

## КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ВЗАИМОСВЯЗИ «СТРУКТУРА–АКТИВНОСТЬ» ИНГИБИТОРОВ ЦИКЛООКСИГЕНАЗЫ-2 СРЕДИ ПРОИЗВОДНЫХ ТЕТРАГИДРО-2Н-ИЗОИНДОЛА

© 2015 В.Р. Хайруллина<sup>1\*</sup>, А.Я. Герчиков<sup>1\*</sup>,  
А.А. Лагунин<sup>2,3</sup>, Ф.С. Зарудий<sup>4</sup>

<sup>1</sup> Башкирский государственный университет, химический факультет,  
450076 Уфа, ул. Заки Валиди, 32; факс: +7(347)229-9707,  
электронная почта: gerchikov@inbox.ru, Veronika1979@yandex.ru

<sup>2</sup> Российский национальный исследовательский медицинский университет  
им. Н.И. Пирогова, 117997 Москва, ул. Островитянова, 1;  
факс: +7(495)434-1422

<sup>3</sup> НИИ биомедицинской химии им. В.Н. Ореховича, 119121 Москва,  
ул. Погодинская, 10, стр. 8; факс: +7(499)245-0857,  
электронная почта: alexey.lagunin@ibmc.msk.ru

<sup>4</sup> Башкирский государственный медицинский университет,  
450000 Уфа, ул. Ленина, 3; факс: +7(347)272-3751,  
электронная почта: zarudii.f@yandex.ru

Поступила в редакцию 07.05.14  
После доработки 10.07.14

С использованием программы GUSAR выполнен количественный анализ взаимосвязи между структурой и эффективностью ингибирования каталитической активности циклооксигеназы-2 двадцатью шестью производными 4,5,6,7-тетрагидро-2Н-изоиндола, 2,3-дигидро-1Н-пирролизина и бензотиофена в концентрационном интервале значений 0,6–700,0 нмоль/л величины  $IC_{50}$ . На основе MNA- и QNA-дескрипторов, а также их сочетания построено шесть статистически значимых консенсус-моделей прогноза численных значений  $IC_{50}$ , характеризующихся высокой точностью прогноза для структур обучающих и тестовой выборок. С помощью той же программы определены фрагменты структур ингибиторов ЦОГ-2, способные усиливать и ослаблять целевое свойство. Эта информация может быть учтена при молекулярном дизайне новых ингибиторов ЦОГ-2. Показано, что в большинстве случаев выявленное с использованием GUSAR влияние фрагментов структур на ингибирующую активность исследованных соединений совпадает с результатами экспертной оценки их влияния на основе известных экспериментальных данных, что может быть использовано для оптимизации структур с целью изменения величины их биологической активности.

**КЛЮЧЕВЫЕ СЛОВА:** ингибиторы циклооксигеназы-2, QSAR-модели, GUSAR, QNA- и MNA-дескрипторы, анализ взаимосвязи «структура–активность».

Циклооксигеназа представляет собой гемсодержащий фермент, катализирующий реакцию превращения арахидоновой кислоты в простагландин H<sub>2</sub> [1–7]. Фермент содержит два активных центра: 1) циклооксигеназный сайт, в котором арахидоновая кислота превращается в простагландин G<sub>2</sub>; 2) гем, обладающий пероксидаз-

ной активностью и способствующий превращению простагландина G<sub>2</sub> в простагландин H<sub>2</sub> [1, 2, 4, 5].

Существуют две гомологичные друг другу на 60% изоформы циклооксигеназы: циклооксигеназа-1 (ЦОГ-1) и циклооксигеназа-2 (ЦОГ-2) [1, 2, 4, 5]. ЦОГ-1 является конститутивным фер-

Принятые сокращения: ЦОГ-1 – циклооксигеназа-1, ЦОГ-2 – циклооксигеназа-2, QSAR (Quantitative Structure–Activity Relationships) – количественная взаимосвязь «структура–активность», GUSAR (General Unrestricted Structure–Activity Relationships) – общая универсальная взаимосвязь «структура–активность», MNA (Multilevel Neighbourhoods of Atoms) – многоуровневые атомные окрестности, QNA (Quantitative Neighborhoods of Atoms) – количественные атомные окрестности,  $IC_{50}$  (half inhibition concentration) – концентрация полумаксимального ингибирования,  $pIC_{50}$  – отрицательный десятичный логарифм величины  $IC_{50}$ .

