

## БИОИНФОРМАТИКА

УДК 577.15.02:541.6.69

© Балакин, Иваненков

### КОМПЬЮТЕРНЫЕ МЕТОДЫ ПРОГНОЗИРОВАНИЯ МЕТАБОЛИЗМА ФИЗИОЛОГИЧЕСКИ АКТИВНЫХ ВЕЩЕСТВ

*К. В. Балакин, Я. А. Иваненков*

Исследовательский институт химического разнообразия, 141700, Московская область, г. Долгопрудный, ул. Первомайская, 1

Рассмотрено современное состояние исследований в области разработки компьютерных методов прогнозирования метаболизма физиологически активных веществ.

**Ключевые слова:** прогнозирование метаболизма, метаболизирующие ферменты, цитохромы Р450, метаболическая стабильность, метаболические реакции, моделирование количественных соотношений структура-метаболизм (КССМ).

**ВВЕДЕНИЕ.** Вероятно, не будет излишне преувеличенным утверждение о том, что прогнозирование метаболических превращений является одной из наиболее сложных областей исследований в современной медицинской и компьютерной химии. Несмотря на значительный опыт подобных работ, накопленный в течение последних десятилетий, в области прогнозирования наиболее вероятных биотрансформаций мы все еще находимся лишь “у подножия Эвереста”, по меткому выражению Дэвида Хоукинса, одного из основоположников современных компьютерных систем анализа и прогнозирования метаболизма. Мы ясно видим цель, осознаем перспективы, открывающиеся при достижении этой цели, но не знаем лучшей дороги к вершине, и лишь смутно представляем возможные препятствия на этом пути.

В настоящем обзоре авторы попытались представить современное состояние разработки компьютерных методов анализа и прогнозирования метаболических превращений, которым подвергается низкомолекулярное органическое соединение при попадании в организм человека. Рассмотрены наиболее типичные подходы, проанализированы достоинства и недостатки существующих методов, а также указаны перспективные направления исследований.

#### **1. Цели исследований.**

Главной движущей силой в современных исследовательских проектах, направленных на прогнозирование метаболизма ксенобиотиков, является возможность существенного повышения с их помощью эффективности исследований и разработок в фармацевтической индустрии. Метаболические трансформации существенно влияют на биодоступность, биологическую активность, токсичность, продолжительность действия, профиль распределения в организме и экскрецию лекарственного соединения. Как следствие, понимание путей метаболических трансформаций и возможности их предсказания имеют большую практическую ценность на любой стадии длительного и дорогостоящего процесса разработки новых фармацевтических субстанций.

Например, в процессе оптимизации фармакокинетического профиля соединений, проявивших активность на ранних этапах биологических испытаний,

метаболическая стабильность соединения является одним из ключевых требований. Зачастую высокий метаболический клиренс критическим образом ограничивает системную экспозицию соединения, проявившего высокую активность при первоначальном тестировании *in vitro*, уменьшает время его действия и, как следствие, существенно снижает его ценность в качестве потенциального фармацевтического препарата. Эффект “первого прохождения” через печень, при котором многие соединения подвергаются интенсивной метаболической деградации, является еще одним примером влияния метаболизма на фармакокинетические параметры лекарственного соединения. В литературе описано значительное количество примеров, в которых метаболиты активных соединений оказывались высокотоксичными агентами, что препятствовало дальнейшей разработке этих соединений.

Знание потенциальных сайтов метаболизма молекул может быть использовано для химической модификации структуры таким образом, чтобы удалить этот сайт, либо защитить его, например, путем уменьшения его стерической доступности. Хрестоматийными примерами на эту тему являются замена сложноэфирной группы прокаина на амидную группу, ведущая к метаболически устойчивому прокаинамиду, или введение объемной *N*-изопропильной группы (вместо *N*-метильной или *N*-этильной) вблизи сайта метаболической реакции *N*-деалкилирования известных ингибиторов  $\beta$ -адренорецепторов [1]. Современная практика также дает нам немалое число примеров оптимизации метаболического профиля соединений в процессе разработки новых лекарственных субстанций (см., например, [2-8]).

Следует упомянуть еще одно важное направление в современной медицинской химии, связанное с использованием информации о путях метаболических трансформаций. Нередко реакции метаболизма используются с целью направленной трансформации неактивного соединения – пролекарства - *in vivo* в его биологически активную форму [9]. В большинстве случаев для этого используют сложноэфирную группу, маскирующую карбоксильный или гидроксильный фрагменты активного соединения [10, 11]. Наоборот, введение метаболически нестабильных элементов в структуру активных соединений позволяет обеспечить кратковременность действия лекарства и тем самым устранить возможные побочные эффекты. Примером такого рода является блокатор  $\beta$ -адренорецепторов эсмолол, обладающий ультракоротким временем действия [12]. Подобные манипуляции с метаболическими системами организма являются эффективным средством достижения желаемых фармакокинетических параметров.

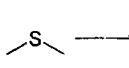
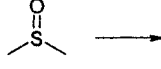
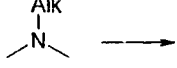
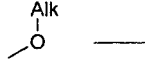
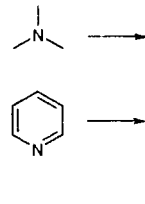
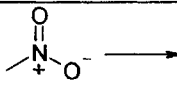
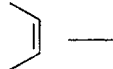

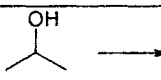
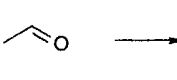
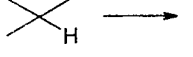
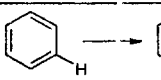
Приведенные примеры, на наш взгляд, в достаточной мере иллюстрируют высокую практическую значимость знаний о путях вероятных метаболических трансформаций органических соединений. Прогнозирование путей метаболизма может быть реализовано лишь на основе знаний о метаболических ферментных системах организма человека.

## 2. Метаболические ферментные системы организма.

Ферменты метаболизма являются уникальным инструментом детоксификации гидрофобных ксенобиотиков, попадающих в организм человека. Гидрофобные соединения, а к таковым относятся, в частности, большая часть лекарственных субстанций, превращаются в более гидрофильные молекулы. Последние, как правило, быстрее выводятся из организма. Подобная биотрансформация является ключевой стадией защитного процесса детоксикации. Главным органом, ответственным за детоксикацию и выведение из организма ксенобиотиков, является печень. В нормально функционирующей печени протекает два типа метаболических реакций, которые принято классифицировать как реакции фаз I и II [13]. Реакции фазы I превращают исходные гидрофобные молекулы в более полярные метаболиты посредством окисления, восстановления или гидролиза. В ходе этих реакций обычно появляются полярные

функциональные группы, такие как  $-\text{OH}$ ,  $-\text{NH}_2$ ,  $-\text{SH}$  и  $\text{COOH}$ . Типичные примеры реакций фазы I, протекающих при участии цитохромов P450, приведены в таблице 1. Во многих случаях в результате реакций фазы I происходит лишь незначительное увеличение гидрофильности молекул, недостаточное для их быстрого выведения. В этом случае метаболиты первой фазы подвергаются превращениям фазы II, которые представляют собой реакции конъюгации с такими молекулами, как глюкуронат, сульфат, ацетат или аминокислота. В некоторых случаях реакции фазы II могут предшествовать реакциям фазы I. Результатом метаболических превращений фаз I и II являются легко выводимые из организма гидрофильные метаболиты.

Таблица 1 Наиболее важные реакции, катализируемые цитохромами P450.

Реакция	Схема
Окисление S(II)	
Окисление S(IV)	
N-деалкилирование	
O-деалкилирование	
Образование N-оксида	
Восстановление нитрогруппы	
Эпоксидирование	
Образование двойной связи	
Окисление спиртов	
Окисление альдегидов	
Алифатическое гидроксирование	
Ароматическое гидроксирование	

## ПРОГНОЗИРОВАНИЕ МЕТАБОЛИЗМА ФИЗИОЛОГИЧЕСКИ АКТИВНЫХ ВЕЩЕСТВ

Организм человека постоянно сталкивается с необходимостью метаболической детоксификации огромного структурного разнообразия ксенобиотиков. Вероятно, именно этим обусловлено наличие весьма значительного числа различных метаболических ферментов с существенно отличающимся профилем субстратной специфичности. Интенсивные исследования последних десятилетий показали, что в метаболизме фазы I участвуют ферменты семейства цитохромов P450 [14,15], а также ряд других ферментов, совместно обладающих очень широкой субстратной специфичностью и многообразием каталитических механизмов [16-21] (табл. 2). К метаболическим ферментам фазы II относят различные трансферазы, наиболее важные из которых перечислены в таблице 3.

Таблица 2 Главные метаболические ферменты фазы I.

Фермент	Реакция	Источник
Цитохромы P450	Реакции окисления (табл. 1)	[14,15]
Флавин-содержащая монооксигеназа (FMO)	N- (или S-) окисление	[16]
Моноаминоксидаза (MAO)	Окислительное дезаминирование	[17]
Алкогольдегидрогеназа	Окисление спиртов	[18]
Альдегиддегидрогеназа	Окисление альдегидов	[19]
Альдегидоксидаза	Окисление альдегидов	[20]
Ксантиноксидаза	C-окисление	[21]

В настоящее время общепризнано, что цитохромы P450 являются главными ферментативными системами биотрансформации ксенобиотиков. В то же время, многие другие ферменты, такие как флавин-содержащая монооксигеназа, алкогольдегидрогеназа, моноаминоксидаза и пероксидазы, способны катализировать реакции, типичные для цитохромов P450. Как следствие, не всегда можно точно оценить относительную роль отдельной ферментативной системы в экспериментах *in vivo*.

Таблица 3. Главные метаболические ферменты фазы II.

Фермент	Реакция	Источник
Глутатион-S-трансфераза (GST)	Присоединение остатка глутатиона	[22,23]
Сульфотрансферазы (ST)	Присоединение сульфатной группы	[23,24]
N-ацетилтрансферазы (NAT)	Присоединение ацетильной группы	[23,25]
Уридиндифосфат-глюкуронозилтрансферазы (UDP-GT)	Присоединение глюкуроновой группы	[26]
Метилтрансферазы	Присоединение метильной группы	[27]

### 3. Экспериментальные методы исследования метаболических превращений.

Выявление существенного влияния метаболических процессов на фармакологический профиль лекарственного соединения привело к разработке целого ряда экспериментальных методов исследования метаболических трансформаций. Наш обзор был бы неполным без краткого упоминания важнейших и наиболее часто используемых в настоящее время технологий.

#### 3.1. Методы оценки метаболической стабильности *in vitro*.

Для проведения данного анализа было разработано несколько систем *in vitro*, включая срезы печени [28-31], изолированные гепатоциты [32-34], субклеточные фракции мембран гепатоцитов [35] и очищенные ферменты P450 [36]. В последнее время появился ряд сообщений об адаптации существующих методик к высокопроизводительным форматам анализа. Так, были разработаны автоматизированные системы анализа для 96-луночных плат, комбинированные с LC/MS-детекцией (жидкостная хроматография, совмещенная с масс-спектрометрической идентификацией продуктов разделения, от англ. Liquid Chromatography/Mass-Spectrometry) [36-39]. С внедрением подобных технологий становится возможным исследовать метаболическую стабильность больших серий соединений. В некоторых методиках метаболическую стабильность соединений оценивают по скорости исчезновения исходного соединения в присутствии микросом или гепатоцитов. В результате подобных экспериментов исследуемые соединения обычно относят к трем классам: стабильные, умеренно стабильные и нестабильные, – используя в качестве сравнения лекарственные соединения с известными параметрами метаболизма. Как правило, соединения с низкой метаболической стабильностью считаются малоперспективными для дальнейших фармакологических испытаний.

#### 3.2. Структурная идентификация метаболитов.

Структурная идентификация метаболитов лекарственных соединений относится к классическим сферам исследований в фармакологии. В настоящее время для этих целей также используются *in vitro* тесты, поскольку, как правило, метаболические профили соединений *in vitro* и *in vivo* в значительной степени совпадают. Наиболее популярными являются системы с LC/MS детекцией [40-43]. О структурах метаболитов, появляющихся в реакционной смеси, судят по появлению в масс-спектрах пиков, соответствующих вероятным метаболитам. Например, для соединения с молекулярной массой М появление пика М-14 соответствует продукту деметилирования, пики М+16 и М+176 свидетельствуют о появлении гидроксированного и глюкуронидного продукта. Более сложные технологии используют методы корреляционного анализа [44, 45], а также ультрафильтрационную масс-спектрометрию [46-48] для оценки метаболитов в сложных смесях, нередко образующихся в результате ферментативных реакций. Описан также автоматизированный метод, при котором анализируемые образцы разделяют хроматографически с последующей жидкостной экстракцией для выделения метаболитов; идентификацию последних затем проводят при помощи ЯМР-спектроскопии [49]. По мере внедрения нового метода LC/NMR (жидкостная хроматография, совмещенная с ЯМР-спектроскопическим способом идентификации продуктов разделения, от англ. Liquid Chromatography/Nuclear Magnetic Resonance), все более перспективными считают комбинированные LC/MS и LC/NMR системы анализа [50-52].

