. The development and validation of a computational model to predict rat liver microsomal clearance // *J. Pharm. Sci.* – 2009. – Vol.98, №8. – P.2857-2867.
14. Comess K.M., Schurdak M.E. Affinity-based screening techniques for enhancing lead discovery // *Curr. Opin. Drug Discov. Devel.* – 2004. – Vol. 7, № 4. – P. 411-416.
15. Cramer R.D., Patterson D.E., Bunce J.D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins // *J. Am. Chem. Soc.* – 1988. – Vol.110, №18. – P.5959-5967.
16. Crivori P., Poggesi I. Computational approaches for predicting CYP-related metabolism properties in the screening of new drugs // *Eur. J. Med. Chem.* – 2006. – Vol.41, №7. – P.795-808.
17. DrugBank. URL: <http://www.drugbank.ca>
18. Fotouhi N., Gillespie P., Goodnow R.A., So S.S., Han Y., Babiss L. E. Application and utilization of chemoinformatics tools in lead generation and optimization // *Comb. Chem. High Throughput Screen.* – 2006. – Vol. 9, № 2. – P. 95-102.
19. Frearson J.A., Collie I.T. HTS and hit finding in academia – from chemical genomics to drug discovery // *Drug Discov. Today* – 2009. – Vol.14, №23-24. – P.1150-1158.

20. Fujita T., Hansch C. Analysis of the structure-activity relationship of the sulfonamide drugs using substituent constants // J. Med. Chem. – 1967. – Vol. 10, № 6. – P. 991-1000.
21. Gao Z., Li H., Zhang H., Liu X., Kang L., Luo X., Zhu W., Chen K., Wang X., Jiang H. PDTD: a web-accessible protein database for drug target identification // BMC Bioinformatics. – 2008. – Vol. 9. – P. 104. URL: <http://www.biomedcentral.com/1471-2105/9/104>
22. Geerts T., Vander Heyden Y. In silico predictions of ADME-Tox properties: drug absorption // Comb. Chem. High Throughput Screen. – 2011. – Vol. 14, № 5. – P. 339-361.
23. Ghose A.K., Viswanadhan V.N., Wendoloski J.J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases // J. Comb. Chem. – 1999. – Vol.1, №1. – P.55-68.
24. Gribbon P., Sewing A. High-throughput drug discovery: what can we expect from HTS? // Drug Discovery Today. – 2005. – Vol.10, №5. – P. 17-22.
25. Guha R. The ups and downs of structure-activity landscapes // Methods Mol. Biol. – 2011. – Vol. 672. – P. 101-117.
26. Haji-Memonian S., Rieger J.M., Macdonald T.L., Brown M.L. Comparative molecular field analysis and QSAR on substrate binding to Cytochrome P450 2D6 // Bioorg. Med. Chem. – 2003. – Vol.11. – P.5545-5554.
27. Hansch C., Fujita T. ρ - δ - π Analysis. A method for the correlation of biological activity and chemical structure // J. Am. Chem. Soc. – 1964. – Vol. 86. – P. 1616-1626.
28. He H.Q., Ma X.H., Liu B., Zhang X.Y., Chen W.Z., Wang C.X., Cheng S.H. High-throughput real-time assay based on molecular beacons for HIV-1 integrase 3'-processing reaction // Acta Pharmacol. Sin. – 2007. – Vol. 28, № 6. – P. 811-817.
29. Helguera A. M., Combes R. D., González M. P., Cordiero M. N. Applications of 2D descriptors in drug design: a DRAGON tale // Curr. Top. Med. Chem. – 2008. – Vol. 8, № 18. – P. 1628-1655.
30. Hong H., Xie Q., Ge W., Qian F., Fang H., Shi L., Su Z., Perkins R., Tong W. Mold(2), molecular descriptors from 2D structures for chemoinformatics and toxicoinformatics // J. Chem. Inf. Model. – 2008. – Vol. 48, № 7. – P. 1337-1344.
31. Hou T., Li Y., Zhang W., Wang J. Recent developments of in silico predictions of intestinal absorption and oral bioavailability // Comb. Chem. High Throughput Screen. – 2009. – Vol. 12, № 5. – P. 497-506.

