

**Дмитриев Александр Викторович**

**КОМПЬЮТЕРНЫЙ ПРОГНОЗ ВЗАИМОДЕЙСТВИЯ  
КСЕНОБИОТИКОВ С ЦИТОХРОМАМИ P450 ЧЕЛОВЕКА**

03.00.28 - биоинформатика

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени

кандидата биологических наук

Москва - 2009 г.

Работа выполнена в Учреждении Российской академии медицинских наук Научно-исследовательском институте биомедицинской химии имени В.Н.Ореховича РАМН

**Научные руководители:**

доктор биологических наук,  
профессор

Поройков Владимир Васильевич

кандидат физико-математических наук

Филимонов Дмитрий Алексеевич

**Официальные оппоненты:**

доктор биологических наук,  
профессор

Каменская Марина Александровна

доктор физико-математических наук,  
профессор

Ефремов Роман Гербертович

**Ведущая организация:** Учреждение Российской академии наук Институт молекулярной биологии им. В.А. Энгельгардта РАН

Защита состоится «3» декабря 2009 года в 11:00 часов на заседании Диссертационного совета Д 001.010.01 при Учреждении Российской академии медицинских наук Научно-исследовательском институте биомедицинской химии имени В.Н.Ореховича РАМН по адресу: 119121, Москва, ул. Погодинская, д. 10.

С диссертацией можно ознакомиться в библиотеке Учреждения Российской академии медицинских наук Научно-исследовательском институте биомедицинской химии имени В.Н.Ореховича РАМН.

Автореферат разослан «\_\_» ноября 2009 г.

Ученый секретарь Диссертационного совета,  
кандидат химических наук

Е.А. Карпова

## 1. ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность проблемы.** Многие ксенобиотики, включая лекарства, подвергаются биотрансформации в организме человека. Информация о метаболитах лекарственных препаратов чрезвычайно важна, поскольку их биологическая активность, токсичность и биодоступность могут значительно отличаться от таковых у исходных веществ. Такая информация также используется при создании пролекарств с целью оптимизации фармакокинетических параметров исходной субстанции и фармакодинамических характеристик активного метаболита.

В организме человека лекарства метаболизируются несколькими ферментными системами, основной из которых является суперсемейство цитохромов P450 (Archakov A., Degtyarenko K., 1993). Около 75% лекарств метаболизируется цитохромами P450, в основном 3A4, 2C9, 2C19, 2D6 и 1A2 (Williams J. *et al.*, 2004). Эти ферменты осуществляют: ароматическое гидроксирование, алифатическое гидроксирование, N-деалкилирование, N-деметилирование, C-окисление с образованием альдегидов, кетонов, карбоновых кислот, O-деалкилирование, O-деметилирование, разрыв сложноэфирной связи, дегалогенирование (Guengerich F., 2001). Общим этапом перечисленных типов биотрансформации является окисление атома углерода субстрата цитохромом P450 (сокращенно **P450CARBONOX**, от англ. P450 carbon oxidation). Данные типы биотрансформации в совокупности составляют около 70% реакций, катализируемых указанными выше цитохромми P450; поэтому предсказание реакций биотрансформации, проходящих через этап P450CARBONOX является важной задачей для прогнозирования метаболитов. Многие ксенобиотики являются субстратами нескольких цитохромов P450. Под воздействием одного и того же цитохрома субстраты могут подвергаться одной или более реакциям; в то же время некоторые из реакций биотрансформации осуществляются более чем одним цитохромом и, напротив, некоторые реакции являются уникальными для конкретной изоформы. Поэтому особый интерес представляет прогноз биотрансформации ксенобиотиков под воздействием различных цитохромов P450 человека.

Исследование указанных выше пяти изоформ цитохрома P450 важно для оценки возможных межлекарственных взаимодействий из-за проявления ксенобиотиками ингибирующей и индуцирующей активности по отношению к цитохромам, поскольку одновременный прием пациентами нескольких препаратов может приводить к усилению либо ослаблению ожидаемых фармакологических эффектов (Hansten P. *et al.*, 2006).

Экспериментальное исследование метаболизма ксенобиотиков и межлекарственного взаимодействия является дорогостоящим и трудоемким (Wang J. *et al.*, 2007). Компьютерные методы позволяют прогнозировать биотрансформацию для многих тысяч химических структур, в том числе виртуальных, еще не синтезированных. Для

прогнозирования взаимодействия ксенобиотиков с цитохромами P450 предложено много подходов, различающиеся используемыми описаниями молекул и реакций, и математическими методами, включая ДСМ-метод (Джон Стюарт Миль) (Финн В.К., 2009), MetaDrug (Ekins S. *et al.*, 2005), MetaSite (Cruciani G. *et al.*, 2005), CATABOL (Jaworska J. *et al.*, 2002), COMPACT (Lewis D. *et al.*, 2001), TIMES, META, METEOR, MetabolExpert (Kulkarni S. *et al.*, 2003) и др.

Ю.В. Бородиной, А.В. Рудик с соавторами, была предложена методика предсказания биотрансформации ксенобиотиков по структурным формулам с использованием **MNA** (Multilevel Neighborhood of Atom, дескрипторы многоуровневых атомных окрестностей) и **RMNA** (Reacting Multilevel Neighborhood of Atom, дескрипторы многоуровневых атомных окрестностей реакций) дескрипторов (Borodina Yu. *et al.*, 2003, 2004; Рудик А., 2007) на основе алгоритмов программы **PASS** (Prediction of Activity Spectra for Substances). Однако в этих работах не реализовано прогнозирование взаимодействия ксенобиотиков с цитохромами P450 человека, прогноз биотрансформаций не учитывал видовую и ферментную специфичность, а прогноз метаболитов осуществлялся только для четырех типов биотрансформации.

**Цель работы:** реализовать компьютерный прогноз взаимодействия ксенобиотиков с цитохромами P450 человека в качестве субстратов, ингибиторов и индукторов, и прогноз сайтов окисления в молекуле субстрата.

**Задачи исследования:**

1. Создать базы данных, содержащие информацию о структурах и биотрансформации ксенобиотиков, и данные об их взаимодействии с цитохромами P450 3A4, 2C9, 2C19, 2D6 и 1A2 человека для формирования обучающих и тестовых выборок при реализации и верификации компьютерного прогноза.

2. Реализовать и верифицировать компьютерный прогноз взаимодействия ксенобиотиков с цитохромами P450 человека в качестве субстратов, ингибиторов и индукторов с использованием созданных баз данных и компьютерной программы PASS.

3. Разработать унифицированный способ записи биотрансформации с учетом положения окисляемого атома углерода при взаимодействии с цитохромами P450; реализовать и верифицировать прогноз сайтов окисления в молекуле субстрата.

4. На примере реакции ароматического гидроксילирования провести сопоставление прогноза сайтов окисления разработанными нами методами с квантовохимическими моделями.