#### 3.3. Идентификация метаболических ферментов.

Для идентификации специфических ферментов, ответственных за метаболизм лекарственных соединений, было разработано несколько экспериментальных подходов [53]. Наибольшее распространение получили следующие методы: (а) метаболизм на микросомах, содержащих кДНК-экспрессированные ферменты, (б) использование неспецифичных микросом совместно с селективными низкомолекулярными ингибиторами, (в)

ингибирование фермент-специфичными антителами. Обычно для точной идентификации метаболических систем требуется сочетание указанных методов. В настоящее время также описаны генно-инженерные системы для клонирования и экспрессии многих других метаболических ферментов, таких как флавиномонаооксигеназа, глутатион-S-трансфераза, UDP-гликозилтрансфераза и др.

#### 4. Компьютерные методы прогнозирования метаболизма.

Начиная с этого раздела, мы переходим к основной теме нашего обзора, посвященного разработке компьютерных систем прогнозирования метаболизма. Важно отметить, что их развитие в значительной мере повторяет пути разработки экспериментальных моделей, что отражает их близкую взаимосвязь. Интенсивное развитие компьютерных предсказательных технологий (или технологий *in silico*) призвано усилить и в ряде случаев заменить традиционные экспериментальные подходы. В перспективе это позволит с наибольшей эффективностью войти в новую эпоху в индустрии разработки лекарственных средств, базирующуюся на высокопроизводительных системах синтеза и биологического скрининга.

##### 4.1. Базы данных и экспертные системы.

Наш экскурс в область систем прогнозирования метаболизма следует начать с описания крупных баз данных (БД) и экспертных систем (ЭС), совмещенных с базами знаний по метаболическим превращениям, играющих центральную роль в большинстве современных исследований в этой области. Прошло уже не менее двух десятилетий с момента появления первых коммерческих баз данных, совмещенных с различными экспертными алгоритмами для прогнозирования структур метаболитов. Некоторые наиболее известные коммерческие продукты представлены в таблице 4.

Таблица 4. Коммерческие базы данных и экспертные системы с базами знаний по метаболизму.

Название	Разработчик/ дистрибьютор	Число соединений	Число биотрансформаций
<i>БД</i>			
Metabolite	MDL Information Systems, Великобритания www.mdl.com	~40 тыс.	~64 тыс.
Metabolism Database	Accelrys, Inc., США www.accelrys.com	~6 тыс.	~40 тыс.
<i>ЭС</i>			
MetabolExpert	CompuDrug, Inc., США www.compudrug.com	-	-
META	Multicase, Inc., США www.multicase.com	-	~1500
METEOR	LHASA, Ltd., Великобритания www.chem.leeds.ac.uk/luk/meteor/	-	-
MetaDrug	GeneGo, Inc., США www.genego.com	~16 тыс.	~ 8 тыс.

Примечание: Данные на середину 2004 г. Прочерк в ячейке означает отсутствие опубликованных данных.

База данных Metabolite [54] предоставляет доступ к широкому набору сведений по метаболическим превращениям. Однако отмечалось, что большой объем содержащегося в ней материала сопровождается определенным снижением его качества [55]. БД Metabolism [56] также обеспечивает доступ к значительному объему информации, которая, по утверждению ее создателей, проходит тщательную экспертную проверку.

Программы MetabolExpert [57, 58] и META [59-61] являются развитыми экспертными системами, созданными на базе тщательного многолетнего отбора экспериментальных материалов из относительно узкого круга специализированных источников. В силу исторических причин, эти системы содержат базы знаний по метаболизму в различных видах млекопитающих. Как следствие, любые алгоритмы прогнозирования метаболизма, основанные на обучающих выборках из этих комбинированных баз, имеют тенденцию к прогнозированию возможных метаболических превращений в некоем гипотетическом, "усредненном" млекопитающем [58]. Еще одним примером экспертной системы, совмещенной с базой знаний для оценки продуктов метаболизма, является программа METEOR [62]. Создатели программы MetaDrug [63] позиционируют ее как новый вид продукта в рассматриваемой сфере, оперирующий с аппаратом системной биологии и интеллектуальными алгоритмами прогнозирования ряда свойств веществ, связанных с метаболизмом, включая их токсичность.

Все указанные здесь системы имеют встроенные возможности поиска по подструктуре, которые позволяют найти опубликованные сведения о возможных метаболических превращениях различных структурных фрагментов. Современные версии программных комплексов MetabolExpert, META и METEOR содержат также специальные алгоритмы, позволяющие генерировать выборки потенциальных метаболитов на основе наборов правил метаболических трансформаций, а также некоторых лимитирующих условий.

Типичным и одним из наиболее современных примеров такого рода, иллюстрирующих работу указанных систем, может служить алгоритм действий программы META [64]. Программа работает во взаимодействии с базой данных, в которой биологические трансформации отнесены к конкретному метаболизирующему ферменту. Эта информация хранится в виде мишень/трансформант, где мишень представляет собой молекулярный фрагмент, который, как предполагается, узнается специфическим ферментом, а трансформантом является структура мишени, подвергшейся метаболическому превращению. Важно при этом отметить, что каждая пара мишень/трансформант имеет определенную степень приоритета, основанную на оценке частоты встречаемости превращений данного типа, полученной с использованием генетического алгоритма [61].

В качестве примера приведем трансформацию группы  $\text{CH}_3\text{-N-NO}$ . Эта группа, принадлежащая, например, диметил-*N*-нитрозамину, распознается программой как мишень для модуля цитохромов P450. Соответствующим трансформантом будет  $\text{HO-CH}_2\text{-N-NO}$ . Используемое правило выглядит следующим образом:

**Найти:**  $\text{CH}_3\text{-N-NO}$  и заменить  
**на:**  $\text{CH}_2\text{-N-NO} <1\text{-OH}>$

Молекула, являющаяся результатом этой трансформации диметил-*N*-нитрозамина, будет выглядеть так:  $\text{HO-CH}_2\text{-N(-NO)-CH}_3$ . После того, как к данной молекуле применены все имеющиеся трансформации, специальный модуль проверяет молекулы на стабильность. Для этого сначала распознаются продукты, содержащие определенные нестабильные сочетания атомов, а затем для этих продуктов рассчитываются наиболее стабильные их формы при помощи методов

молекулярной механики. Например, в случае *N*-нитрозо-соединений этот модуль сигнализирует о том, что группа HO-CH<sub>2</sub>-N-NO является в общем случае нестабильной и что она расщепляется на фрагменты O=CH<sub>2</sub> и N=N-OH. Таким образом, продукт первого уровня метаболизма диметил-*N*-нитрозамина будет трансформирован в формальдегид CH<sub>2</sub>O и CH<sub>3</sub>-N=N-OH в соответствии со следующим правилом:

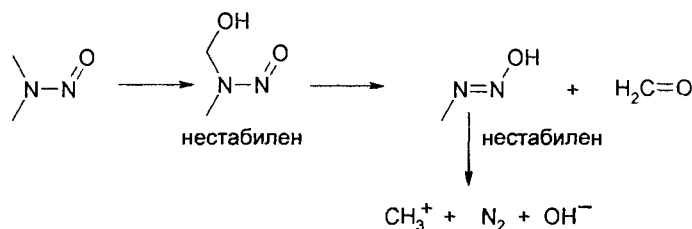
Найти:  $\text{HO}-\text{CH}_2-\text{N}-\text{N}=\text{O}$  и заменить

на:  $\text{O}=\text{CH}_2$ ,  $\text{N}=\text{N}-\text{OH}$

Далее программа вновь обращается к модулю стабильности до тех пор, пока не останутся только стабильные метаболиты. Модуль стабильности работает в фоновом режиме, и по умолчанию программа выводит в конечном итоге только стабильные первичные метаболиты. В то же время, интерфейс программы позволяет при необходимости отследить все промежуточные трансформации.

После определения первичных метаболитов программа по запросу пользователя может начать определять вторичные метаболиты. В результате подобного итеративного процесса программа генерирует дерево метаболизма. Процесс может быть автоматически прекращен после поступления сигнала о том, что полученные метаболиты достигают определенной степени гидрофильности, при которой наиболее вероятным событием будет их быстрая экскреция из организма. Программа МЕТА использует для этого специально разработанный авторами модуль для расчета величины  $\log P$ , основанный на оценке индивидуальных вкладов отдельных фрагментов молекулы [65]. Предполагается, что при достижении определенной величины  $\log P$  соединение теряет способность к реабсорбции в почках и, как следствие, выводится с мочой. Хотя подобная зависимость действительно существует, но для разных типов соединений (основания, нейтральные молекулы, кислоты) пороговые значения  $\log P$  могут существенно отличаться [66]. Кроме того, очевидно, существует определенное влияние на реабсорбцию и других параметров молекул, таких как размер, форма, площадь полярной поверхности и пр. Дополнительным фактором, вносящим неопределенность в момент завершения метаболизма, например, ферментами фазы I, является возможность быстрого метаболизма ферментами фазы II, при котором метаболиты становятся высокогидрофильными соединениями и быстро выводятся из организма.

На рисунке 1 показано метаболическое дерево для диметил-*N*-нитрозамина, сгенерированное программой META.



**Рисунок 1.**

Метаболизм диметил-*N*-нитрозамина, предсказанный программой МЕТА.

Сходные алгоритмы используются большинством других экспертных систем для оценки продуктов метаболизма.

Одной из типичных проблем в работе подобных экспертных систем является определение наиболее вероятных сайтов метаболизма, что до сих пор является в значительной мере нерешенной задачей [1]. Более подробно этот вопрос будет



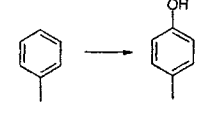
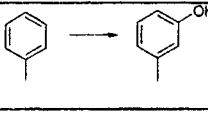
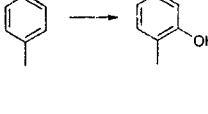
рассмотрен в одном из следующих разделов настоящего обзора. Здесь же нам следует упомянуть о возможном подходе к решению этой проблемы, основанном на относительно простом статистическом алгоритме.

Принцип приоритизации может быть проиллюстрирован на примере гидроксилирования толуильного фрагмента в *пара*-, *орто*- и *мета*- положения [67]. Все расчеты проводились с использованием базы Metabolite. Посредством встроенной возможности подструктурного поиска был проведен поиск всех возможных метаболических реакций гидроксилирования, протекающих в различные положения фенильной группы. Результаты поиска представлены в таблице 5. Первый столбец показывает сумму всех реакций ароматического гидроксилирования, содержащихся в базе, в положение  $i$  ( $\Sigma t_i$ ). Второй столбец показывает общее число структур исходных субстратов метаболических реакций в базе Metabolite, содержащих толуильную группу; она одинакова для всех трех рассматриваемых случаев ( $\Sigma s_i$ ). В третьем столбце приведено отношение числа реакций к числу потенциальных субстратов ( $\Sigma t_i / \Sigma s_i$ ), а в четвертом столбце - значения этой же величины, нормализованные относительно максимальной из них, принятой за 1. Эти нормализованные значения вероятности протекания данной реакции ( $P_i$ ), рассчитанные в соответствии с уравнением 1,

$$P_i = (\Sigma t_i / \Sigma s_i) / (\Sigma t_\alpha / \Sigma s_\alpha) \quad (1),$$

где соотношение ( $\Sigma t_\alpha / \Sigma s_\alpha$ ) равно наибольшей вероятности, могут служить для статистической оценки приоритетности того или иного сайта гидроксилирования толуильной группы. Как видно из таблицы 5, наиболее вероятным является гидроксилирование в *пара*-положение.

Таблица 5. Статистический подход к приоритизации сайтов ароматического гидроксилирования толуильного фрагмента.