32. Hou T., Wang J., Li Y. ADME evaluation in drug discovery. 8. The prediction of human intestinal absorption by a support vector machine // J. Chem. Inf. Model. – 2007. – Vol. 47, № 6. – P. 2408-2415.
33. Hou T., Wang J., Zhang W., Xu X. ADME evaluation in drug discovery. 6. Can oral bioavailability in humans be effectively predicted by simple molecular property-based rules? // J. Chem. Inf. Model. – 2007. – Vol. 47, № 2. – P. 460-463.
34. Hou T., Xu X.J. ADME evaluation in drug discovery. 3. Modeling blood-brain barrier partitioning using simple molecular descriptors // J. Chem. Inf. Comput. Sci. – 2003. – Vol.43, №6. – P.2137-2152.
35. Hou T.J., Wang J.M., Zhang W., Wang W., Xu X. Recent advances in computational prediction of drug absorption and permeability in drug discovery // Curr. Med. Chem. – 2006. – Vol. 13, № 22. – P. 2653-2667.
36. Hou T.J., Zhang W., Xia K., Qiao X.B., Xu X.J. ADME evaluation in drug discovery. 5. Correlation of Caco-2 permeation with simple molecular properties // J. Chem. Inf. Comput. Sci. – 2004. – Vol.44, №5. – P.1585-1600.
37. Iwasa J., Fujita T., Hansch C. Substituent constants for aliphatic functions obtained from partition coefficients // J. Med. Chem. – 1965. – Vol. 8. – P. 150-153.
38. Jäger S., Brand L., Eggeling C. New fluorescence techniques for high-throughput drug discovery // Curr. Pharm. Biotechnol. – 2003. – Vol. 4, № 6. – P. 463-476.
39. Karelson M., Lobanov V.S., Katritzky A.R. Quantum-Chemical Descriptors in QSAR/QSPR Studies // Chem. Rev. – 1996. – Vol.96, №3. – P.1027-1044.
40. Katritzky A.R., Fara D.C., Petrukhin R.O., Tatham D.B., Maran U., Lomaka A., Karelson M. The Present Utility and Future Potential for Medicinal Chemistry of QSAR/QSPR with Whole Molecule Descriptors // Curr. Topics Med. Chem. – 2002. – Vol.2. – P.1333-1356.
41. Keller T.H., Pichota A., Yin Z. A practical view of 'druggability' // Curr. Opin.Chem. Biol. – 2006. – Vol. 10, № 4. – P. 357-361.
42. Kharkar P.S. Two-dimensional (2D) in silico models for absorption, distribution, metabolism, excretion and toxicity (ADME/T) in drug discovery // Curr. Top. Med. Chem. – 2010. – Vol. 10, № 1. – P. 116-126.
43. Kriegl J.M., Eriksson L., Arnhold T., Beck B., Johansson E., Fox T. Multivariate modeling of cytochrome P450 3A4 inhibition // Eur. J. Pharm. Sci. – 2005. – Vol. 24, № 5. – P. 451-463.

44. Kuentz M. Prediction of drug absorption: different modeling approaches from discovery to clinical development // *Exp. Rev. Clin. Pharmacol.* – 2009. – Vol. 2, № 3. – P. 217-219.
45. Kümmel A., Parker C.N. The Interweaving of Cheminformatics and HTS // *Methods Mol. Biol.* – 2011. – Vol. 672. – P. 435-457.
46. Lajiness M.S., Vieth M., Erickson J. Molecular properties that influence oral drug-like behavior // *Curr. Opin. Drug Discov. Devel.* – 2004. – Vol. 7, № 4. – P. 470-477.
47. Lipinski C.A., Lombardo F., Dominy B.W., Feeney P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings // *Adv. Drug Del. Rev.* – 2001. – Vol. 46, № 1-3. – P. 3–26.
48. Liu B., Li S., Hu J. Technological advances in high-throughput screening // *Am. J. Pharmacogenomics.* – 2004. – Vol. 4, № 4. – P. 263-276.
49. Mayr L.M., Bojanic D. Novel trends in high-throughput screening // *Curr. Opin. Pharmacol.* – 2009. – Vol. 9, №5. – P. 580-588.
50. Mishra N.K., Agarwal S., Raghava G.P.S. Prediction of cytochrome P450 isoform responsible for metabolizing a drug molecule // *BMC Pharmacology* – 2010. 10:8; URL: <http://www.biomedcentral.com/1471-2210/10/8>
51. Nakao K., Fujikawa M., Shimizu R., Akamatsu M. QSAR application for the prediction of compound permeability with in silico descriptors in practical use // *J. Comput. Aided. Mol. Des.* – 2009. – Vol. 23, № 5. – P. 309-319.
52. Oprea T.I. Property distribution of drug-related chemical databases // *J. Comput. Aided Mol. Des.* – 2000. – Vol. 14, № 3. – P. 251-264.
53. Oprea T.I., Gottfries J. Toward minimalistic modeling of oral drug absorption // *J. Mol. Graph. Model.* – 1999. – Vol. 17, № 5-6. – P. 261-274.
54. Pirogova E., Istivan T., Gan E., Cosic I. Advances in methods for therapeutic peptide discovery, design and development // *Curr. Pharm. Biotechnol.* – 2011. –Vol.12, №8. – P.1117-1127.
55. Selassie C.D. History of Quantitative Structure-Activity Relationships / In: *Burger's Medicinal Chemistry and Drug Discovery*. 6-th Ed., Vol.1: Drug Discovery / Ed. by D.J. Abraham. – John Wiley & Sons, Inc. – 2003. – P. 1-48.
56. Susnow R.G., Dixon S.L. Use of robust classification techniques for the prediction of human cytochrome P450 2D6 inhibition // *J. Chem. Inf. Model.* – 2003. – Vol. 43, № 4. – P. 1308-1315.