\* Адресат для корреспонденции.

ментом и, соответственно, присутствует в тканях практически повсеместно [1, 2, 4, 5]. В тромбоцитах она обеспечивает превращение арахидоновой кислоты в тромбоксан [1, 4, 5]. Ингибирование каталитической активности ЦОГ-1 при приеме неселективных нестероидных противовоспалительных препаратов приводит к повреждению стенки желудка и развитию язв [1, 3–7]. ЦОГ-2 в обычных условиях присутствует в мозге и корковом слое почек [1, 2, 4, 5]. В других тканях ЦОГ-2 индуцируется при воспалении [2, 5]. Имеются экспериментальные доказательства участия ЦОГ-2 в канцерогенезе кишечника и молочных желез, полученные на животных моделях с применением неселективных и селективных ингибиторов ЦОГ-2 [2, 6, 7]. В частности повышенный уровень ЦОГ-2 обнаруживается более чем у 50% пациентов со злокачественной опухолью молочной железы, простаты и т.д. [6, 7]. В этой связи поиск эффективных и вместе с тем селективных ингибиторов ЦОГ-2 является актуальной задачей медицинской химии и фармакологии.

В настоящее время накоплен значительный объем экспериментальных данных по эффективности ингибирования изоформ ЦОГ разными классами биологически активных соединений [1–7]. Это позволяет использовать методы виртуального скрининга, базирующиеся на анализе взаимосвязей «структура–активность» – (Q)SAR (Quantitative Structure–Activity Relationships) для целенаправленного поиска новых природных и синтетических гетероциклических соединений с выраженной селективностью в отношении ЦОГ-2. Методы, реализующие подходы QSAR, позволяют уже на досинтетической стадии создания потенциальных лекарственных соединений количественно предсказывать их биологическую активность [8–13]. Актуальность данных исследований обусловлена тем, что в арсенале современной фармакологии находится несколько тысяч разных видов биологической активности [12, 13]. В этих условиях экспериментальное изучение фармакологического профиля новых потенциальных лекарственных средств в различных модельных системах *in vitro* и *in vivo* без привлечения методов компьютерной химии представляет собой задачу, требующую значительных временных и финансовых затрат [12]. Другой не менее актуальной задачей является целенаправленный молекулярный дизайн новых потенциальных ингибиторов ЦОГ-2 на основе их структурных аналогов. Рациональным путем ее решения является структурный анализ биологически активных соединений для определения в них структурных фрагментов, снижающих или повышающих био-

логическую активность. Создание QSAR-моделей, описывающих и визуализирующих вклад различных частей исследуемых молекул в изменение величины биологической активности, может стать основой для оптимизации и рационального конструирования новых биологически активных соединений, в т.ч. ингибиторов ЦОГ-2. Традиционно в таких работах используют методы молекулярного моделирования и 3D-QSAR [12, 13].

Целью настоящей работы было построение и валидация QSAR-моделей селективных ингибиторов ЦОГ-2 для производных тетрагидро-2H-изоиндола, 2,3-дигидро-1H-пирролизина и бензотиофена на основе двухмерного представления их структурной формулы, а также анализ влияния атомов и функциональных групп на эффективность ингибирования каталитической активности ЦОГ-2 с использованием созданных QSAR-моделей.

## МЕТОДЫ ИССЛЕДОВАНИЯ

Количественный анализ взаимосвязей «структура–активность» ингибиторов ЦОГ-2 проводили с использованием компьютерной программы GUSAR (General Unrestricted Structure–Activity Relationships) [13–18]. Подходы, реализованные в программе GUSAR, являются достаточно новыми в QSAR-моделировании. Они сочетают в себе идеологию как SAR-, так и традиционных 2D-QSAR-методов. В связи с этим для объективного понимания результатов нижеописанных исследований следует дать краткое описание возможностей данной программы и идеологию построения в ней количественных взаимосвязей «структура–активность».