Трансформация	$\Sigma t_i$	$\Sigma s_i$	$\Sigma t_i / \Sigma s_i$	$P_i$
	767	5859	0,1309	1,0
	71	5859	0,012	0,0926
	34	5859	0,006	0,0443

Подобный подход может быть рекомендован в качестве простого способа определения наиболее вероятного сайта метаболизма для большого числа структурных фрагментов. Очевидным его недостатком является жесткая привязка результата к содержимому конкретной базы данных. Необходимым условием обеспечения приемлемого качества прогнозирования является достаточно большое число трансформаций интересующего нас типа в доступной базе. Следует отметить, что по этой причине из указанных в таблице 4 баз данных и экспертных систем, вероятно, наиболее эффективно этот алгоритм будет работать с базой Metabolite, являющейся наиболее представительной. Другим принципиальным недостатком этого способа является тот факт, что в нем не учитываются структура и свойства молекулы в целом, например, стерическое микроокружение сайта гидроксилирования, что может приводить к весьма

существенным ошибкам прогнозирования.

В разделе “Исследования фермент-специфичности субстратов метаболических реакций” будет рассмотрен усовершенствованный алгоритм статистического анализа сайтов метаболических реакций при помощи баз Metabolism и Metabolite, описанный в недавней работе отечественных исследователей [68]. Еще один статистический подход к приоритизации метаболических реакций, реализованный в одной из последних версий программы METEOR, изложен в работе [69].

Можно констатировать, что главными недостатками рассмотренных методов, реализованных в экспертных системах, являются следующие: (1) несовершенные алгоритмы оценки способности молекулы служить субстратом метаболической реакции; как следствие, программы могут неверно находить конечные продукты метаболических превращений, а также оценивать “изначальную склонность” соединения к метаболической деградации (метаболическую стабильность); (2) при наличии в молекуле нескольких потенциальных сайтов метаболизма (что на практике является наиболее типичным случаем) рассмотренные программы редко способны дать корректную оценку наиболее вероятных направлений метаболических превращений; особенно сложным является определение положения ароматического и алифатического гидроксирования; (3) существующие программы не способны оценить скорость протекания той или иной метаболической реакции.

В целом необходимо отметить, что, по мере накопления новых данных по метаболизму ксенобиотиков и улучшения качества информации, базы данных по метаболическим превращениям будут играть все более важную роль в аналитических исследованиях. Накопленные за прошедшие годы массивы данных уже обладают значительным потенциалом для поддержки исследований в сфере прогнозирования метаболизма, и большое число теоретических исследований в данной области основано на их анализе. Очевидно, что в будущем можно ожидать их дальнейшее количественное и качественное развитие, наряду с внедрением дополнительных усовершенствованных алгоритмов анализа данных, приспособленных для решения определенных типов задач. Наиболее типичным задачам такого рода будет посвящена следующая часть настоящего обзора.

#### 4.2. *Метаболическая стабильность соединений.*

Методы прогнозирования метаболической стабильности соединений, хотя и не ведут напрямую к структурам потенциальных метаболитов, играют все более заметную роль в современных исследованиях метаболизма и фармакокинетики лекарственных веществ [70,71]. Подтверждением значимости такого рода информации может служить значительное число экспериментальных методов анализа метаболической стабильности, кратко рассмотренных выше. Долгое время этим экспериментальным процедурам практически не было альтернативы, однако ситуация стала меняться с появлением высокопроизводительных методов синтеза, а также с все более широким внедрением концепции виртуального скрининга. В настоящее время уже ни у кого не вызывает сомнений тот факт, что при помощи рациональных методов *in silico* скрининга виртуальных баз соединений можно отбирать для последующего синтеза и анализа наиболее интересные и перспективные структуры. Качественное прогнозирование метаболической стабильности соединений, являющейся одним из важнейших факторов, влияющих на фармакологический профиль лекарственного препарата, может эффективно служить целям рациональной селекции в программах виртуального скрининга. Кроме того, как мы уже отмечали в предыдущем разделе, подобные предсказания могут служить ценным дополнением в работе экспертных систем для прогнозирования структур метаболитов.

Практически все известные нам модели появились в течение последних нескольких лет. Так, был проведен компьютерный анализ свойств 875 разнообразных соединений с известными величинами метаболической

стабильности, определенными в результате эксперимента с микросомами печени человека *in vitro* [72]. Была построена компьютерная модель с использованием алгоритма рекурсивного разбиения (англ. Recursive Partitioning), которая обладала неплохой предсказательной способностью: квадрат коэффициента корреляции между экспериментальными и предсказанными значениями ( $r^2$ ) для внутренней валидационной выборки оказался равным 0,71. Полученная модель была с успехом использована для прогнозирования системного клиренса 41 соединения [72].

Сообщалось также об использовании другого экспериментального набора из 500 соединений, разнообразных по химическому строению, в моделировании метаболической стабильности [73]. Каждое соединение инкубировали с микросомальным препаратом, содержащим экспрессированный цитохром CYP3A4. После соответствующей обработки определяли процентное содержание соединения (относительно его исходного количества) в анализируемом образце при помощи LC/MS анализа. Образцы соединений, содержащие не менее 90% исходного вещества, относили к стабильным; нестабильными считались образцы, содержащие менее 20% исходного материала. Компьютерный анализ этой экспериментальной базы позволил эффективно разделить указанные категории соединений. Сначала при помощи программы VolSurf были рассчитаны специальные 3D дескрипторы, а затем была проведена процедура анализа главных компонент [74]. На рисунке 2 показана двумерная карта распределения стабильных и нестабильных соединений в координатах двух первых главных компонент. Хотя некоторая часть нестабильных соединений (черные кружки) попала в зоны стабильных, тем не менее, совершенно отчетливо видны зоны преимущественной локализации соединений, принадлежащих двум указанным категориям.

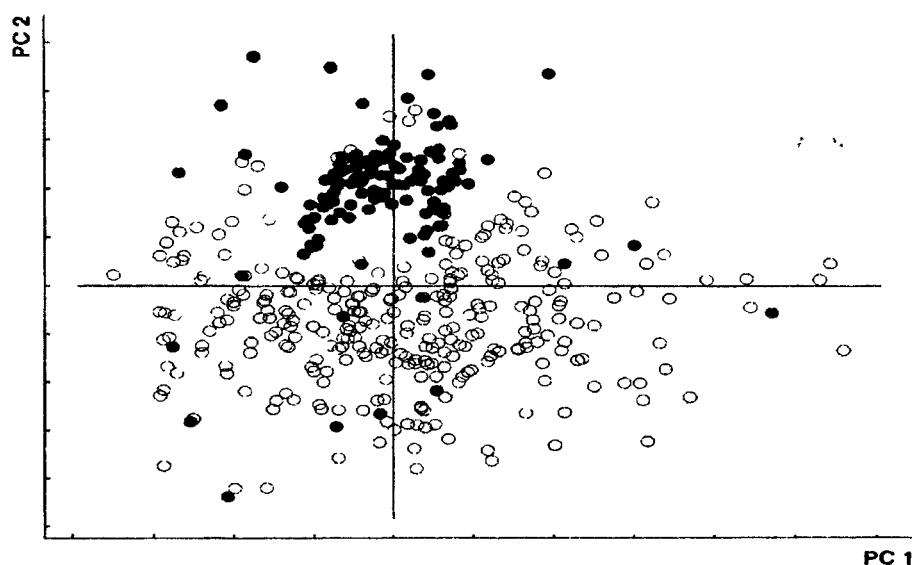


Рисунок 2.

Распределение метаболически стабильных (белые кружки) и нестабильных (черные кружки) соединений в координатах двух первых главных компонент (по [73] с изменениями).

Недавно была опубликована работа по прогнозированию метаболической стабильности [75], основанная на алгоритме  $k$ -ближайших соседей (англ.,  $k$ -Nearest Neighbors - KNN). Метод основан на принципе существования подобия профиля биологической активности соединений (в данном случае, метаболического профиля), обладающих подобными структурными и физико-химическими свойствами. Каждое соединение описывают некоторым набором

молекулярных дескрипторов, после чего рассчитывают степень подобия между соединениями как Евклидово расстояние между объектами в многомерном пространстве свойств. В указанной работе метаболическую стабильность каждого соединения рассчитывали как средневзвешенную активность его  $k$  ближайших соседей (т. е. наиболее подобных соединений) в обучающей выборке. В качестве обучающей выборки использовали 631 соединение с экспериментально определенной метаболической стабильностью в присутствии клеточного гомогената печени человека. В соответствии с относительным количеством исходного соединения, оставшегося в реакционной смеси после инкубирования, все соединения были разделены на четыре группы: стабильные ( $>75\%$  исходного соединения), умеренно-стабильные (50-75%), умеренно-нестабильные (25-50%), нестабильные ( $<25\%$ ).

В качестве расчетных молекулярных параметров использовали дескрипторы типа MolConnZ [76] и атомно-парные (AP) [77]. Дескрипторы первого типа включают в себя различные топологические индексы, индексы электротопологического состояния, индексы Винера и Платта, индексы формы, степени разветвленности, а также ряд других. Ключевыми компонентами при расчете AP-дескрипторов являются определенные элементы в структуре молекулы. К таковым, например, относятся, отрицательно заряженный центр, положительно заряженный центр, акцептор водородной связи, донор водородной связи, центр ароматического кольца, атом азота, центр тройной связи, центр двойной связи и пр. Атомно-парные дескрипторы кодируют расстояния между двумя такими элементами, в виде наименьшего числа атомов, разделяющих эти элементы.

В результате экспериментов с различными типами дескрипторов и параметров настройки, были сконструированы модели, позволяющие с хорошей точностью предсказывать соединения, относящиеся к крайним классам, то есть стабильным и нестабильным. В частности, для модели, построенной на 30 MolConnZ дескрипторах, точность прогнозирования стабильных и нестабильных соединений равнялась, соответственно 84% и 75%. При этом использование MolConnZ и AP дескрипторов приводило к сопоставимым результатам. Полученные модели были успешно оценены при помощи внешней тестовой выборки, содержащей 107 соединений с экспериментально определенной метаболической стабильностью.

Во всех описанных случаях в качестве обучающих выборок использовались данные *in vitro* экспериментов, проведенных на базе крупных фармацевтических корпораций. В качестве альтернативного подхода была описана нейронно-сетевая модель метаболической стабильности [78], основанная на обучающей выборке, заимствованной из электронной базы данных MetaDrug компании GeneGo. Были использованы данные для более чем 1000 соединений, являющихся субстратами или продуктами метаболических реакций, катализируемых 38 различными цитохромами CYP450 человека. Все соединения были разделены на две категории: “субстраты” и “несубстраты”, являющиеся близкими аналогами нестабильных и стабильных соединений. Для каждого из соединений были рассчитаны десятки молекулярных дескрипторов; в результате направленного отбора при помощи анализа главных компонент (PCA) были отобраны шесть дескрипторов, участвующих в построении конечной модели.

Затем была построена карта Кохонена для базы “субстратов” (485 соединений) и “несубстратов” (523 соединения). Алгоритм самоорганизующихся карт Кохонена [79] позволяет построить двумерное отображение многомерного пространства свойств для исследуемых объектов. При этом объекты, близко расположенные на карте Кохонена, обладают близкими свойствами в исследуемом многомерном пространстве. Эта интересная методология часто позволяет выявлять тонкие различия в свойствах различных групп объектов, в частности, наличие кластеров. Кроме того, этот алгоритм совместим с высокопроизводительными методами виртуального скрининга: после того, как

карта Кохонена построена, с ее помощью можно гестировать в реальном времени даже очень большие виртуальные библиотеки. В описываемом случае при помощи карты Кохонена удалось выявить выраженные зоны преимущественной заселенности "субстратов" и "несубстратов", в которые попали соответственно 77% и 63% соединений, принадлежащих одноименной категории. Внутри этих зон относительное преобладание соединений определенной категории составляло 85-90%.

Методология Кохонена предоставляет широкие возможности для анализа и визуализации данных. Так, на рисунке 3 представлена карта Кохонена, в которой контурная линия определяет зоны преимущественной локализации "субстратов", то есть метаболически нестабильных соединений, а градиентный фон отражает величину параметров HBD и  $\log D_{7.4}$ . Видно, что способность вещества служить субстратом метаболической реакции в общем случае положительно коррелирует с уменьшением числа доноров водородной связи, а также с ростом липофильности.