57. Terfloeth L., Bienfait B., Gasteiger J. Ligand-based models for the isoform specificity of cytochrome P450 3A4, 2D6, and 2C9 substrates // J. Chem. Inf. Model. – 2007. – Vol. 47, № 4. – P. 1688-1701.
58. Todeschini R., Consonni V. Handbook of Molecular Descriptors. WILEY-WCH Publishers, Weinheim, 2000.
59. Van de Waterbeemd H. High-throughput and in silico techniques in drug metabolism and pharmacokinetics // Curr. Opin. Drug Discov. Devel. – 2002. – Vol. 5, № 1. – P. 33-43.
60. Veber D.F., Johnson S.R., Cheng H.Y., Smith B.R., Ward K.W., Kopple K.D. Molecular properties that influence the oral bioavailability of drug candidates // J. Med. Chem. – 2002. – Vol.45, №12. – P.2615-2623.
61. Vilar S., Chakrabarti M., Costanzi S. Prediction of passive blood-brain partitioning: straightforward and effective classification models based on in silico derived physicochemical descriptors // J. Mol. Graph. Model. – 2010. – Vol.28, №8. – P.899-903.
62. Villoutreix B.O., Renault N., Lagorce D., Sperandio O., Montes M., Miteva M.A. Free resources to assist structure-based virtual ligand screening experiments // Curr. Protein Pept. Sci. – 2007. – Vol. 8, № 4. – P. 381-411.
63. Warr W.A. Some Trends in Chem(o)informatics // Methods Mol. Biol. – 2011. – Vol. 672. – P. 1-37.
64. Waterbeemd H.v.d., Gifford E. ADMET in silico modeling: Towards prediction paradise? // Nat. Rev. Drug Discov. – 2003. – Vol. 2, № 3. – P. 192-204.
65. Wishart D.S., Knox C., Guo A.C., Cheng D., Shrivastava S., Tzur D., Gautam B., Hassanali M. DrugBank: a knowledgebase for drugs, drug actions and drug targets // Nucleic Acid Res. – 2008. – Vol. 36 (Database issue). – D901-D906.
66. Wolf C.R, Smith G., Smith R.L: Science, medicine and the future pharmacogenetics // Br. Med. J. – 2000. – Vol.320. – P.987-990.
67. Yan A., Wang Z., Cai Z. Prediction of human intestinal absorption by GA feature selection and support vector machine regression // Int. J. Mol. Sci. – 2008. – Vol. 9, № 10. – P. 1961-1976.
68. Yao L.X., Wu Z.C., Ji Z.L., Chen Y.Z., Chen X. Internet resources related to drug action and human response: a review // Appl. Bioinformatics. – 2006. – Vol. 5, № 3. – P. 131-139.

69. Yap C.W., Chen Y.Z. Prediction of Cytochrome P450 3A4, 2D6 and 2C9 inhibitors and substrates by using support vector machines // J. Chem. Info. Model. – 2005. – Vol.45. – P.982-992.
70. Zefirov N.S., Palyulin V.A. Fragmental Approach in QSPR // J. Chem. Inf. Comput. Sci. – 2002. – Vol.42, №5. – P.1112-1122.
71. Zhang M.Q., Wilkinson B. Drug discovery beyond the 'rule-of-five' // Curr. Opin. Biotechnol. – 2007. – Vol.18, №6. – P.478-488.
72. Zhao Y.H., Abraham M.H., Ibrahim A., Fish P.V., Cole S., Lewis M.L., de Groot M.J., Reynolds D.P. Predicting penetration across the blood-brain barrier from simple descriptors and fragmentation schemes // J. Chem. Inf. Model. – 2007. – Vol. 47, № 1. – P. 170-175.
73. Zhou J.Z. Chemoinformatics and library design // Methods Mol. Biol. – 2011. – Vol. 685. – P. 27-52.
74. Zhou S.F., Liu J.P., Chowbay B. Polymorphism of human cytochrome P450 enzymes and its clinical impact // Drug Metab. Rev. – 2009. – Vol.41, №2. – P.89-295.