**Научная новизна.** Впервые реализован компьютерный прогноз взаимодействия ксенобиотиков с цитохромами P450 3A4, 2C9, 2C19, 2D6 и 1A2 человека в качестве субстратов, ингибиторов и индукторов. Предложен новый метод прогноза сайтов

биотрансформации, проходящих через стадию окисления атома углерода (P450CARBONOX), позволяющий предсказывать реакции ароматического и алифатического гидроксирования, N-деалкилирования, N-деметилирования, C-окисления с образованием альдегидов, кетонов, карбоновых кислот, O-деалкилирования, O-деметилирования, разрыва сложноэфирной связи и дегалогенирования.

**Научно-практическая значимость.** Результаты работы могут быть использованы для прогноза метаболизма органических соединений при поиске и создании новых лекарств, оценки риска воздействия ксенобиотиков на организм человека, анализа межлекарственного взаимодействия при назначении пациентам нескольких лекарственных препаратов одновременно.

**Апробация работы.** Основные положения диссертационной работы представлены на симпозиумах «Биоинформатика и компьютерное конструирование лекарств», состоявшихся в рамках Российских национальных конгрессов «Человек и лекарство» XII (Москва, 2005 г.), XIII (Москва, 2006 г.), XIV (Москва, 2007 г.), XV (Москва, 2008 г.); на Московской международной конференции «Биотехнология и медицина» (Москва, 2006 г.), IV Московском международном конгрессе «Биотехнология: состояние и перспективы развития» (Москва, 2007 г.), научной конференции ИБМХ РАМН (Москва, 2007 г.), 4-ом Международном симпозиуме «Компьютерные методы в токсикологии и фармакологии, включающие Интернет-ресурсы (СМТPI-2007)» (Москва, 2007 г.), на Германской конференции по хемоинформатике (Гослар, Германия, 2008 г.).

**Работа выполнена при поддержке грантов** Международного фонда технологий и инвестиций (МФТИ) № DPG.55229907.00106 и № DPG.55229907.00190; Государственного контракта ГК № 02.740.11.0306.

**Публикации.** По материалам диссертации опубликовано 18 печатных работ в российских и зарубежных научных изданиях, в том числе 4 статьи в рецензируемых журналах, входящих в список ВАК, 14 публикаций в сборниках трудов научных конференций.

**Объем и структура диссертации.** Диссертационная работа изложена на 129 страницах машинописного текста, содержит 28 рисунков и 25 таблиц. Работа состоит из введения, обзора литературы, описания объектов и методов исследования, полученных результатов, их обсуждения, выводов, списка цитированной литературы, включающего 147 источников, и приложения.

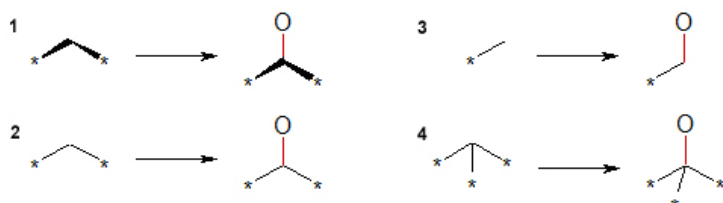
**Обзор литературы** содержит описание основных цитохромов P450 человека, принимающих участие в метаболизме ксенобиотиков, экспериментальных методик исследования метаболизма ксенобиотиков, компьютерных методов моделирования структуры и функции различных цитохромов P450, и прогноза биотрансформации.

## 2. ОБЪЕКТЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Для создания базы данных (БД) субстратов, ингибиторов, индукторов цитохромов P450 человека, и БД реакций биотрансформации использован программный комплекс **ISIS BASE, ISIS DRAW, ISIS HOST**; и следующие источники информации: БД **Metabolite (Symyx MDL)**, БД **Metabolism V 2002.1 (Accelrys)**, база знаний **Cytochrome P450 Knowledgebase (СРК)** (Lisitsa A. *et al.*, 2001), данные из публикаций в научных журналах.

Анализ взаимосвязей «структура-свойство» и прогноз взаимодействия органических соединений с цитохромами P450 3A4, 2C9, 2C19, 2D6 и 1A2 в качестве субстратов, ингибиторов и индукторов выполнен с использованием **программы PASS** (Свидетельство Роспатента № 2006613275), которая на основе обучающей выборки прогнозирует по структурной формуле органического соединения вероятность его принадлежности (**Pa**) или не принадлежности (**Pi**) к определенному классу.

Прогноз сайтов окисления атома углерода субстратов цитохромов P450 выполнен с использованием **программы METAPREDICT** (Свидетельство Роспатента № 2004610666), которая осуществляет генерацию гипотетических метаболитов с использованием фрагментов реакций. Фрагменты реакций содержат по одному ближайшему соседу атома, к которому присоединена изменяющаяся в ходе биотрансформации связь (см. рис. 1). Фрагмент 1 описывает ароматическую структуру, остальные фрагменты - окисление алифатических соединений.



**Рис. 1.** Фрагменты реакций P450CARBONOX.

Для предсказания сайтов P450CARBONOX реакции из обучающей выборки кодируются RMNA дескрипторами, разработанными ранее с нашим участием (Borodina Yu. *et al.*, 2004). Для каждой возможной реакции биотрансформации программа METAPREDICT подсчитывает вероятности её наличия (**Pt**) и отсутствия (**Pf**). Чем больше величина  $\Delta P = Pt - Pf$ , тем вероятнее прохождение данной реакции.

**Оценка точности прогноза.** Точность прогноза оценивали в компьютерных экспериментах: 1) скользящий контроль с исключением по одному для обучающих выборок; 2) двадцатикратное случайное разбиение исходной выборки на обучающую и тестовую подвыборки в пропорции 80:20%; 3) прогноз субстратов, ингибиторов, индукторов цитохромов P450 и сайтов P450CARBONOX для тестовых выборок.

Использовались следующие статистические критерии:

**Инвариантная точность прогноза IAP** (Invariant Accuracy of Prediction) – рассчитывается для каждого класса по следующей формуле:

$$IAP = \frac{N(Pa(Sa) > Pa(Si))}{Na \cdot Ni} \cdot 100\%$$

где  $N(Pa(Sa) > Pa(Si))$  – число случаев, когда  $Pa$  для соединений, принадлежащих к определенному классу (положительные примеры,  $Sa$ ) превышает  $Pa$  для соединений, к нему не принадлежащих (отрицательные примеры,  $Si$ );  $Na$  и  $Ni$  – число положительных и отрицательных примеров.

Чувствительность, специфичность, точность и индекс Йодена, рассчитываемые по формулам:

$$SENS = TP / (TP + FN)$$

$$SPEC = TN / (FP + TN)$$

$$ACR = (TP + TN) / (TP + FN + FP + TN)$$

$$YDN = SENS + SPEC - 1$$

где: SENS - чувствительность, SPEC - специфичность, ACR - точность прогноза, YDN - индекс Йодена, TP - число истинно положительных (число положительных примеров, спрогнозированных как положительные), TN - число истинно отрицательных, FP - число ложно положительных, FN - число ложно отрицательных предсказаний.

Индекс Йодена включает как специфичность, так и чувствительность, и отражает сбалансированность прогноза.