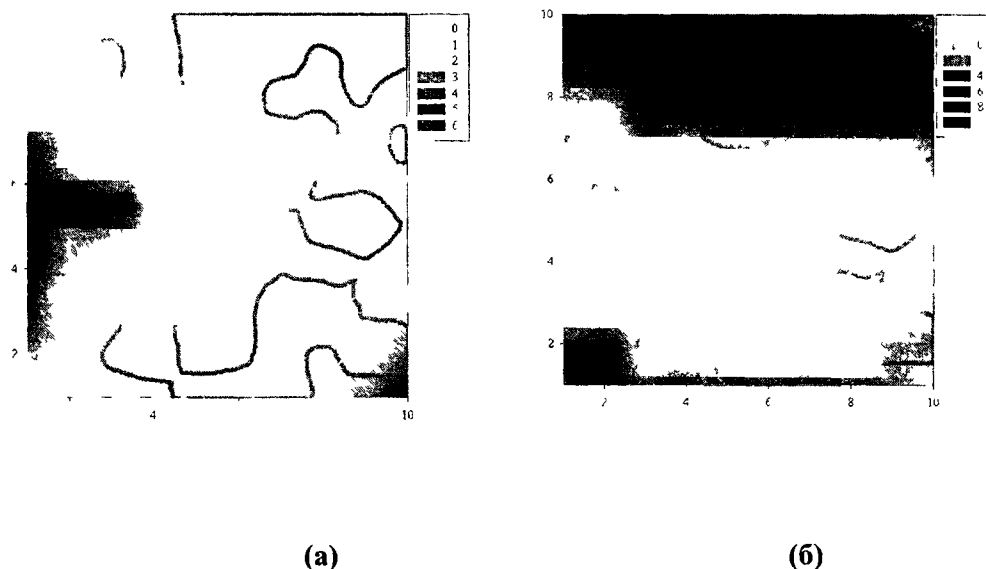


Рисунок 3.

Градиентные карты Кохонена параметров HBD (а) и  $\log D_{7.4}$  (б). Контур ограничивает зоны преимущественной локализации метаболически нестабильных соединений.

Построенная модель также позволила сделать гипотезы о фермент-специфичных зонах на карте Кохонена, то есть зонах, в которые преимущественно попадают субстраты определенных ферментов семейства цитохромов P450. В настоящий момент ведется работа по усовершенствованию описанного алгоритма и его интеграции с экспертной системой по прогнозированию метаболизма лекарственных соединений компании GeneGo (некоторые характеристики и особенности этой системы описаны в недавнем обзоре [80]).

#### 4.3 Прогнозирование наиболее вероятных сайтов метаболизма

Прогнозирование наиболее вероятных сайтов метаболизма, то есть участков молекулы, которые будут подвергаться метаболическим трансформациям, является весьма важной и трудной задачей. Предлагались различные варианты решения указанной проблемы. В одном из подходов, рассмотренном выше, был предложен статистический анализ определенных подфрагментов метаболических реакций с использованием специальных баз данных [67]. Как отмечалось, этот метод отличается простотой и в некоторых случаях приводит к неплохим результатам, но имеет ряд существенных принципиальных недостатков.

Современные подходы развиваются в направлении анализа соответствия

трехмерных структур потенциальных субстратов структурам активных центров метаболических ферментов, что позволяет выявлять фрагменты молекул, оказывающиеся в непосредственном контакте с реакционной частью фермента (например, с гемовой оксиферрильной группой цитохромов P450). Другой подход предполагает анализ энергетики метаболических реакций, позволяющий выявить фрагменты молекулы субстрата, в наибольшей степени подверженные метаболическим реакциям определенного типа. В некоторых работах самого последнего времени предложены варианты комбинирования обеих вышеуказанных стратегий, что, вероятно, способно обеспечить наивысшую точность прогнозирования, хотя и ценой существенных затрат вычислительных ресурсов.

4.3.1. *Оценка энергии разрыва связи в ходе метаболической реакции.* Исторически первыми появились методы прогнозирования вероятных направлений протекания метаболических реакций, основанные на оценке энергии диссоциации C-H связи в молекуле субстрата. Эти методы основаны на предположении, что, подобно любой химической реакции, метаболическая реакция должна протекать в том месте, где энергетический барьер будет наименьшим, а относительная стабильность переходного состояния - наивысшей.

В начале 90-х годов прошлого века была опубликована модель для прогнозирования региоселективности протекания метаболических реакций, катализируемых цитохромами P450 [81]. Метод включал в себя расчет энергий основных энергетических состояний исходной молекулы и радикала, образующегося в результате реакции, по методу AM1. Используя этот метод, разработчики успешно предсказали реакции гидроксилирования, N-деалкилирования и O-деалкилирования для нескольких органических субстратов. Эта модель впоследствии была успешно испытана для прогнозирования скоростей метаболических реакций галогенированных углеводов [82,83], а также для оценки региоселективности метаболизма нитрила *in vivo* и *in vitro* [84]. Более поздние усовершенствования позволили создать комбинированные модели, с успехом предсказывающие сайты ароматического и алифатического гидроксилирования [85]. Указанные разработки легли в основу создания патентованной методологии [86-90].

Одним из важных усовершенствований описанного подхода служит введение дополнительного фактора, влияющего на подверженность C-H связей метаболическим реакциям. Так, в недавно опубликованной работе [91] было показано, что для лекарственных молекул необходимым условием расщепления C-H связи ферментативным путём является определенный уровень стерической доступности атома водорода для микроокружения. В качестве меры стерической доступности был предложен параметр "площадь поверхности атома H, доступная для растворителя". Было выяснено, что метаболическая реакция не может протекать в том случае, если этот параметр оказывается меньше  $8,0 \text{ \AA}^2$ , даже в случае потенциальной неустойчивости данной C-H связи. Неустойчивость определяли при помощи полуэмпирического метода расчета энергии молекулярных орбиталей AM1. Авторы данной работы показали на примере 50 различных лекарственных соединений, что в 78% случаев атомы водорода, обладающие энергией отрыва  $\leq 27$  ккал/моль и площадью поверхности, доступной для растворителя  $\geq 8 \text{ \AA}^2$ , являются главными объектами метаболической атаки. В целом ряде случаев введение указанного стерического фактора позволило уточнить предсказанное направление метаболизма. Например, на рисунке 4 показаны структуры трех молекул, для которых единственное условие минимальной энергии диссоциации C-H связи приводит к некорректному прогнозированию метаболизма цитохромом CYP3A4 (незакрашенная стрелка). Введение стерического фактора позволяет правильно предсказать действительный сайт метаболизма (закрашенная стрелка).

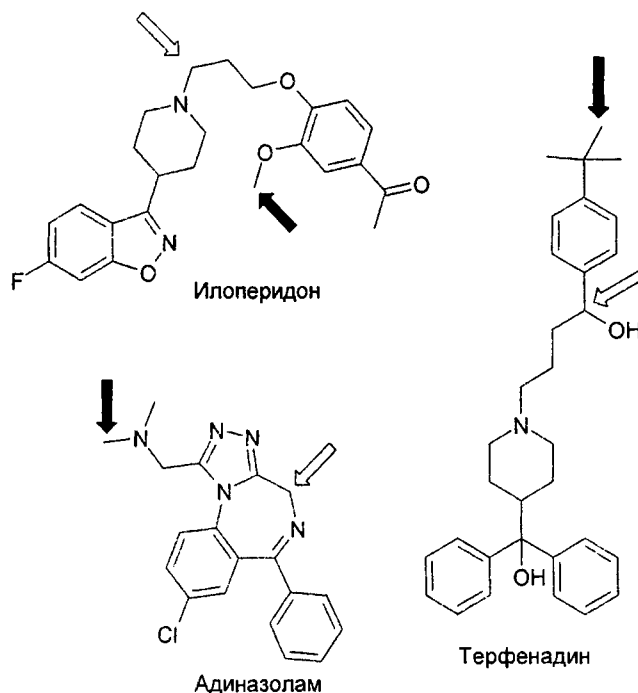


Рисунок 4.

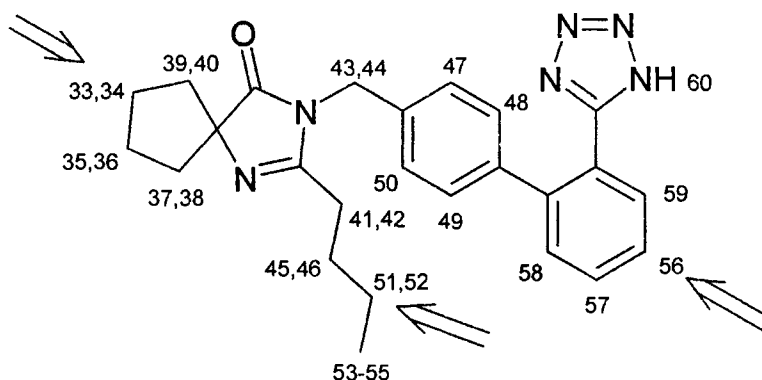
Введение фактора стерической доступности увеличивает точность прогнозирования сайта метаболизма. Незакрашенная стрелка указывает на сайт с минимальной энергией диссоциации C-H связи; закрашенная стрелка указывает на действительный сайт протекания метаболической реакции.

Важным условием успешного применения описанного подхода является возможность свободного вращения молекулы в активном сайте метаболизирующего фермента [92]. На практике подобное ограничение позволяет использовать указанные модели лишь для прогнозирования сайтов метаболизма малых молекул, либо в случае ферментов, обладающих достаточно гибким и пространственно незатрудненным активным центром. К таковым, например, относится, цитохром CYP3A4, ответственный за метаболизм примерно 50% известных лекарственных соединений и обладающий широчайшим спектром субстратной специфичности [93]. Во всех остальных случаях для корректного прогнозирования необходимо учитывать факторы, ориентирующие молекулу субстрата в активном центре фермента [94-98].

**4.3.2. Моделирование взаимодействий "фермент-субстрат".** Опубликовано несколько работ, посвященных компьютерному моделированию специфических пространственных, электростатических и гидрофобных взаимодействий субстратов с активными центрами метаболических ферментов. Поскольку кристаллографическая информация о структурах комплексов субстратов с метаболическими ферментами человека до сих пор весьма ограничена, в настоящее время широко используют построение моделей трехмерной структуры белков по гомологии. В частности, в одной из широко известных работ на указанную тему [99] авторы использовали гомологичную модель одного из важнейших цитохромов (CYP2D6) построенную на основе кристаллических структур цитохромов CYP101, CYP102 и CYP108. Ориентацию трехмерных структур субстратов CYP2D6 в активном сайте модельного белка исследовали при помощи программы SYBYL [100]. Подобную стратегию, основанную на докинге, авторы дополнили независимо построенной фармакофорной моделью, основанной на трехмерных структурах 40 известных субстратов CYP2D6. Авторы данной

работы подчеркивают, что лишь сочетание докинга и фармакофорной модели позволяет делать выводы о пространственном соответствии субстрата и активного окисляющего фрагмента фермента. К сожалению, в описываемой работе авторами не приведены примеры прогнозирования сайтов метаболизма, сделанных при помощи разработанных моделей.

В другой публикации описана перспективная методология прогнозирования сайта метаболизма ферментом CYP2C9 [101]. Она основана на сравнении специальных дескрипторов, рассчитанных для активного центра фермента CYP2C9 и молекул субстрата при помощи технологии GRID MIF (англ. - Molecular Interaction Fields) [102, 103]. Для оценки взаимного расположения реакционного фрагмента фермента в его активном центре и низкомолекулярного субстрата используются специальные индексы комплементарности, учитывающие гидрофобные взаимодействия, а также наличие доноров и акцепторов водородной связи. Эта многоступенчатая запатентованная технология была использована в целом ряде работ за последние несколько лет (см., например, [104, 105]). Она является жизнеспособной альтернативой более традиционному моделированию с использованием докинга, требующей меньших вычислительных ресурсов. В результате работы данного алгоритма для каждой отдельной молекулы генерируется список всех атомов водорода в субстрате, ранжированный в порядке уменьшения вероятности отрыва в ходе метаболической реакции окисления. В описываемой работе было проанализировано 87 метаболических реакций с участием цитохрома CYP2C9. Было показано, что более чем в 90% случаев атомы водорода, находящиеся на первых трех позициях в ранжированном списке, являются действительными сайтами метаболических реакций. Как и в вышеописанном примере, использовали гомологичную белковую модель, основанную на кристаллической структуре фермента CYP2C5 кролика.



ранжированный список атомов H:

53, 56, 57, 54, 52, 55, 34, 51, 39, 35, 33, 36 ...

Рисунок 5.

Направления метаболических реакций гидроксилирования ирбесартана, катализируемых цитохромом CYP2C9, и ранжированный список атомов H, используемый для прогнозирования наиболее вероятных сайтов метаболизма.