**Краткое описание программы GUSAR.** В GUSAR для построения (Q)SAR-моделей используется метод самосогласованной регрессии [16]. Описание структуры и вычисление регрессионных коэффициентов с целью дальнейшего построения QSAR-моделей для прогноза количественных величин основано на использовании двух типов подструктурных дескрипторов атомных окрестностей: MNA (Multilevel Neighborhoods of Atoms) и QNA (Quantitative Neighborhoods of Atoms) [15, 16]. Их расчет осуществляется программой GUSAR автоматически из структурных формул химических соединений с учетом валентности и парциальных зарядов содержащихся в них атомов, но без указания специфичности типов связи. MNA-дескрипторы генерируются на основе структурной формулы химических соединений без использования какого-либо заранее составленного списка структур-

ных фрагментов [11–13]. На основе MNA-дескрипторов с использованием В-статистики, рассчитанной по алгоритму, реализованному в программе PASS (Prediction of Activity Spectra for Substances), осуществляется прогноз спектра биологической активности химического соединения [11, 15, 16]. Результаты прогноза, представленные в виде списка видов биологической активности с оценкой вероятности их проявления, являются переменными для расчета регрессионных коэффициентов. Регрессионное уравнение, построенное на основе MNA-дескрипторов, отражает специфику действия соединения, но не отражает в явном виде физико-химические параметры химических соединений [15, 16].

Для вычисления QNA-дескрипторов используются значения потенциала ионизации и сродства к электрону каждого атома молекулы. Они рассчитываются с учетом связей между всеми атомами структуры. Таким образом, QNA-дескрипторы с одной стороны описывают каждый атом молекулы, а с другой зависят от строения молекулы в целом [16]. Значения QNA-дескрипторов являются базовой информацией для расчета 2D полиномов Чебышева, которые в дальнейшем используются в качестве переменных для построения регрессионного уравнения, учитывающего как специфику, так и физико-химические свойства каждого атома, входящего в обучающую выборку [16]. Следует отметить, что программа способна строить QSAR-модели как опираясь исключительно на один из этих типов дескрипторов, так и на их сочетание в консенсусном подходе [15, 16]. При этом, исходя из идеологии консенсусного подхода, модели количественного прогноза биологической активности по каждому из данных типов дескрипторов рассчитываются независимо друг от друга. Примеры готовых QSAR-моделей GUSAR по предсказанию токсических эффектов химических соединений, которыми можно свободно воспользоваться, представлены в интернете на веб-сайте: <http://www.way2drug.com/GUSAR>.

Программа GUSAR позволяет создавать QSAR-модели, а также визуализировать вклад каждого атома в прогнозируемую величину [13, 15–18]. Такая возможность реализована в QSAR-моделях, построенных на основе QNA-дескрипторов, и, соответственно, в консенсусном сочетании построенных разными способами QSAR-моделей. Она позволяет определять «сильные» и «слабые» места в молекулах биологически активных веществ и, следовательно, делать объективные выводы о замене тех или иных фрагментов при молекулярном дизайне, направленном либо на усиление, либо на ослабление целевого свойства. Следует отметить, что

подобная задача может быть успешно решена и с помощью других компьютерных программ и систем, например, с помощью компьютерной системы SARD-21 [19–21]. Однако в отличие от аналогичной возможности, реализованной в компьютерной системе SARD-21, процедура оценки вклада атомов в целевую активность в программе GUSAR может осуществляться на небольшой по числу структур обучающей выборке (от 20 структур). При этом не требуется наличие противокласса соединений – т.е. альтернативной обучающей выборки, все структуры в которой не обладают заданным свойством либо обладают им в незначительной степени. По умолчанию атомы, не влияющие на активность молекулы, изображаются зеленым цветом. Атомы, усиливающие либо ослабляющие изучаемую активность, изображаются красным и синим цветом соответственно [13, 15–18].