Для квантовохимических расчетов при построении известных из литературы моделей ароматического гидроксирования ксенобиотиков под воздействием цитохромов P450, использовался комплекс программ **МОРАС 6.0**. (Stewart J., *et al.*, 1990), интегрированный в **Sybyl 6.9** фирмы **Tripos** (Clark M., *et al.*, 1989).

### 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

#### 3.1. Создание баз данных о метаболизме ксенобиотиков

##### 3.1.1. БД по взаимодействию ксенобиотиков с цитохромами P450

В основе БД по взаимодействию ксенобиотиков с цитохромами P450 лежит информация из БД Metabolite (Symyx MDL), обзора Слободана Рандича «Summary of information on human CYP enzymes: human P450 metabolism data» (Rendic S., 2002), и БД СРК (Lisitsa A. *et al.*, 2001).

БД по взаимодействию ксенобиотиков с цитохромами P450 включает в себя 2051 запись, содержащие структурные формулы веществ, информацию об их взаимодействии с

различными изоформами цитохрома P450 (субстрат, ингибитор, индуктор), ссылки на первоисточники. Сводная информация о содержании БД приведена в таблице 1.

**Таблица 1.** Содержание БД по взаимодействию ксенобиотиков с цитохромами P450

Тип взаимодействия	Кол-во молекул	Тип взаимодействия	Кол-во молекул	Тип взаимодействия	Кол-во молекул
CYP1A2 substrate	650	CYP1A2 inhibitor	127	CYP1A2 inducer	24
CYP2C19 substrate	406	CYP2C19 inhibitor	84	CYP2C19 inducer	8
CYP2C9 substrate	499	CYP2C9 inhibitor	143	CYP2C9 inducer	24
CYP2D6 substrate	592	CYP2D6 inhibitor	178	CYP2D6 inducer	8
CYP3A4 substrate	1103	CYP3A4 inhibitor	225	CYP3A4 inducer	71

### 3.1.2. БД биотрансформации

Созданная нами БД содержит 3042 записи реакций биотрансформации с участием цитохромов P450 3A4, 2C9, 2C19, 2D6 и 1A2, наблюдаемых при метаболизме ксенобиотиков в экспериментах *in vivo* и *in vitro* (таблица 2), из БД Metabolite (Symyx MDL).

**Таблица 2.** Содержание БД биотрансформации

Изоформа цитохрома P450	Количество субстратов	Количество реакций
CYP1A2	658	1178
CYP2C19	422	703
CYP2C9	503	836
CYP2D6	613	1037
CYP3A4	1069	2017

БД содержит реакции биотрансформации, структурные формулы субстратов и метаболитов, наименования классов биотрансформации и соответствующих изоформ цитохрома P450.

## 3.2. Прогноз субстратов, ингибиторов и индукторов цитохромов P450

### 3.2.1. Обучающая выборка

БД по взаимодействию ксенобиотиков с цитохромами P450 (см. п. 3.1.1.) была использована для создания обучающей выборки, содержащей структурные формулы органических соединений, названия изоформ цитохрома P450 с указанием типа взаимодействия (субстрат, ингибитор, индуктор). Данная обучающая выборка содержит 2051 соединение, взаимодействующее с цитохромами P450, далее - выборка «O1».



### 3.2.2. Обучение компьютерной программы PASS

В ходе обучения программы PASS с использованием выборки O1 осуществлялся скользящий контроль с исключением по одному, а также было произведено двадцатикратное случайное разбиение исходной выборки на обучающую и тестовую подвыборки в пропорции 80:20%, с вычислением значений 10%, 50% и 90% квантилей значений IAP. Полученные результаты приведены в таблице 3.

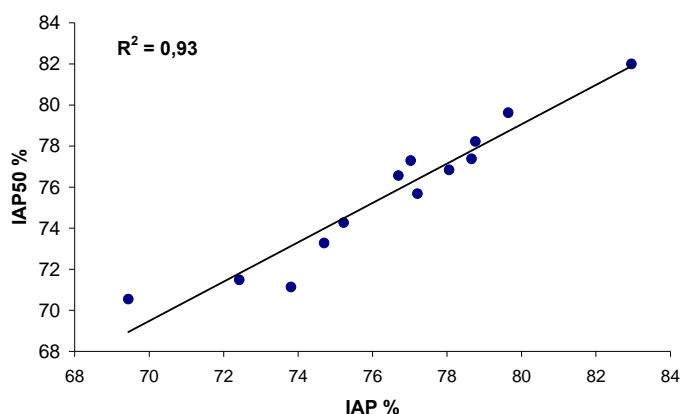
**Таблица 3.** Оценка качества прогноза для выборки O1

Тип взаимодействия (число молекул)	IAP, %	IAP90, %	IAP50, %	IAP10, %
CYP1A2 substrate (650)	77	79	77	73
CYP2C19 substrate (406)	78	81	77	74
CYP2C9 substrate (499)	75	76	73	70
CYP2D6 substrate (592)	79	81	78	76
CYP3A4 substrate (1103)	75	77	74	72
<b>Среднее значение по субстратам:</b>	<b>77</b>	<b>79</b>	<b>76</b>	<b>73</b>
CYP1A2 inhibitor (127)	77	84	77	71
CYP2C19 inhibitor (84)	74	78	71	64
CYP2C9 inhibitor (143)	79	84	77	72
CYP2D6 inhibitor (178)	80	83	80	74
CYP3A4 inhibitor (225)	72	77	71	67
<b>Среднее значение по ингибиторам:</b>	<b>76</b>	<b>81</b>	<b>75</b>	<b>69</b>
CYP1A2 inducer (24)	83	95	82	67
CYP2C9 inducer (24)	69	87	71	50
CYP3A4 inducer (71)	77	85	76	66
<b>Среднее значение по индукторам:</b>	<b>77</b>	<b>89</b>	<b>76</b>	<b>61</b>
<b>Среднее значение по 13 типам взаимодействия</b>	<b>77</b>	<b>82</b>	<b>76</b>	<b>69</b>

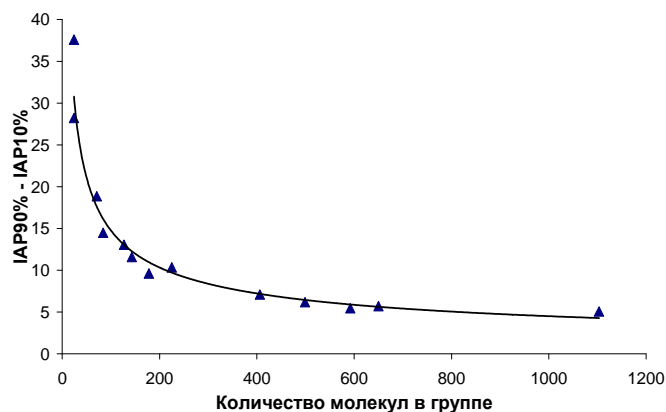
Число соединений с определенным типом взаимодействия в обучающей выборке O1 варьирует от 24 (для индукторов CYP1A2 и CYP2C9) до 1103 (субстраты CYP3A4). Индукторов цитохромов P450 2D6 и 2C19 в обучающей выборке по 8 молекул, что недостаточно для корректного построения прогноза и проведения теста с двадцатикратным случайным разбиением на обучающую и тестовую подвыборки. Из данных, приведенных в таблице 3, следует, что доверительные интервалы различны для групп субстратов, ингибиторов и индукторов, и зависят от количества молекул в соответствующей группе.