Важным достоинством представленной методологии является хорошая точность прогнозирования для молекул, имеющих несколько альтернативных сайтов метаболизма. Например, ирбесартан, синтетический непептидный



антагонист ангиотензина II, представляет собой пример соединения с тремя возможными направлениями метаболического гидроксирования, катализируемого CYP2C9 (рис. 5). В таких случаях, когда анализируемые молекулы являются конформационно гибкими, критически важным фактором является детальное рассмотрение конформаций, которые они могут принимать при взаимодействии с активным сайтом фермента. В описываемой работе [101] для ирбесартана было сгенерировано 25 различных конформаций, для каждой из которых был составлен свой ранжированный список атомов водорода. На рисунке 5 представлен ранжированный список, усредненный по всем 25 конформациям. Видно, что атомы водорода, находящиеся в начале списка, действительно находятся в положениях, для которых наблюдается метаболическая реакция. В ряде случаев на первых позициях ранжированного списка находятся атомы, расположенные в соседних положениях относительно реального сайта метаболизма. Можно предположить, что введение дополнительных критериев, таких как энергетика разрыва связи C-H, способно существенно увеличить предсказательную способность подобных моделей.

К настоящему времени нам удалось найти лишь один пример такого “тотального” подхода к решению подобной задачи. В работе [106] мы встречаем комбинацию докинга субстрата в активном сайте цитохрома CYP2E1 с последующей оценкой энергии разрыва C-H связей вблизи реакционного центра. По аналогии с предыдущими работами, эксперимент основан на модели цитохрома CYP2E1, построенной по гомологии со структурой цитохрома CYP2C5 млекопитающих. Исследовалась всего одна молекула, для которой было обнаружено хорошее соответствие между экспериментальными и теоретически предсказанными направлениями метаболизма. Несмотря на ограниченность изученного материала, данная работа, на наш взгляд, является интересным и ценным примером, демонстрирующим вероятные направления прогресса в исследуемой нами области.

Во всех упомянутых в данном разделе исследованиях ключевым элементом являлась информация о трехмерном строении метаболизирующего фермента, а также его активного сайта. В связи с этим в заключительной части раздела необходимо упомянуть работы отечественных исследователей, работающих в области построения гомологичных пространственных моделей цитохромов P450 [107-109].

#### *4.4. Моделирование кинетики взаимодействия веществ с активными центрами метаболических ферментов.*

Вопросам прогнозирования кинетических параметров взаимодействия с метаболическими ферментами, преимущественно цитохромами P450, посвящено значительное число исследований. Стандартными подходами к решению подобных задач служат нахождение количественных взаимосвязей “структура-активность” (КССА), либо построение фармакофорных моделей субстратов/ингибиторов для соответствующих ферментов.

История КССА-моделирования в сфере предсказательного метаболизма ведет свой отсчет от момента опубликования классических работ Hansch по моделированию кинетических параметров взаимодействия цитохромов P450 с субстратами [110, 111]. В дальнейшем эта тема получила широкое развитие. Так, в работах [112, 113] были исследованы и описаны факторы, влияющие на аффинность связывания субстратов с цитохромами P450. В частности, в работе [113] были построены КССА модели (табл. 6), связывающие величину  $\Delta G_{\text{bind}}$  (свободная энергия связывания) с расчетными параметрами молекул, для нескольких цитохромов, играющих наиболее важную роль в метаболизме ксенобиотиков в организме человека. На основе построенных моделей был сделан практически важный вывод о центральной роли водородных связей в обеспечении прочного связывания субстрата и фермента. Следует, однако, отметить, что недостаточная представительность использованных в этой работе обучающих выборок, к сожалению, не позволяет делать серьезные обобщения.

Таблица 6. КССА модели энергии связывания соединений с активными сайтами цитохромов P450 (по [113] с изменениями).

Цитохром P450	Уравнение	N	R
CYP3A4	$\Delta G_{\text{bind}} = 1,92(\pm 0,45)E_L - 2,99(\pm 0,61)E_H -$ $0,52(\pm 0,14)N_{\text{HB}} + 3,55(\pm 0,71)N_{\pi-\pi} - 35,10$	10	0,96
CYP2E1	$\Delta G_{\text{bind}} = 0,56(\pm 0,09)N_{\text{HB}} - 2,18(\pm 0,17)N_{\pi-\pi} -$ $3,37(\pm 0,37)\log M_r - 0,36(\pm 0,12)N_{\text{HB}} + 1,86$	10	0,99
CYP2D6	$\Delta G_{\text{bind}} = 492,03(\pm 100,04)\log M_r - 0,88(\pm 0,18)M_r -$ $5,08(\pm 0,90)N_{\text{HB}} - 3,76(\pm 0,82)N_{\pi-\pi} - 947,68$	10	0,94
CYP2C19	$\Delta G_{\text{bind}} = 201,69(\pm 19,87)\log M_r - 0,31(\pm 0,03)M_r -$ $2,50(\pm 0,23)\log P + 0,12(\pm 0,11)N_{\text{BD}} - 405,68$	8	0,99
CYP2C9	$\Delta G_{\text{bind}} = 8,62(\pm 0,82)\log D_{7,4} - 8,02(\pm 0,79)\log P -$ $6,26(\pm 0,62)pK_a + 0,57(\pm 0,19)N_{\text{BD}} + 42,74$	8	0,99
CYP2B6	$\Delta G_{\text{bind}} = 3,99(\pm 0,76)N_{\text{HB}} - 5,41(\pm 0,86)N_{\text{BD}} -$ $1,92(\pm 0,23)N_{\text{BA}} - 4,19$	10	0,97
CYP2A6	$\Delta G_{\text{bind}} = 0,77(\pm 0,19)N_{\text{HB}} - 0,78(\pm 0,27)\log P - 4,45$	6	0,97
CYP1A2	$\Delta G_{\text{bind}} = 0,30(\pm 0,12)\mu - 0,90(\pm 0,36)l/w + 2,11(\pm 0,26)\Delta E -$ $0,50(\pm 0,17)N_{\text{HB}} - 22,41$	11	0,97

Примечание: N – число соединений в обучающей выборке; R – коэффициент корреляции.  $\Delta G_{\text{bind}}$  – свободная энергия связывания;  $E_L$  – энергия низшей незанятой молекулярной орбитали;  $E_H$  – энергия высшей занятой молекулярной орбитали;  $\Delta E = E_L - E_H$ ;  $N_{\text{HB}}$  – число водородных связей между молекулой субстрата и активным сайтом фермента;  $N_{\pi-\pi}$  – число  $\pi$ - $\pi$  стэкинг-взаимодействий между молекулой субстрата и активным сайтом фермента;  $M_r$  – относительная молекулярная масса;  $N_{\text{BA}}$  – число акцепторов водородной связи;  $N_{\text{BD}}$  – число доноров водородной связи;  $\mu$  – дипольный момент;  $l/w$  – отношение длины молекулы к ее ширине;  $\log P$  – логарифм коэффициента распределения в системе 1-октанол/вода.

Недавно были описаны статистически значимые нейронно-сетевые КССА-модели, описывающие скорость реакции *N*-деалкилирования, катализируемой цитохромами CYP3A4 и CYP2D6 (31 и 36 соединений соответственно) [114]. Авторы исследовали три различных типа структур в исследуемых молекулах: (1) полная молекулярная структура, (2) структура топологического центроида реакционного центра с радиусом, равным трем молекулярным связям от атома азота, (3) структура уходящего фрагмента. Было обнаружено, что все три указанных уровня молекулярной организации влияют на скорость реакции, причем их относительная значимость убывает в следующем порядке: полная структура, центроид, уходящий фрагмент. Важнейшими факторами, влияющими на скорость реакции *N*-деалкилирования, являются плотность центроида как мера стерической доступности сайта реакции, полярные свойства поверхности молекул, а также геометрические параметры уходящего фрагмента.

Большинство работ, посвященных прогнозированию кинетических

параметров взаимодействия низкомолекулярных соединений с метаболическими ферментами, основано на построении трехмерных фармакофорных моделей. Мы можем упомянуть здесь лишь некоторые работы из обширного ряда исследований, посвященных разработке фармакофорных моделей субстратов и ингибиторов цитохромов P450 [96, 115-130]. Специфические характеристики некоторых опубликованных фармакофорных моделей приведены в таблице 7. Как правило, построенные фармакофорные модели в дальнейшем используются для построения трехмерных КССА моделей, описывающих кинетические параметры связывания, такие как константа Михаэлиса ( $K_m$ ) или константа ингибирования ( $K_i$ ). Рамки настоящей статьи не позволяют более подробно останавливаться на деталях данных экспериментальных работ, которым уделено значительное внимание в целом ряде свежих обзорных публикаций [93, 97, 131-133].

Таблица 7 Особенности фармакофорных моделей субстратов и ингибиторов некоторых цитохромов P450

Цитохром P450	Специфические параметры фармакофора	Ссылка
1A2	<i>ингибиторами</i> являются плоские молекулы с малой площадью поверхности	[120]
2B6	<i>субстраты</i> содержат как минимум три гидрофобных области и один акцептор водородной связи	[121]
2C9	<i>субстраты</i> : (модель 1) донор водородной связи на расстоянии 7 Å от сайта метаболизма; (модель 2) расстояние между сайтом гидроксирования и анионом равно 7,8 Å; гидрофобная зона между сайтом гидроксирования субстрата и катионным центром фермента;  <i>ингибиторы</i> содержат как минимум одну гидрофобную зону и один акцептор водородной связи на расстоянии от 3 до 5,8 Å	[122] [123,124]  [125]
2D6	<i>субстраты</i> содержат основной атом азота на расстоянии 5-7 Å от места окисления, а также копланарные ароматические кольца;  <i>ингибиторы</i> включают донор и акцептор водородной связи, а также две или три гидрофобных зоны	[126]  [127]
3A4	<i>субстраты</i> : (модель 1) акцептор водородной связи на расстоянии 5,5-7,8 Å от места метаболизма; (модель 2) два акцептора и один донор водородной связи, одна гидрофобная зона;  <i>ингибиторы</i> (модель 1) три гидрофобных зоны на расстоянии 5,2-8,8 Å от акцептора водородной связи; (модель 2) одна гидрофобная зона на расстоянии 8,1-16,3 Å от двух (из числа трех) самых удаленных акцепторов водородной связи	[128] [129]  [130]

В последние годы для построения фармакофорных моделей субстратов и ингибиторов метаболических ферментов активно используют информацию не только о низкомолекулярных соединениях, но также о пространственном строении ферментов. Мы уже описывали ряд подобных работ в разделе о прогнозировании сайтов метаболизма [114, 134]. Аналогичные подходы описаны для прогнозирования кинетических параметров взаимодействия структурно разнообразных субстратов и ингибиторов с активными центрами цитохромов. Так, уже упоминавшаяся методология GRID была использована в серии недавних работ интернациональной группы исследователей [134-136]. Результатом каждой из этих работ являлась КССА модель ингибирования цитохрома CYP2C9, основанная на оценке особенностей стерических, водородных и гидрофобных взаимодействий соединений с активным центром фермента.

Моделированию взаимодействий низкомолекулярных субстратов и ингибиторов с метаболическими ферментами, отличными от цитохромов P450, посвящено сравнительно малое число публикаций [137-142]. В работе [140] была построена КССА модель, описывающая скорость гидролиза 80 соединений, содержащих сложноэфирную группу, карбоксилэстеразами. Ключевым фактором, ответственным за скорость ферментативного гидролиза, оказалась стерическая доступность гидролизуемой связи. В числе других молекулярных свойств, влияющих на метаболическую стабильность, оказались липофильность (оцененная как  $\log P$ ) и заряд на карбонильном атоме углерода.

Следует отметить публикацию, в которой была построена трехмерная фармакофорная модель субстратов UDP-зависимой глюкуронозилтрансферазы 1A4 (UGT1A4) [142]. Обучающая выборка включала в себя 24 структурно разнообразных субстрата UGT1A4. В результате построения фармакофорной модели при помощи программного комплекса Catalyst [143] авторам удалось построить эффективную КССА-модель, имевшую более высокую предсказывающую способность, чем модель, основанная на двумерных дескрипторах.

Недавно была описана модель, основанная на методологии самоорганизующихся карт Кохонена, которая позволяет эффективно классифицировать низкомолекулярные органические соединения по их способности связываться с активными центрами важнейших метаболических ферментов [144]. В качестве обучающей выборки использовалась база, состоящая из более 500 соединений, импортированных из коммерческой базы MetaDrug [63]. Первоначально было рассчитано несколько десятков молекулярных дескрипторов, из которых были отобраны шесть параметров, кодирующих липофильность, топологические параметры и способность молекулы образовывать водородные связи. В результате авторам удалось построить несколько карт Кохонена, соответствующих разным цитохромам P450, на которой сильно и слабо связывающиеся соединения оказывались расположенными в разных зонах (рис. 6; в качестве примера показана карта цитохрома CYP3A4). Использование подобных карт в режиме тестирования позволяет эффективно и вполне надежно классифицировать соединения по их способности связываться с метаболическими ферментами. Подобный подход может быть рекомендован для высокопроизводительной оценки больших виртуальных баз соединений, не требующей использования значительных вычислительных ресурсов.