Построение QSAR-моделей с использованием программы GUSAR проводили в несколько этапов.

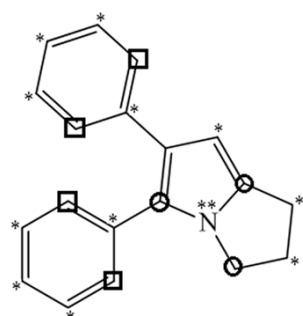
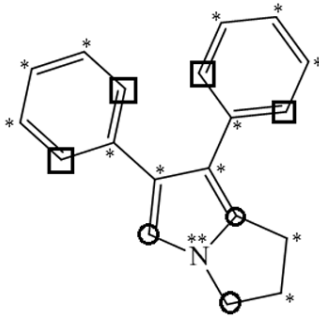
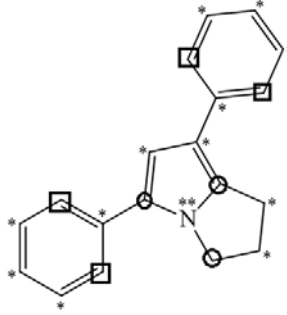
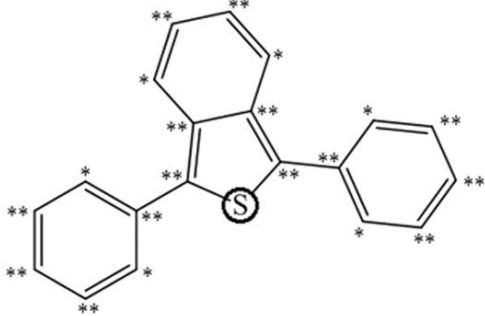
#### **Формирование обучающих и тестовой выборок.**

Структуры соединений обучающей и тестовой выборок создавались в программе MarvinSketch 5.9.1. [22] и с использованием программы Discovery Studio Visualiser [23] переводились в SDF-формат.

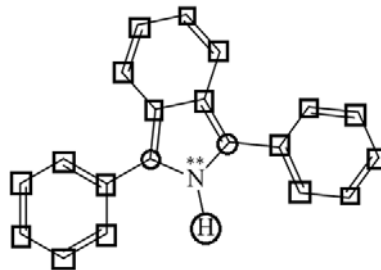
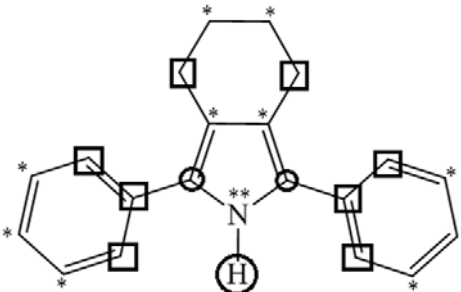
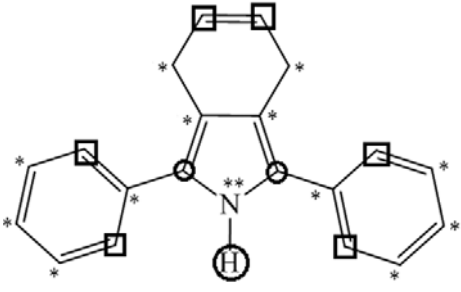
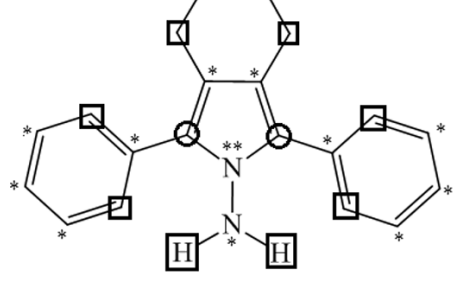
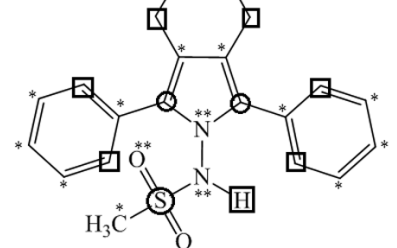
Обучающая выборка (OB1) сформирована на основе 26 соединений, изученных в качестве ингибиторов ЦОГ-2 в работе Портевин с соавт. [24] (табл. 1).