Среднее значение точности прогноза для 13 типов взаимодействия с цитохромами P450 равно 77%. При сопоставлении значений 50% квантилей IAP, полученных при двадцатикратном случайном разбиении исходной выборки, и значений IAP, полученных при скользящем контроле с исключением по одному, установлена корреляция со значением  $R^2 = 0,93$  (см. рис. 2). Это свидетельствует об устойчивости прогноза для 13

исследованных типов взаимодействия ксенобиотиков с цитохромами P450. Широта полученных 10-90% доверительных интервалов для субстратов (в среднем 6%) и ингибиторов (в среднем 12%) всех пяти изоформ цитохрома P450 свидетельствует об устойчивости прогноза. Прогноз же индукторов, в сравнении с субстратами и ингибиторами, является менее устойчивым, поскольку доверительный интервал в среднем по группе индукторов равен 28%, что может быть объяснено небольшим количеством положительных примеров в обучающей выборке. Зависимость широты доверительных интервалов от количества молекул в каждой группе субстратов, ингибиторов и индукторов может быть описана графиком степенной функции, представленным на рисунке 3.



**Рис. 2.** Сопоставление IAP и IAP50, %.



**Рис. 3.** Зависимость широты доверительных интервалов от количества молекул в каждой группе субстратов, ингибиторов и индукторов.

Для верификации прогноза и оценки предсказательной способности программы PASS проведен компьютерный эксперимент с тестовой выборкой.

### 3.2.3. Тестовая выборка для оценки предсказательной способности

Для оценки способности прогнозирования взаимодействия ксенобиотиков с цитохромами P450 в качестве субстратов, ингибиторов и индукторов была использована тестовая выборка, подготовленная на основе ежегодного издания «The Top 100 Drug Interaction 2006 N&H Publications» (Hansten P. *et al.*, 2006). Тестовая выборка содержит структурные формулы 260 молекул лекарств и данные об их клинически наблюдаемом взаимодействии с цитохромами P450 1A2, 2C9, 2C19, 2D6 и 3A4, полученные в экспериментах *in vivo* для человека. Далее по тексту эта выборка будет именоваться как выборка «Т1». Выборка Т1 представлена в приложении А диссертации.

### 3.2.4. Тестирование предсказательной способности алгоритмов PASS

В эксперименте с тестовой выборкой Т1 использовалась обучающая выборка О1 из которой предварительно были удалены молекулы тестовой выборки и данные об их

взаимодействии с цитохромами. Для каждого типа взаимодействия, предсказанного для конкретной молекулы, программой PASS рассчитываются вероятности наличия взаимодействия ( $P_a$ ) и его отсутствия ( $P_i$ ). Все типы взаимодействия, предсказанные с вероятностями  $P_a > P_i$ , рассматривались как возможные. Для каждого спрогнозированного типа взаимодействия были рассчитаны чувствительность, специфичность, точность прогноза и индекса Йодена, значения которых представлены в таблице 4.

**Таблица 4.** Оценка предсказательной способности PASS для независимой тестовой выборки

<b>Субстраты</b>	<b>TP</b>	<b>TN</b>	<b>FP</b>	<b>FN</b>	<b>SENS</b>	<b>SPEC</b>	<b>ACR</b>	<b>YDN</b>
CYP1A2 substrate	17	180	52	11	0,61	0,78	0,76	0,38
CYP2C9 substrate	29	155	69	7	0,81	0,69	0,71	0,50
CYP2C19 substrate	18	151	84	7	0,72	0,64	0,65	0,36
CYP2D6 substrate	54	143	57	6	0,90	0,72	0,76	0,62
CYP3A4 substrate	118	61	44	37	0,76	0,58	0,69	0,34
<b>Среднее:</b>	<b>47,2</b>	<b>138</b>	<b>61,2</b>	<b>13,6</b>	<b>0,76</b>	<b>0,68</b>	<b>0,71</b>	<b>0,44</b>
<b>Ингибиторы</b>	<b>TP</b>	<b>TN</b>	<b>FP</b>	<b>FN</b>	<b>SENS</b>	<b>SPEC</b>	<b>ACR</b>	<b>YDN</b>
CYP1A2 inhibitor	5	204	46	5	0,50	0,82	0,80	0,32
CYP2C9 inhibitor	17	190	43	10	0,63	0,82	0,80	0,45
CYP2C19 inhibitor	3	179	65	13	0,19	0,73	0,70	-0,08
CYP2D6 inhibitor	21	151	82	6	0,78	0,65	0,66	0,43
CYP3A4 inhibitor	20	167	61	12	0,63	0,73	0,72	0,36
<b>Среднее:</b>	<b>13,2</b>	<b>178,2</b>	<b>59,4</b>	<b>9,2</b>	<b>0,54</b>	<b>0,75</b>	<b>0,74</b>	<b>0,29</b>
<b>Индукторы</b>	<b>TP</b>	<b>TN</b>	<b>FP</b>	<b>FN</b>	<b>SENS</b>	<b>SPEC</b>	<b>ACR</b>	<b>YDN</b>
CYP1A2 inducer	1	238	19	2	0,33	0,93	0,92	0,26
CYP2C9 inducer	4	219	34	3	0,57	0,87	0,86	0,44
CYP3A4 inducer	8	181	62	9	0,47	0,74	0,73	0,22
<b>Среднее:</b>	<b>4,3</b>	<b>212,7</b>	<b>38,3</b>	<b>4,7</b>	<b>0,46</b>	<b>0,85</b>	<b>0,83</b>	<b>0,30</b>
<i>Среднее по 13 типам взаимодействия:</i>	<i>24,23</i>	<i>170,69</i>	<i>55,23</i>	<i>9,85</i>	<i>0,61</i>	<i>0,75</i>	<i>0,75</i>	<i>0,35</i>

В среднем значение точности предсказания 13 типов взаимодействия для тестовой выборки  $ACR=0,75$ , значение индекса Йодена  $YDN=0,35$ , что отражает хорошую точность и сбалансированность прогноза. При рассмотрении отдельных типов взаимодействия можно отметить лучшую сбалансированность прогноза для субстратов цитохромов ( $YDN=0,44$ ) по сравнению с ингибиторами ( $YDN=0,29$ ) и индукторами ( $YDN=0,30$ ). В ряду ингибиторов выделяется низким значением чувствительности (SENS), прогноз ингибиторов цитохрома P450 2C19, составляющий 0,19 и обуславливающий отрицательное значение индекса Йодена. Более низкая сбалансированность прогноза для

ингибиторов по сравнению с субстратами может быть обусловлена структурой обучающей выборки, содержащей как неконкурентные, так и конкурентные ингибиторы. Прогноз индукторов также является менее сбалансированным, что объясняется небольшим количеством положительных примеров в обучающей выборке и разнообразием механизмов, опосредующих процесс индукции цитохромов (взаимодействие с различными ядерными рецепторами) и химической гетерогенностью индукторов.