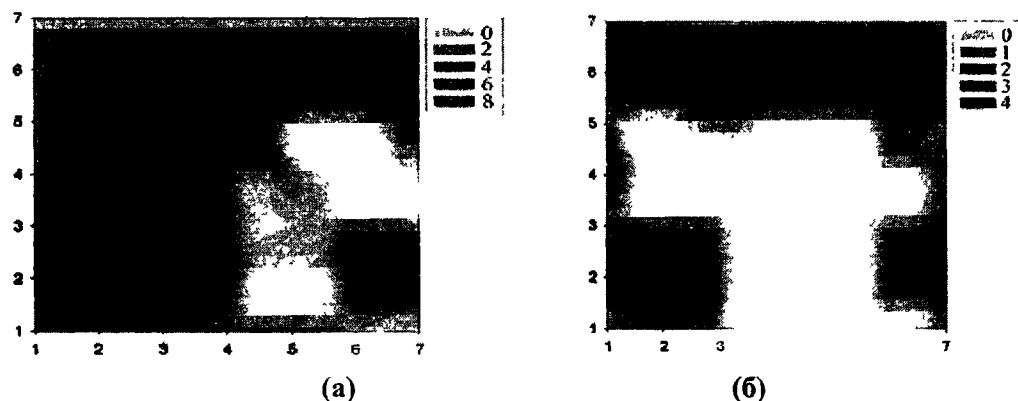


Рисунок 6.

Зоны расположения на карте Кохонена соединений, эффективно связывающихся (а,  $K_m < 10$ ), а также слабо связывающихся (б,  $K_m > 100$ ) с активным сайтом СУРЗА4. Цифры по осям абсцисс и ординат являются безразмерными координатами ячеек карты.

#### 4.5 Исследования ферментной специфичности субстратов метаболических реакций

Анализ субстратов, связанный с прогнозированием типа ферментативной системы, вовлеченной в их метаболическую трансформацию, является еще одной важной стороной исследований метаболических эффектов. Получение информации о наиболее вероятном типе метаболизирующего фермента для конкретного соединения позволяет далее с наибольшей эффективностью использовать различные модели, аналогичные рассмотренным выше, так как практически все эти модели привязаны к определенным типам метаболических ферментов. Кроме того, подобная информация может быть использована для анализа и оценки ряда актуальных фармакокинетических и токсикокинетических свойств соединений, таких как метаболический клиренс, профиль распределения, нежелательные взаимодействия между лекарственными препаратами (англ. Drug-Drug Interactions) и др.

Несмотря на очевидную актуальность, в настоящий момент эта тема, на наш взгляд, пока не получила должного развития. В ряде работ мы указываем на возможности фермент-специфичной дифференциации субстратов [78], а также ингибиторов [144] ферментов семейства цитохромов P450. Эти работы связаны с построением классификационных КССА-моделей, основанных на статистических алгоритмах снижения размерности и визуализации многомерных пространств свойств молекул.

Наиболее законченное выражение к настоящему времени эта тема получила в работе российских исследователей [145]. Используя собственный оригинальный аналитический модуль PASS-BioTransfo, авторы проанализировали профили реакционной и фермент-специфичной активности субстратов, содержащихся в базах Metabolite и Metabolism (см. табл. 4). Методология работы системы PASS (англ. Prediction of Activity Spectra for Substances), описанная в серии публикаций [146-148], основана на КССА-анализе соединений из обучающей выборки, кодированных специальными дескрипторами “многоуровневого атомного окружения” (или MNA, от англ. Multilevel Neighborhoods of Atoms). В результате обучения с использованием указанных баз, программа оказалась способной корректно предсказывать соответственно 88% и 85% классов метаболических трансформаций. Для многих важных трансформаций, таких как реакции ароматизации, деароматизации, эпоксицирования, *N*-деалкилирования (цитохромы P450), деалогенирование (глутатион-трансфераза), гидролиз (эстераза), *N*-ацетилирование (*N*-ацетилтрансфераза) и других, количество корректно классифицированных соединений превышает 90%. В качестве примеров таблица 8 содержит количественные сведения о корректно предсказанных классах некоторых

важных метаболических трансформаций, содержащихся в базе Metabolite.

Таблица 8. Примеры прогнозирования классов биотрансформаций программой PASS-BioTransfo при использовании в качестве обучающей выборки базы Metabolite.

Класс биотрансформации	Число соединений в базе, относящихся к данному классу	Точность прогнозирования (%), оцененная в результате процедуры скользящего контроля с исключением по одному
Алифатическое гидроксилирование (цитохром P450)	443	81,35
Алифатическое гидроксилирование (цитохром CYP3A4)	130	80,73
Ароматическое гидроксилирование (цитохром P450)	403	85,54
Ароматическое гидроксилирование (цитохром CYP3A4)	78	84,95
Ароматизация (цитохром P450)	102	96,10
Глюкуронидация (UDP- глюкуронозилтрансфераза)	276	85,88
Гидролиз (эстераза)	231	92,51
Дезаминирование (цитохром P450)	46	86,75
Деароматизация (цитохром P450)	161	90,37
Дегалогенирование (глутатион- трансфераза)	65	97,68
Дегидратация (цитохром P450)	43	87,61
Дегидрирование (цитохром P450)	116	85,98
Присоединение глутатиона (глутатион- трансфераза)	220	89,83
Эпоксидирование (цитохром P450)	185	90,52
Присоединение глутатиона (глутатион- трансфераза)	220	89,83
Гидролиз (эстераза)	231	92,51
N-ацетилирование (N- ацетилтрансфераза)	88	96,26
N-деалкилирование (цитохром P450)	370	92,40
N-окисление (цитохром P450)	132	88,71
S-окисление (цитохром P450)	92	94,85

Логическим развитием рассмотренной методологии явилась разработка теми же исследователями нового подхода к статистической оценке наиболее вероятных сайтов протекания метаболических превращений [68]. Подход, реализованный в рамках программной среды PASS, основан на статистическом анализе метаболических трансформаций определенного типа, содержащихся в экспериментальных базах. Все возможные трансформации, а также ближайшее структурное окружение сайта реакции кодировались посредством специальных дескрипторов “реакционного многоуровневого атомного окружения” (или RMNA, от англ. Reacting Multilevel Neighborhoods of Atoms), являющихся близкими аналогами MNA-дескрипторов. Вероятность каждой реакции оценивалась при помощи статистического уравнения, учитывающего вероятность протекания подобной трансформации, кодированной RMNA-дескрипторами, в обучающей выборке. Разработанный алгоритм был применен для прогнозирования сайтов ароматического гидроксирования для базы структурно неоднородных низкомолекулярных субстратов. Для обучения использовалась база данных Metabolism (Accelrys), а в качестве тестовой выборки были использованы 1552 ароматических соединения из БД Metabolite (MDL Information Systems). Средняя точность прогнозирования сайтов ароматического гидроксирования для тестовой выборки превысила 84%. Сравнение данного алгоритма с двумя альтернативными алгоритмами, основанными на оценке энергетики реакции ароматического гидроксирования [81, 83, 85], продемонстрировало сопоставимую точность результатов прогноза при существенно меньших вычислительных затратах метода [68].

Следует отметить некоторые ограничения подхода, реализованного в программном комплексе PASS. Как и большинство других систем, использующих информацию только о структурных фрагментах, эта программа не учитывает особенности физико-химического профиля анализируемых соединений, играющие важную роль в оценке метаболических эффектов, а также трехмерное пространственное строение молекул. Однако, как отмечалось выше, данный подход может быть рекомендован в качестве эффективного средства предварительного анализа данных, а полученная в результате информация может быть использована для более тонкой оценки пространственных, кинетических и энергетических аспектов специфических субстрат-ферментных взаимодействий.

#### *4.6. Хемогеномика и прогнозирование метаболизма.*

Наш обзор был бы неполным без упоминания интенсивно разрабатываемых в последние годы специализированных баз данных, связывающих геномную информацию со строением метаболических ферментов, а также их низкомолекулярных субстратов и модуляторов. Подобная информация о связи геномов и низкомолекулярных физиологически активных веществ играет все возрастающую роль в научных исследованиях в области биомедицинской химии. В частности, на ее основе интенсивно развивается новая дисциплина *хемогеномика* [149], позволяющая интенсифицировать процессы поиска новых лекарственных субстанций и валидации новых терапевтически значимых белковых мишеней. В настоящее время разработано несколько подобных баз знаний, относящихся к метаболическим системам организма [150-152]. Как правило, они создаются в университетских исследовательских центрах и включают в себя сведения о структурах генов, кодирующих метаболические ферменты, информацию о метаболических реакциях, важнейших субстратах, ингибиторах и индукторах. Информационная система по цитохромам P450, разработанная российскими исследователями [150], дополнительно содержит ссылки на литературные источники, а также обладает гибкими возможностями поиска, просмотра и отбора информации.

В настоящее время подобные базы данных являются незаменимым инструментом исследований, связанных с реконструкцией трехмерных структур метаболических ферментов. Можно предположить, что по мере накопления

данных и совершенствования наших знаний о законах функционирования геномов и их взаимосвязях с низкомолекулярными соединениями подобные информационные системы будут играть важнейшую роль в области предсказательного метаболизма. Вероятно, их дальнейшее развитие пойдет по пути интегрирования с аналитическими алгоритмами прогнозирования метаболизма, аналогичными описанным в настоящем обзоре.

**ЗАКЛЮЧЕНИЕ.** Анализ текущего состояния дел в сфере компьютерных методов прогнозирования метаболизма низкомолекулярных органических соединений позволяет нам нарисовать вероятную картину развития в этой области на ближайшие годы. Очевидным направлением работ останется дальнейшее развитие и усовершенствование информационных баз данных метаболических превращений. Количественное и качественное усовершенствование подобных коммерчески доступных БД будет неизбежно сопровождаться внедрением в них интеллектуальных алгоритмов прогнозирования сайтов метаболических превращений, метаболической стабильности соединений, а также оценки фермент-специфического профиля веществ.

Накопление и систематизация информации из различных экспериментальных источников приведет к созданию больших (тысячи или даже десятки тысяч структур веществ) общедоступных баз данных, содержащих сведения о метаболической стабильности соединений по отношению к различным метаболическим системам. Параллельно будет происходить совершенствование методов анализа данных, а соответствующие статистические алгоритмы будут рассматриваться в качестве стандартных инструментов при планировании схем синтеза соединений. Это приведет к тому, что прогнозирование метаболической стабильности станет рутинной операцией в программах виртуального скрининга на ранних этапах поиска и разработки лекарственных субстанций.

Увеличение вычислительных возможностей, а также получение кристаллографических данных для большинства наиболее важных метаболических ферментов позволит построить достаточно надежные модели для оценки ориентации низкомолекулярных субстратов в активных центрах ферментов. В сочетании с предсказанными данными об энергии разрыва и стерической доступности молекулярных связей, потенциально подверженных метаболической деградации, подобные пространственные модели позволят с высокой степенью надежности предсказывать сайты метаболизма и структуры продуктов. Наиболее вероятно, что подобные системы будут широко использоваться на поздних этапах разработки лекарственных субстанций для оптимизации фармакокинетических и токсикокинетических свойств соединений-кандидатов до проведения доклинических испытаний.

Создание качественных трехмерных моделей взаимодействия субстрат-фермент будет основано на информации, полученной из геномных баз данных по метаболическим системам организма. В будущем можно предвидеть создание полностью интегрированных систем, в которых сведения о геномах, трехмерных структурах кодируемых ими ферментов, метаболических реакциях, а также интеллектуальные алгоритмы анализа данных будут объединены в мощные информационно-вычислительные комплексы, способные эффективно прогнозировать метаболизм ксенобиотиков.

Вероятно, даже реализация всех этих ожиданий, пока еще кажущихся чрезмерно оптимистическими, не позволит нам достигнуть “вершины Эвереста”. Однако, они, безусловно, откроют исследователям новые горизонты и еще более широкие перспективы в этой чрезвычайно интересной области. Области науки, насыщенной конкуренцией идей, многообразием подходов и яркими достижениями человеческой интуиции и передовых технологий.



## ЛИТЕРАТУРА.

1. Erhardt P.W. (1999) in: *Drug metabolism: Drug metabolism data: past, present and future considerations* (P.W. Erhardt ed.) Blackwell Science, pp. 17-22.
2. Dragovich P.S., Prins T.J., Zhou R., Johnson T.O., Hua Y., Luu H.T., Sakata S.K., Brown E.L., Maldonado F.C., Tuntland T., Lee C.A., Fuhrman S.A., Zalman L.S., Patick A.K., Matthews D.A., Wu E.Y., Guo M., Borer B.C., Nayyar N.K., Moran T., Chen L., Rejto P.A., Rose P.W., Guzman M.C., Dovalsantos E.Z., Lee S., McGee K., Mohajeri M., Liese A., Tao J., Kosa M.B., Liu B., Batugo M.R., Gleeson J.P., Wu Z.P., Liu J., Meador J.W. 3<sup>rd</sup>, Ferre R.A. (2003) *J. Med. Chem.*, **46**, 4572-4585.
3. Tagat J.R., Steensma R.W., McCombie S.W., Nazareno D.V., Lin S.I., Neustadt B.R., Cox K., Xu S., Wojcik L., Murray M.G., Vantuno N., Baroudy B.M., Strizki J.M. (2001) *J. Med. Chem.*, **44**, 3343-3346.
4. Palani A., Shapiro S., Josien H., Bara T., Clader J.W., Greenlee W.J., Cox K., Strizki J.M., Baroudy B.M. (2002) *J. Med. Chem.*, **45**, 3143-3160.
5. Humphreys W.G., Obermeier M.T., Barrish J.C., Chong S., Marino A.M., Murugesan N., Wang-Iverson D., Morrison R.A. (2003) *Xenobiotica*, **33**, 1109-1123.
6. Burgey C.S., Robinson K.A., Lyle T.A., Sanderson P.E.J., Lewis D.S., Lucas B.J., Krueger J.A., Singh R., Miller-Stein C., White R.B., Wong B. (2003) *J. Med. Chem.*, **46**, 461-473.
7. Clader J.W. (2004) *J. Med. Chem.*, **47**, 1-9.
8. Bell I.M., Gallicchio S.N., Abrams M., Beese L.S., Beshore D.C., Bhimnathwala H., Bogusky M.J., Buser C.A., Culberson J.C., Davide J., Ellis-Hutchings M. (2002) *J. Med. Chem.*, **45**, 2388-2409.
9. Higuchi T., Stella V. (1975) *Pro-drugs as novel drug delivery systems*, Am. Chem. Soc., Washington.
10. Stella V.J., Himmelstein K.J. (1980) *J. Med. Chem.*, **23**, 1275-1282.
11. Deutsch H.M., Glinski J.A., Hernandez M., Haugwitz R.D., Narayanan V.L., Suffness M., Zalkow L.H. (1989) *J. Med. Chem.* **32**, 788-792.
12. Erhardt P.W. (1993) *Chronicles Drug Discov.*, **3**, 191-206.
13. Testa B., Jenner P. (1976) *Drug Metabolism-Chemical and Biochemical Aspects*, Marcel Dekker Inc., New York.
14. Ioannides C. (1996) *Cytochromes P450 Metabolic and Toxicological Aspects*, CRC Press Inc., New York.
15. Lewis D.F.V. (1996) *Cytochromes P450*, Taylor & Francis Inc., Bristol.
16. Cashman J.R. (2000) *Curr. Drug Metab.*, **1**(2), 181-191.
17. Benedetti M.S., Tipton K.F. (1998) *J. Neural Transm. Suppl.*, **52**, 149-171.
18. Yin S.-J., Han C.-Li., Lee A.-I., Wu C.-W. (1999) *Adv. Exp. Med. Biol.*, **463**, 265-274.
19. Vasiliou V., Pappa A., Petersen D.R. (2000) *Chem.-Biol. Interact.*, **129**(1-2), 1-19.
20. Kitamura S., Nakatani K., Ohashi K., Sugihara K., Hosokawa R., Akagawa Y., Ohta S. (2001) *Biol. Pharm. Bull.*, **24**(7), 856-859.
21. Pritsos C.A. (2000) *Chem.-Biol. Interact.*, **129**(1-2), 195-208.
22. Armstrong R.N. (1997) *Chem. Res. Toxicol.*, **10**, 2-18.
23. Vereczkey L., Jemnitz K., Gregus Z. (1998) *Acta. Pharm. Hung.*, **68**(5), 284-288.
24. Negishi M., Pedersen L.G., Petrotchenko E., Shevtsov S., Gorokhov A., Kakuta Y., Pedersen L.C. (2001) *Arch. Biochem. Biophys.*, **390**(2), 149-157.
25. Ilett K.F., Kadlubar F.F., Minchin R.F. (1999) *Drug Metab. Dispos.*, **27**(9), 957-959.
26. Burchell B., Brierly C.H., Rance D. (1995) *Life Sci.*, **57**, 1819-1831.
27. Weinshilboum R. (1989) *Pharmacol. Ther.*, **43**, 77-90.
28. Ferrero J.L., Brendel K. (1997) *Adv. Pharmacol.*, **43**, 131-169.
29. Fisher R.L., Ulreich J.B., Nakazato P.Z., Brendel K. (2001) *Toxicol. Methods*, **11**(2), 59-79.
30. Vanhulle V.P., Martiat G.A., Verbeeck R.K., Horsmans Y., Calderon P.B., Eeckhoudt S.L., Taper H.S., Delzenne N. (2001) *Life Sci.*, **68**(21), 2391-2403.
31. Ekins S., Ring B.J., Grace J., McRobie-Belle D.J., Wrighton S.A. (2000) *J. Pharm.*

- Toxicol. Methods, **44**(1), 313-324.
32. Cross D.M., Bayliss M.K. (2000) Drug Metab. Rev., **32**(2), 219-240.
33. Guillouzo A., Rialland L., Fautrel A., Guyomard C. (1999) Chem.-Biol. Interact., **121**(1), 7-16.
34. Li A.P. (1999) Chem.-Biol. Interact., **121**(1), 1-5.
35. Umehara K., Kudo S., Hirao Y., Morita S., Uchida M., Odomi M., Miyamoto G. (2000) Drug Metab. Dispos., **28**(8), 887-894.
36. Crespi C.L. (1999) Curr. Opin. Drug Discov. Dev., **2**(1), 15-19.
37. Caldwell G.W., Masucci J.A., Chacon E. (1999) Comb. Chem. High Throughp. Scr., **2**(1), 39-51.
38. Watt A.P., Morrison D., Evans D.C. (2000) Drug Discovery Today, **5**(1), 17-24.
39. White R.E. (2000) Annu. Rev. Pharm. Toxicol., **40**, 133-157.
40. Cox K.A., Clarke N.J., Rindgen D., Korfmacher W.A. (2001) Am. Pharm. Rev., **4**(1), 45-46.
41. Zhou L., Thakker D.R., Voyksner R.D., Anbazhagan M., Boykin D.W., Hall J.E., Tidwell R.R. (2004) J. Mass Spectrom., **39**(4), 351-360.
42. Yu X., Cui D., Davis M.R. (1999) J. Am. Soc. Mass Spectrom., **10**(2), 175-183.
43. Lopez L.L., Yu X., Cui D., Davis M.R. (1998) Rapid Commun. Mass Spectrom., **12**(22), 1756-1760.
44. Fernandez-Metzler C.L., Owens K.G., Baillie T.A., King R.C. (1999) Drug Metab. Dispos., **27**(1), 32-40.
45. Owens K.G. (1991) Appl. Spectrosc. Rev., **27**(1), 1-49.
46. Van Breemen R.B., Nikolic D., Bolton J.L. (1998) Drug Metab. Dispos., **26**(2), 85-90.
47. Nikolic D., Fan P.W., Bolton J.L., Van Breemen R.B. (1999) Comb. Chem. High Throughp., Scr. **2**(3), 165-175.
48. Nikolic D., Habibi-Goudarzi S., Corley D.G., Gafner S., Pezzuto J.M., Van Breemen R.B. (2000) Anal. Chem., **72**(16), 3853-3859.
49. Wilson I.D., Nicholson J.K., Lindon J.C. (1999) in: *Handbook of Drug Metabolism: The role of nuclear magnetic resonance spectroscopy in drug metabolism* (T.F. Woolf ed.) Marcel Dekker Inc., New York, pp. 523-550.
50. Dear G.J., Plumb R.S., Sweatman B.C., Parry P.S., Roberts A.D., Lindon J.C., Nicholson J.K., Ismail I.M. (2000) J. Chromatogr.; B: Biomed. Sci. Appl., **748**(1), 295-309.
51. Dear G.J., Plumb R.S., Sweatman B.C., Ayrton J., Lindon J.C., Nicholson J.K., Ismail I.M. (2000) J. Chromatogr.; B: Biomed. Sci. Appl., **748**(1), 281-293.
52. Lindon J.C., Nicholson J.K., Wilson I.D. (2000) J. Chromatogr.; B: Biomed. Sci. Appl., **748**(1), 233-258.
53. Erhardt P.W. (2002) in: *High-throughput ADMETox estimation* (F. Darvas and G. Dorman eds.) BioTechniques Press, pp. 49-68.
54. MDL Information Systems, Inc. <http://www.mdl.com>.
55. Erhardt P.W. (2002) in: *High-throughput ADMETox estimation: Metabolism Prediction* (F. Darvas and G. Dorman eds.) BioTechniques Press, Eaton Publishing, pp. 41-48.
56. Accelrys. Inc.; <http://www.accelrys.com>.
57. Darvas F. (1987) in: *QSAR Environmental Toxicology II: MetabolExpert: an expert system for predicting the metabolism of substances* (K.L. Kaiser ed.) Reidel Co., Dordrecht, pp. 71-81.
58. Darvas F.J. (1988) Mol. Graphics, **6**, 80-86.
59. Klopman G., Dimayuga M., Talafous J. (1994) J. Chem. Inf. Comput. Sci., **34**, 1320-1325.
60. Talafous J., Sayre L.M., Mieyal J.J., Klopman G. (1994) J. Chem. Inf. Comput. Sci., **34**, 1326-1333.
61. Klopman G., Tu M., Talafous J. (1997) J. Chem. Inf. Comput. Sci., **37**, 329-334.
62. Greene N. (1999) *Drug Metabolism* (P.W. Erhardt ed.) Blackwell Science, pp. 289-296.
63. MetaDrug database, GeneGo, Inc.; <http://www.genego.com>.
64. Klopman G., Tu M. (1999) in: *Drug Metabolism: META: a program for the prediction of the product of mammal metabolism of xenobiotics* (P.W. Erhardt ed.)