Полученные результаты позволяют говорить о хорошей предсказательной способности компьютерного прогноза взаимодействия ксенобиотиков с цитохромами P450 человека с использованием алгоритмов PASS при пороге  $P_a > P_i$ .

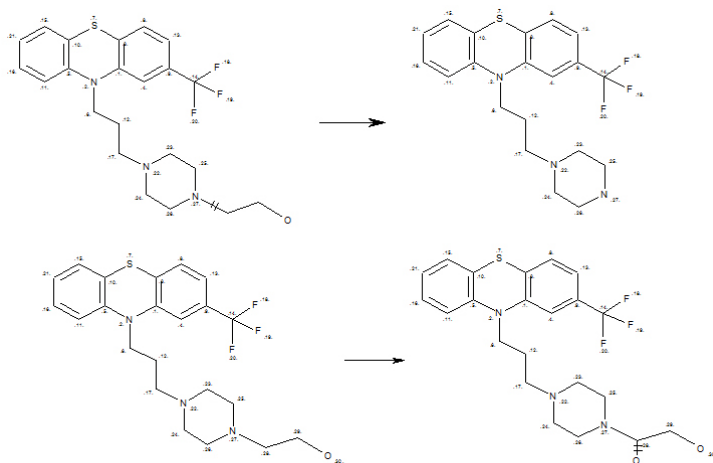
### 3.3. Построение и верификация моделей прогноза сайтов P450CARBONOX

После определения того, что молекула является субстратом того или иного цитохрома P450, для прогноза метаболитов, необходимо определить сайт биотрансформации. Для оценки сайтов окисления атома углерода ксенобиотиков была разработана унифицированная форма записи биотрансформации, названная P450CARBONOX.

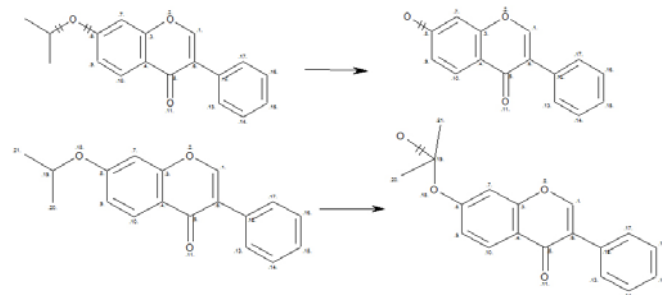
#### 3.3.1. Унифицированный способ записи биотрансформации P450CARBONOX

При окислении субстратов цитохром P450 взаимодействует с отдельным атомом углерода или гетероатомом (Guengerich F., 2001). Окисление атома углерода в молекуле ксенобиотиков под воздействием цитохрома P450 приводит к образованию промежуточных продуктов, таких, как радикал углерода или карбокатион, которые в последующем превращаются в продукты метаболизма. Окисление атома углерода может быть отображено как присоединение гидроксильной группы в молекуле продукта реакции, т.е. представлено аналогично биотрансформации гидрокселирования углерода. Примеры записей реакций, с применением разработанного нами унифицированного способа записи, приведены ниже.

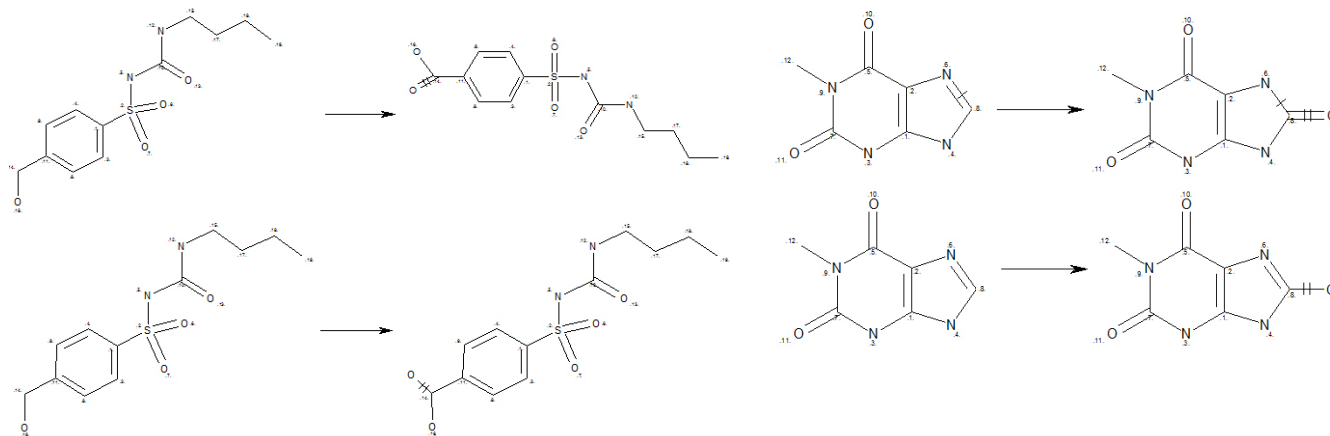
#### *N*-деалкилирование



#### *O*-деалкилирование



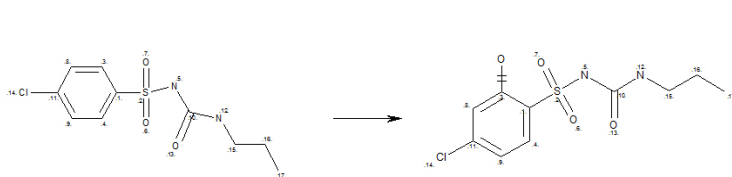
## Преобразование гидрокси-группы в карбокси-группу      Присоединение оксо-группы



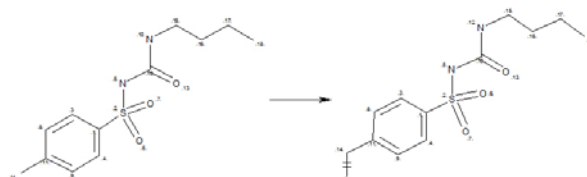
Для каждого типа биотрансформации верхняя запись представляет собой исходную схему, нижняя - схему, представленную на основе присоединения гидроксильной группы к углероду в молекуле продукта.

Схемы биотрансформаций ароматического и алифатического гидроксирования не требуют применения унифицированного способа записи, поскольку изначально описаны как присоединение гидроксила к окисляемому атому углерода субстрата:

### Ароматическое гидроксирование



### Алифатическое гидроксирование



Предложенный нами унифицированный способ записи реакций биотрансформации с учетом положения окисляемого цитохромом P450 атома углерода в молекуле субстрата был использован для прогноза сайтов ароматического гидроксирования, алифатического гидроксирования, N-деалкилирования, N-деметилирования, C-окисления с образованием альдегидов, кетонов, карбоновых кислот, O-деалкилирования, O-деметилирования, разрыва сложноэфирной связи и дегалогенирования.

### 3.3.2. Подготовка выборок P450CARBONOX

На основе БД биотрансформаций были созданы выборки для обучения программы METAPREDICT с целью прогноза сайтов P450CARBONOX под воздействием различных изоформ P450. Было проведено преобразование различных реакций биотрансформации, происходящих в результате окисления атома углерода цитохромом P450 к единой форме с использованием унифицированного способа записи P450CARBONOX. При этом для описания сайтов P450CARBONOX необходимо лишь четыре фрагмента реакций биотрансформации, представленных на рисунке 1. Всем модифицированным реакциям, а также реакциям ароматического и алифатического гидроксирования, в обучающих

выборках были присвоены дополнительные наименования P450CARBONOX. В итоге, в один тип «окисление атома углерода цитохромом P450», было объединено более десяти различных типов биотрансформации, реализующихся через окисление атома углерода.

### 3.3.3. Выборки для прогнозирования сайтов P450CARBONOX

Были подготовлены пять выборок, содержащих информацию о реакциях биотрансформации P450CARBONOX под воздействием различных изоформ P450. Количественные характеристики выборок представлены во втором столбце (количество реакций) табл. 5.

### 3.3.4. Обучение компьютерной программы METAPREDICT

Для каждой из пяти выборок с использованием программы METAPREDICT были рассчитаны значения IAP (см. таблицу 5) при различных уровнях RMNA дескрипторов (с первого по пятый).

**Таблица 5.** Значения IAP в зависимости от используемого уровня RMNA дескрипторов

Обучающая выборка	Количество реакций	RMNA 1, IAP%	RMNA 2, IAP%	RMNA 3, IAP%	RMNA 4, IAP%	RMNA 5, IAP%
CYP1A2	819	69	73	71	68	63
CYP2C9	636	67	71	69	66	61
CYP2C19	541	65	70	69	66	61
CYP2D6	818	72	75	72	69	64
CYP3A4	1385	69	72	71	69	65
<i>Среднее значение:</i>	<b>840</b>	<b>68</b>	<b>72</b>	<b>70</b>	<b>68</b>	<b>63</b>

Из данных, приведенных в таблице 5, следует, что лучшая точность прогноза достигается при использовании RMNA дескрипторов второго уровня (72%). Полученные результаты позволяют прийти к заключению о применимости предложенного нами унифицированного способа записи P450CARBONOX, и о возможности использования программы METAPREDICT для предсказания сайтов окисления под воздействием цитохромов P450 1A2, 2C9, 2C19, 2D6 и 3A4.

### 3.3.5. Тестовые выборки для верификации прогноза сайтов P450CARBONOX

В БД биотрансформации 363 реакции обобщенного класса P450CARBONOX, которые могут быть осуществлены одной из пяти изоформ P450: для 1A2 - 261 реакция, для 2C19 - 248 реакций, для 2C9 - 249 реакций, для 2D6 - 258 реакций и для 3A4 - 265 реакций. В этих реакциях в качестве субстратов фигурируют 158 молекул ксенобиотиков.

### 3.3.6. Верификация прогноза сайтов P450CARBONOX

Для пяти тестовых выборок был проведен прогноз сайтов P450CARBONOX с использованием обучающих выборок, созданных с использованием различных уровней RMNA дескрипторов. Из обучающих выборок предварительно были удалены реакции, известные для 158 молекул тестовой выборки. Все предсказанные с вероятностями  $\Delta P > 0$  сайты P450CARBONOX рассматривались как возможные. В таблице 6 представлены значения чувствительности, специфичности, точности и индекса Йодена.

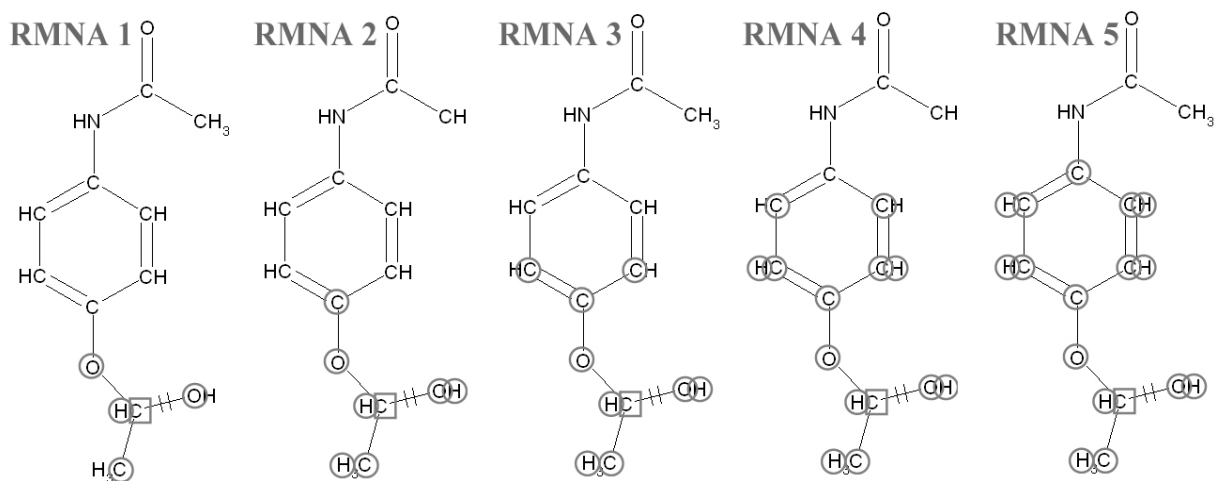
**Таблица 6.** Оценки качества прогноза сайтов биотрансформаций