- Blackwell Science, pp. 271-276.
65. Klopman G., Li J.-Y., Wang S., Dimayuga, M. (1994) *J. Chem. Inf. Comp. Sci.*, **34**, 752-781.
  66. Smith D.A., van de Waterbeemd H., Walker D.K. (2001) in: *Pharmacokinetics and Metabolism in Drug Design: Renal Clearance* (D.A. Smith, H. van de Waterbeemd and D.K. Walker eds.) Wiley-VCH, Weinheim, pp. 67-74.
  67. Boyer S., Zamora I. (2002) *J. Comp.-Aid. Mol. Des.*, **16**(5-6), 403-413.
  68. Borodina Yu., Rudik A., Filimonov D., Kharchevnikova N., Dmitriev A., Blinova V., Poroikov V. (2004) *J. Chem. Inform. Comput. Sci.*, **44** (6), 1998-2009.
  69. Button W.G., Judson P.N., Long A., Vessey J.D. (2003) *J. Chem. Inf. Comp. Sci.*, **43**(5), 1371-1377.
  70. Gombar V.K., Silver I.S., Zhao Z. (2003) *Curr. Top. Med. Chem.*, **3**(11), 1205-1225.
  71. Musimirembwa C.M., Bredberg U., Andersson T.B. (2003) *Clin. Pharmacokinet.*, **42**(6), 515-528.
  72. Ekins S. (2003) *Biochem. Soc. Trans.*, **31**, 611-614.
  73. Cruciani G., Meniconi M., Carosati E., Zamora I., Mannhold R. (2003) in: *Drug Bioavailability: VOLSURF: a tool for drug ADME-properties prediction* (H. van de Waterbeemd, H. Lennemäs and P. Artursson eds.) Wiley-VCH, pp. 406-419.
  74. Jolliffe I.T. (2002) *Principal Component Analysis, 2nd edition*, Springer-Verlag, New York.
  75. Shen M., Xiao Y., Golbraikh A., Gombar V.K., Tropsha A. (2003) *J. Med. Chem.*, **46**, 3013-3020.
  76. Quincy M.A. (1998) MolConnZ, version 3.5; Hall Associates Consulting.
  77. Carhart R.E., Smith D.H., Venkataraghavan R. (1985) *J. Chem. Inf. Comput. Sci.*, **25**, 64-73.
  78. Korolev D., Balakin K.V., Nikolsky Y., Kirillov E., Ivanenkov Y.A., Savchuk N.P., Ivashchenko A.A., Nikolskaya T. (2003) *J. Med. Chem.*, **46**, 3631-3643.
  79. Kohonen T. (1989) *Self-Organization and Associative Memory, 3rd edition*, Springer-Verlag, Berlin.
  80. Bugrim A., Nikolskaya T., Nikolsky Y. (2004) *Drug Discov. Today*, **9**(3), 127-135.
  81. Korzekwa K.R., Jones J.P., Gillette J.R. (1990) *J. Am. Chem. Soc.*, **112**, 7042-7046.
  82. Harris J.W., Jones J.P., Martin J.L., LaRosa A.C., Olson M.J., Pohl L.R., Anders M.W. (1992) *Chem. Res. Toxicol.*, **5**, 720-725.
  83. Yin H., Anders M.W., Korzekwa K.R., Higgins L., Thummel K.E., Kharasch E.D., Jones J.P. (1995) *Proc. Nat. Acad. Sci.*, **92**, 11076-11080.
  84. Grogan J., DeVito S.C., Pearlman R.S., Korzekwa K.R. (1992) *Chem. Res. Toxicol.*, **5**, 548-552.
  85. Jones J.P., Mysinger M., Korzekwa K.R. (2002) *Drug Metab. Dispos.*, **30**, 7-12.
  86. Jones J.P., Korzekwa K.R. (1996) *Methods Enzymol.*, **272**, 326-335.
  87. Ortiz de Montellano P.R. (1999) in: *Handbook of Drug Metabolism: The Cytochrome P450 Oxidative System* (T.F. Woolf ed.) Dekker, New York, Chapter 4, pp. 109-130.
  88. Omura T. (1999) *Biochem. Biophys. Res. Commun.*, **266**, 690-698.
  89. Korzekwa K.R., Jones J.P., Higgins L.A. (2003) US Pat. 6,643,591. Der. Abstr., **C2000**, 224032.
  90. Korzekwa K.R., Jones J.P. (2003) US Pat. 6,625,547. Der. Abstr., **C2000**, 224032.
  91. Singh S.B., Shen L.Q., Walker M.J., Sheridan R.P. (2003) *J. Med. Chem.*, **46**, 1330-1336.
  92. Higgins L., Korzekwa K.R., Rao S., Shou M., Jones J.P. (2001) *Arch. Biochem. Biophys.*, **385**, 220-230.
  93. Smith D.A., Ackland M.J., Jones B.C. (1997) *Drug Discov. Today*, **2**, 479-486.
  94. Szklarz G.D., Halpert J.R. (1997) *Life Sci.*, **61**, 2507-2520.
  95. Lightfoot T., Ellis S.W., Mahling J., Ackland M.J., Blaney F.E., Bijloo G.J., de Groot M.J., Vermeulen N.P., Blackburn G.M., Lennard M.S., Tucker G.T. (2000) *Xenobiotica*, **30**, 219-233.
  96. Rao S., Raoyama M., Schrag W.F., Truger A., Rettie, Jones J.P. (2000) *J. Med. Chem.*, **43**, 2789-2796.

97. Long A., Walker J.D. (2003) Environ. Toxicol. Chem., **22**(8), 1894-1899.
98. Xue L.L., Wang H.F., Wang Q.M., Szklarz G.D., Domanski T.L., Halpert J.R., Correia M.A. (2001) Chem. Res. Toxicol., **14**, 483-491.
99. de Groot M.J., Ackland M.J., Horne V.A., Alex A.A., Jones B.C. (1999) J. Med. Chem., **42**, 1515-1524.
100. SYBYL (1997) version 6.4, Tripos Inc., <http://www.tripos.com>.
101. Zamora I., Afzelius L., Cruciani G. (2003) J. Med. Chem., **46**, 2313-2324.
102. Goodford P.J.A. (1985) J. Med. Chem., **28**, 849-857.
103. GRID V.20 (2001) Molecular Discovery Ltd, (<http://www.moldiscovery.com>).
104. Ridderstrom M., Zamora I., Fjellstrom O., Andersson T.B. (2001) J. Med. Chem., **44**(24), 4072-4081.
105. Pastor M., Cruciani G., McLay I., Pickett S., Clementi S. (2000) J. Med. Chem., **43**(17), 3233-3243.
106. Park J.-Y., Harris D. (2003) J. Med. Chem., **46**, 1645-1660.
107. Иванов А.С., Скворцов В.С., Арчаков А.И. (2000) Вопр. мед. химии, **46**(6), 615-625.
108. Сеченых А.А., Дубанов А.В., Скворцов В.С., Иванов А.С., Арчаков А.И., Вильямс П., Косме Дж. Джонсон Е.Ф., Макри Д.Е. (2002) Вопр. мед. химии, **48**(5), 526-538.
109. Белкина Н.В., Скворцов В.С., Иванов А.С., Арчаков А.И. (1998) Вопр. мед. химии, **44**(5), 464-473.
110. Hansch C. (1972) Drug Metab. Rev., **1**, 1-14.
111. Hansch C., Lien E.J., Helmer F. (1968) Arch. Biochem. Biophys., **128**, 319-330.
112. Lewis D.F. (2000) Biochem. Pharmacol., **60**, 293-306.
113. Lewis D.F., Modi S., Dickins M. (2002) Drug Metab. Rev., **34**, 69-82.
114. Balakin K.V., Ekins S., Bugrim A., Ivanenkov Y.A., Korolev D., Nikolsky Y.V., Ivashchenko A.A., Savchuk N.P., Nikolskaya T. (2004) Drug Metab. Dispos., in press.
115. De Rienzo F., Fanelli F., Menziani M.C., De Benedetti P.G. (2000) J. Comp.-Aid. Mol. Des., **14**, 93-116.
116. Jones J.P., He M., Trager W.F., Rettie A.E. (1996) Drug Metab. Dispos., **24**, 1-6.
117. Strobl G.R., von Kruedener S., Stockigt J., Guengerich F.P., Wolff T. (1993) J. Med. Chem., **36**, 1136-1145.
118. de Groot M.J., Ackland M.J., Horne V.A., Alex A.A., Jones B.C. (1999) J. Med. Chem., **42**, 4062-4070.
119. Talele T.T., Kulkarni V.M. (1999) J. Chem. Inf. Comput. Sci., **39**, 204-210.
120. Lee H., Yeom H., Kim Y.G., Yoon C.N., Jin C., Choi J.S., Kim B.-R., Kim D.-H. (1998) Biochem. Pharmacol., **55**, 1369-1375.
121. Ekins S., Bravi G., Ring B.J., Gillespie T.A., Gillespie J.S., VandenBranden M., Wrighton S.A., Wikel J.H. (1999) J. Pharmacol. Exp. Ther., **288**, 21-29.
122. Jones B.C., Hawksworth G., Horne V.A., Newlands A., Morsman J., Tute M.S., Smith D.A. (1996) Drug Metab. Dispos., **24**, 260-266.
123. Mancy A., Broto P., Dijols S., Dansette P.M., Mansuy D. (1995) Biochemistry, **34**, 10365-10375.
124. Mancy A., Dijols S., Poli S., Guengerich P., Mansuy D. (1996) Biochemistry, **35**, 16205-16212.
125. Ekins S., Bravi G., Binkley S., Gillespie J.S., Ring B.J., Wikel J.H., Wrighton S.A. (2000) Drug Met. Dispos., **28**, 994-1002.
126. Islam S.A., Wolf C.R., Lennard M.S., Sternberg J.E. (1991) Carcinogenesis, **12**, 2211-2219.
127. Ekins S., Bravi G., Binkley S., Gillespie J.S., Ring B.J., Wikel J.H., Wrighton S.A. (1999) Pharmacogenetics, **9**, 477-489.
128. Lewis D.F., Eddershaw P.J., Goldfarb P.S., Tarbit M.H. (1996) Xenobiotica, **10**, 1067-1086.
129. Ekins S., Bravi G., Wikel J.H., Wrighton S.A. (1999) J. Pharmacol. Exp. Ther., **291**, 424-433.
130. Ekins S., Bravi G., Binkley S., Gillespie J.S., Ring B.J., Wikel J.H., Wrighton S.A.

- (1999) *J. Pharmacol. Exp. Ther.*, **290**, 429-438.
131. Ekins S., de Groot M., Jones J.P. (2001) *Drug Metab. Dispos.*, **29**, 936-944.
  132. Smith D.A., Ackland M.J., Jones B.C. (1997) *Drug Discov. Today*, **2**, 406-414.
  133. de Groot M.J., Ekins S. (2002) *Adv. Drug Deliv. Rev.*, **54**, 367-383.
  134. Afzelius L., Zamora I., Masimirembwa C.M., Karlen A., Andersson T.B., Mecucci S., Baroni M., Cruciani G. (2004) *J. Med. Chem.*, **47**(4), 907-914.
  135. Afzelius L., Masimirembwa C.M., Karlen A., Andersson T.B., Zamora I. (2002) *J. Comput. Aided. Mol. Des.*, **16**(7), 443-458.
  136. Afzelius L., Zamora I., Ridderstrom M., Andersson T.B., Karlen A., Masimirembwa C.M. (2001) *Mol. Pharm.*, **59**(4), 909-919.
  137. de Groot M.J., Vermeulen N.P.E. (1997) *Drug Metab. Rev.*, **29**, 747-799.
  138. Miners J.O., Smith P.A., Sorich M.J., McKinnon R.A., Mackenzie P.I. (2004) *Ann. Rev. Pharm.*, **44**, 1-25.
  139. Sorich M.J., McKinnon R.A., Miners J.O., Smith P.A. (2003) *J. Chem. Inf. Comput. Sci.*, **43**, 2019-2024.
  140. Buchwald P., Bodor N. (1999) *J. Med. Chem.*, **42**, 5160-5168.
  141. Sorich M.J., Smith P.A., McKinnon R.A., Miners J.O. (2002) *Pharmacogenetics*, **12**(8), 635-645.
  142. Smith P.A., McKinnon R.A., Sorich M.J., Miners J.O. (2002) *J. Med. Chem.*, **46**(9), 1617-1626.
  143. Sprague P.W., Hoffmann R. (1997) in: *Computer Assisted Lead Finding and Optimization Current Tools for Medicinal Chemistry: Catalyst Pharmacophore Models and their Utility as Queries for Searching 3D Databases* (H.van de Waterbeemd, B. Testa and G. Folkers eds.) VHCA, Basel, pp. 233-240.
  144. Balakin K.V., Ekins S., Bugrim A., Ivanenkov I.A., Korolev D., Nikolsky Y.V., Skorenko A.V., Ivashchenko A.A., Savchuk N.P., Nikolskaya T. (2004) *Drug Metab. Dispos.*, in press.
  145. Borodina Y., Sadyam A., Filimonov D., Blinova V., Dmitriev A., Poroikov V.V. (2003) *J. Chem. Inf. Comput. Sci.*, **43**(5), 1636-1646.
  146. Poroikov V.V., Filimonov D.A., Borodina Yu. V., Lagunin A.A., Kos A. (2000). *J. Chem. Inform. Comput. Sci.*, **40**(6), 1349-1355.
  147. Poroikov V.V., Filimonov D.A. (2002). *J. Comput. Aid. Molec. Des.*, **16**(11), 819-824.
  148. Stepanchikova A.V., Lagunin A.A., Filimonov D.A., Poroikov V.V. (2003). *Current Med. Chem.*, **10**(3), 225-233.
  149. Jacoby E., Schuffenhauer A., Floersheim P. (2003) *Drug News Perspect.*, **16**, 93-102.
  150. Lisitsa A.V., Gusev S.A., Karuzina I.I., Archakov A.I., Koymans L. (2001) *SAR QSAR Environ. Res.*, **12**(4), 359-366.
  151. Ellis L.B., Hershberger C.D., Wackett L.P. (1999) *Nucl. Acids Res.*, **27**(1), 373-376.
  152. Selkov E., Jr., Grechkin, Y., Mikhailova N., Selkov E. (1998) *Nucl. Acids Res.*, **26**(1), 43-45.

Поступила: 31. 05. 2004

#### COMPUTATIONAL METHODS FOR PREDICTION OF METABOLISM OF PHYSIOLOGICALLY ACTIVE SUBSTANCES

K.V. Balakin, Y.A. Ivanenkov

Chemical Diversity Research Institute, NIIRP Bldg. 1, ul. Rabochaya 2a, Khimki, Moscow reg., Russia

This article reviews the current status of computational methods for prediction of metabolism of physiologically active substances.

**Key words:** metabolism prediction, metabolic enzymes, cytochromes P450, metabolic stability, metabolic reactions, quantitative structure-metabolism relationship (QSMR) modeling.