Обучающая выборка	TP	TN	FP	FN	SENS	SPEC	ACR	YDN
CYP1A2 RMNA 1	0	597	0	229	0	1,00	0,72	0
CYP2C9 RMNA 1	3	608	0	221	0,01	1,00	0,73	0,01
CYP2C19 RMNA 1	2	600	0	224	0,01	1,00	0,73	0,01
CYP2D6 RMNA 1	3	596	0	223	0,01	1,00	0,73	0,01
CYP3A4 RMNA 1	4	596	1	236	0,02	1,00	0,72	0,01
<b>Среднее значение:</b>	<b>2,4</b>	<b>599,4</b>	<b>0,2</b>	<b>226,6</b>	<b>0,01</b>	<b>1,00</b>	<b>0,73</b>	<b>0,01</b>
CYP1A2 RMNA 2	73	1298	57	181	0,29	0,96	0,85	0,25
CYP2C9 RMNA 2	152	1188	188	93	0,62	0,86	0,83	0,48
CYP2C19 RMNA 2	135	1152	207	109	0,55	0,85	0,80	0,40
CYP2D6 RMNA 2	111	1229	121	138	0,45	0,91	0,84	0,36
CYP3A4 RMNA 2	147	1200	181	113	0,57	0,87	0,82	0,43
<b>Среднее значение:</b>	<b>123,6</b>	<b>1213,4</b>	<b>150,8</b>	<b>126,8</b>	<b>0,49</b>	<b>0,89</b>	<b>0,83</b>	<b>0,38</b>
CYP1A2 RMNA 3	144	1217	202	117	0,55	0,86	0,81	0,41
CYP2C9 RMNA 3	178	1085	353	71	0,71	0,75	0,75	0,47
CYP2C19 RMNA 3	148	1085	335	100	0,60	0,76	0,74	0,36
CYP2D6 RMNA 3	154	1178	234	104	0,60	0,83	0,80	0,43
CYP3A4 RMNA 3	181	1148	298	84	0,68	0,79	0,78	0,48
<b>Среднее значение:</b>	<b>161</b>	<b>1142,6</b>	<b>284,4</b>	<b>95,2</b>	<b>0,63</b>	<b>0,80</b>	<b>0,77</b>	<b>0,43</b>
CYP1A2 RMNA 4	168	1103	324	93	0,64	0,77	0,75	0,42
CYP2C9 RMNA 4	161	1027	419	88	0,65	0,71	0,70	0,36
CYP2C19 RMNA 4	154	1020	408	94	0,62	0,71	0,70	0,34
CYP2D6 RMNA 4	122	1104	316	136	0,47	0,78	0,73	0,25
CYP3A4 RMNA 4	179	1094	360	86	0,68	0,75	0,74	0,43
<b>Среднее значение:</b>	<b>156,8</b>	<b>1069,6</b>	<b>365,4</b>	<b>99,4</b>	<b>0,61</b>	<b>0,75</b>	<b>0,73</b>	<b>0,36</b>
CYP1A2 RMNA 5	176	880	548	85	0,67	0,62	0,63	0,29
CYP2C9 RMNA 5	176	833	614	73	0,71	0,58	0,59	0,28
CYP2C19 RMNA 5	166	811	618	82	0,67	0,57	0,58	0,24
CYP2D6 RMNA 5	169	909	512	89	0,66	0,64	0,64	0,29
CYP3A4 RMNA 5	119	975	480	146	0,45	0,67	0,64	0,12
<b>Среднее значение:</b>	<b>161,2</b>	<b>881,6</b>	<b>554,4</b>	<b>95</b>	<b>0,63</b>	<b>0,61</b>	<b>0,62</b>	<b>0,24</b>

Из приведенных в таблице 6 данных следует, что наилучшая средняя точность прогноза (83%) получена с использованием RMNA дескрипторов второго уровня, но при

этом чувствительность не высока (SENS=0,49). Наиболее сбалансированный по специфичности и чувствительности прогноз (SENS=0,63, SPEC=0,80, YDN=0,43) получен с использованием RMNA дескрипторов третьего уровня.

RMNA дескрипторы первого уровня описывают непосредственное окружение отдельных атомов, что, как следует из результатов теста, недостаточно для прогноза сайтов окисления. RMNA дескрипторы последующих уровней охватывают окрестности большего размера (см. рис. 4).



**Рис. 4.** Фрагменты, описываемые RMNA дескрипторами различных уровней для окисляемого атома углерода (обозначен прямоугольником).

Проведенная нами верификация прогноза сайтов P450CARBONOX свидетельствует о хорошей предсказательной способности при использовании RMNA дескрипторов второго и третьего уровней.

Результаты эксперимента позволяют прийти к заключению о правомочности применения предложенного нами унифицированного способа записи P450CARBONOX для окисления атома углерода, и о возможности предсказания сайтов окисления атомов углерода под воздействием различных цитохромов P450 1A2, 2C9, 2C19, 2D6 и 3A4.

Поскольку для прогнозов сайтов биотрансформации в литературе нередко используются квантовохимические методы, представляло интерес сопоставить с ними предложенный нами статистический подход. Такое сопоставление было проведено на примере реакции ароматического гидроксирования.

### 3.4. Сопоставление статистического и квантовохимических методов прогноза ароматического гидроксирования

Ароматическое гидроксирование составляет примерно 20% всех реакций, осуществляемых цитохромами P450 в организме человека. Точность предсказания сайтов



биотрансформации ароматического гидроксирования с помощью статистической модели (программа METAPREDICT) была сопоставлена с точностью прогноза ароматического гидроксирования, полученной с использованием двух квантовохимических моделей активных частиц кислорода, образованных в результате взаимодействия с гемовым железом цитохрома P450. Для сравнения были использованы оксеноидная модель (Korzekwa, K. *et al.* 1985) и модель метокси-радикала (Jones, J. *et al.* 2002). Подробные описания моделей и проведенных компьютерных экспериментов представлены в п.п. 1.3.6 и 3.4. диссертации, соответственно.

Таблица 7 содержит рассчитанные с использованием этих моделей величины IAP для трех тестовых выборок, содержащих производные бензола, соединения гетероциклической структуры и субстраты цитохрома CYP2D6.

**Таблица 7.** Точность предсказания (IAP%), рассчитанная для моделей: статистической (s), оксеноидной (o) и с метокси радикалом (m)

	<b>Производные бензола</b>	<b>Гетероциклы</b>	<b>Субстраты CYP2D6</b>
Количество соединений	15	17	15
IAPs, %	85	83	90
IAPo %	95	Не применим	Не применим
IAPm %	84	45	70

Как следует из приведенных в таблице 7 данных, наиболее высокая точность предсказания ароматического гидроксирования замещенных бензолов (IAP=95%) получена для оксеноидной модели. Однако, эта модель может использоваться только для предсказания гидроксирования в замещенных бензолах и не применима для ароматических гетеро- и полициклов. Для статистической модели и модели метокси-радикала точности предсказания для замещенных бензолов составляют 85% и 84%, соответственно. Для веществ более сложной структуры точность предсказания с использованием модели метокси радикала оказывается неудовлетворительной (IAP=45%), в то время как разработанный нами статистический метод дает неплохие результаты (IAP=83%).

Средние значения точности прогноза, полученные для выборки субстратов цитохрома P450 2D6, с использованием статистической модели и модели метокси-радикала, равны 90% и 70%, соответственно.

Для всех трех выборок точность прогноза с использованием статистической модели оказалась выше или сопоставимой с точностью квантовохимических методов. При этом статистическая модель применима для всех рассмотренных нами классов химических

соединений, содержащихся в тестовых выборках, а скорость расчетов с её использованием на два порядка превышает скорость квантовохимических методов.

## **ВЫВОДЫ**

1. Созданы базы данных по взаимодействию ксенобиотиков с цитохромами P450 человека и осуществляемых ими реакций биотрансформации, которые были использованы для формирования обучающих и тестовых выборок при реализации и верификации компьютерного прогноза взаимодействия ксенобиотиков с цитохромами P450 человека и сайтов окисления в молекуле субстрата.

2. Реализован компьютерный прогноз взаимодействия ксенобиотиков с цитохромами P450 1A2, 2C9, 2C19, 2D6 и 3A4 человека; проведенная верификация показала хорошую точность и устойчивость прогноза субстратов и ингибиторов (77 и 76%, соответственно). Средняя точность прогноза индукторов цитохромов P450 составляет 75%; при этом прогноз индукторов цитохромов P450 1A2, 2D6 и 3A4 является менее устойчивым, а прогноз индукторов 2C19 и 2D6 в настоящее время не может быть реализован из-за малого количества соединений в обучающей выборке.

3. Разработанный унифицированный способ записи реакций биотрансформации с учетом положения окисляемого цитохромом P450 атома углерода в молекуле субстрата был использован для прогноза сайтов ароматического гидроксирования, алифатического гидроксирования, N-деалкилирования, N-деметилирования, C-окисления с образованием альдегидов, кетонов, карбоновых кислот, O-деалкилирования, O-деметилирования, разрыва сложноэфирной связи, и дегалогенирования. Показано, что использование RMNA дескрипторов второго уровня обеспечивает наилучшую среднюю точность прогноза (72%).

4. Показано, что разработанный нами статистический метод обеспечивает точность прогноза сайтов ароматического гидроксирования под воздействием цитохромов P450, сопоставимую с квантовохимическими моделями, а скорость расчетов на два порядка превышает скорость квантовохимических методов.

## СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Sadym A.V., Borodina Yu.V., Dmitriev A.V., Blinova V.G, Filimonov D.A., Poroikov V.V. Computer generation of metabolites for drug-like molecules. // Abstr. 2<sup>nd</sup> International Symposium «Computational Methods in Toxicology and Pharmacology Integrating Internet Resources (СМТПИ-2003)». - Thessaloniki (Greece). - 2003. - P. 67.
2. Sadym A.V., Borodina Yu.V., Filimonov D.A., Poroikov V.V., Dmitriev A.V., Blinova V.G. Computer predicting of drug-like substances metabolism: from biotransformation reactions to metabolic network. // Abstr. 4<sup>th</sup> International Symposium on Pharmaceutical Chemistry - Istanbul (Turkey). - 2003. - P. 259-260.
3. Borodina Yu.V., Sadym A.V., Filimonov D.A., Blinova V.G., Dmitriev A.V., Poroikov V.V. Predicting Biotransformation Potential from Molecular Structure. // Journal of Chemical Information and Computer Sciences. - 2003. - V. 43. № 5. - P. 1636-1646.
4. Borodina Yu.V., Rudik A.V., Filimonov D.A., Kharchevnikova N.V., Dmitriev A.V., Blinova V.G., Poroikov, V.V. A New Statistical Approach to Predicting Aromatic Hydroxylation Sites. Comparison with Model-Based Approaches. // Journal of Chemical Information and Computer Sciences. - 2004. - V. 44 № 6. - P. 1998-2009.
5. Dmitriev A.V., Rudik A.V., Borodina Yu.V., Filimonov D.A., Poroikov V.V., Blinova V.G., Kharchevnikova N.V. Computer predicting of aromatic hydroxylation site for drug-like compounds. // Abstr. The 15th European Symposium on Quantitative Structure-Activity Relationships & Molecular Modelling (Euro QSAR 2004). - Istanbul (Turkey). - 2004. - P. 104.
6. Бородина Ю.В., Рудик А.В., Филимонов Д.А., Блинова В.Г., Дмитриев А.В., Харчевникова Н.В., Пороиков В.В. Компьютерный прогноз биотрансформации ксенобиотиков. // Сборник материалов XII Международной конференции и дискуссионного научного клуба «Новые информационные технологии в медицине, биологии, фармакологии и экологии». - Гурзуф (Украина). - 2004. - С. 77-79.
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. Дмитриев А.В., Харчевникова Н.В., Рудик А.В., Пороиков В.В. Разработка метода для предсказания региоселективности изоформы 3А4 цитохрома Р450. // Сборник материалов XII Российского национального конгресса «Человек и лекарство». - Москва (Россия). - 2005. - С. 750.
9. Lagunin A.A., Zaharov A.V., Dmitriev A.V., Filts O.A., Poroikov V.V. Database of drug-drug interactions. // The 3<sup>rd</sup> International Conference «Genomics, Proteomics, Bioinformatics and Nanotechnologies for Medicine». - Novosibirsk (Russia). - 2006. - P. 111.

10. Дмитриев А.В., Рудик А.В., Блинова В.Г., Филимонов Д.А., Поройков В.В. Предсказание классов биотрансформации ксенобиотиков на основе базы фрагментов. // Сборник материалов XIII Российского национального конгресса «Человек и лекарство». - Москва (Россия). - 2006. - С. 14.
11. Харчевникова Н.В., Дмитриев А.В., Дьячков П.Н., Поройков В.В. Прогноз сайта N-деалкилирования под действием цитохрома P450 на основе анализа топологических признаков и квантово химических расчетов. // Сборник материалов XIII Российского национального конгресса «Человек и лекарство». - Москва (Россия). - 2006. - С. 39.
12. Dmitriev A.V., Lagunin A.A., Zakharov A.V., Filz O.A., Filimonov D.A., Poroikov V.V. Computer prediction of substrates, inhibitors and inducers of cytochrome P450. // Abstr. 4<sup>th</sup> International Symposium «Computational Methods in Toxicology and Pharmacology Integrating Internet Resources (СМТПИ-2007)». - Moscow (Russia) - 2007. - P. 91.
13. D'yachkov P.N., Kharchevnikova N.V., Dmitriev A.V., Kuznetsov AV., Poroikov V.V. Quantum Chemical Simulation of Cytochrome P450 Catalyzed Aromatic Oxidation: Metabolism, Toxicity, and Biodegradation of Benzene Derivatives. // International Journal of Quantum Chemistry. - 2007. - V. 107 № 13. - P. 2454-2478.
14. Дмитриев А.В., Лагунин А.А., Захаров А.В., Фильц О.А., Филимонов Д.А., Поройков В.В. Компьютерное предсказание ферментной специфичности, ингибирующей и индуцирующей активностей молекул по отношению к цитохромам P450. // Тез. докл. IV Московского международного конгресса «Биотехнология: состояние и перспективы развития». - 2007. - Москва (Россия). С. 401.
15. Лагунин А.А., Филимонов Д.А., Захаров А.В., Фильц О.А., Дмитриев А.В., Поройков В.В. Компьютерный прогноз взаимодействия лекарств. // Сборник материалов XIV Российского национального конгресса «Человек и лекарство». - Москва (Россия). - 2007. - С. 297.
16. Дмитриев А.В., Рудик А.В., Блинова В.Г., Филимонов Д.А., Поройков В.В. Предсказание сайтов биотрансформации окисления с использованием программы METAPREDICT и набора фрагментов. // Сборник материалов XIV Российского национального конгресса «Человек и лекарство». - Москва (Россия). - 2007. - С. 283.
17. Дмитриев А.В., Лагунин А.А., Рудик А.В., Филимонов Д.А., Поройков В.В. Прогноз взаимодействия ксенобиотиков с основными изоформами цитохрома P450 человека. // Сборник материалов XV Российского национального конгресса «Человек и лекарство». - Москва (Россия). - 2008. - С. 521.
18. Dmitriev A.V., Filimonov D.A., Lagunin A.A., Poroikov V.V. Tools for prediction of xenobiotics interaction with human cytochrome P450. // Abstr. 4<sup>th</sup> German conference on chemoinformatics. - Goslar (Germany). - 2008. - P. 61.