Обучающая и тестовые выборки OB2 и ТВ включают в себя 20 и 6 структур ингибиторов ЦОГ-2 соответственно. Они получены в результате разделения предварительно отсортированной по возрастанию значений  $IC_{50}$  OB1 в соотношении  $\sim 3 : 1$ , т.е. из OB1 исключали каждое четвертое соединение. Как видно из табл. 1, структуры соединений выборок OB1, OB2 и ТВ характеризуются достаточно широким диапазоном величин 50%-ной ингибирующей концентрации ( $IC_{50}$ ). Ингибирующая активность этих соединений измерена методом связывания по степени снижения каталитической активности ЦОГ-2 резидентных макрофагов у самцов белых мышей и представлена в виде количественного параметра  $IC_{50}$ , выраженного в величинах нмоль/л. Отобранные соединения являются производными 4,5,6,7-тетрагидро-2H-изоиндола, 2,3-дигидро-1H-пирролизина и бензотиофена. Они характеризуются наличием в своей структуре двух бензольных заместителей, связанных с гетероциклическими ароматическими фрагментами. По способности ингибировать каталитическую активность ЦОГ-2 эти соединения лежат в интервале 0,6–700,0 нмоль/л. Для построения

**Таблица 1.** Структуры ингибиторов ЦОГ-2, вошедшие в обучающие и тестовую выборки ОВ1, ОВ2, ТВ с соответствующими им величинами IC<sub>50</sub>, а также влияние вклада атомов в ингибирующую активность, оцененное по консенсус-модели М5

№	Шифр вещества по статье Портевин с соавт. [24]	Структурная формула соединений с соответствующими вкладами атомов в активность	ОВ1	ОВ2	ТМ	IC <sub>50</sub> , нМ [24]*
1	2	3	4	5	6	7
1	3		+	+	-	10,0
2	4		+	+	-	28,7
3	5		+	+	-	50,0
4	6		+	+	-	10,9

Продолжение табл. 1

1	2	3	4	5	6	7
5	7		+	+	-	10,0
6	11		+	+	-	1,5
7	12		+	+	-	3,3
8	14		+	-	+	1,8
9	15		+	+	-	500,0

Продолжение табл. 1

1	2	3	4	5	6	7
10	17	<p>Chemical structure of a cyclooxygenase-2 inhibitor. It features a central piperidine ring with a hydrogen atom on the nitrogen. Two phenyl rings are attached to the piperidine ring at the 2 and 6 positions. Each phenyl ring has a fluorine atom at the 2-position. The piperidine ring has asterisks at the 2, 3, 4, and 5 positions. The phenyl rings have asterisks at the 1, 3, and 4 positions and double asterisks at the 2 and 6 positions.</p>	+	+	-	1,7
11	18	<p>Chemical structure of a cyclooxygenase-2 inhibitor. It features a central piperidine ring with a hydrogen atom on the nitrogen. Two phenyl rings are attached to the piperidine ring at the 2 and 6 positions. Each phenyl ring has a methylsulfanyl group (-S-CH<sub>3</sub>) at the 2-position. The piperidine ring has asterisks at the 2, 3, 4, and 5 positions. The phenyl rings have asterisks at the 1, 3, and 4 positions and double asterisks at the 2 and 6 positions.</p>	+	-	+	500,0
12	19	<p>Chemical structure of a cyclooxygenase-2 inhibitor. It features a central piperidine ring with a hydrogen atom on the nitrogen. Two phenyl rings are attached to the piperidine ring at the 2 and 6 positions. Each phenyl ring has a methyl group (-CH<sub>3</sub>) at the 2-position. The piperidine ring has asterisks at the 2, 3, 4, and 5 positions. The phenyl rings have asterisks at the 1, 3, and 4 positions and double asterisks at the 2 and 6 positions.</p>	+	+	-	16,7
13	20	<p>Chemical structure of a cyclooxygenase-2 inhibitor. It features a central piperidine ring with a hydrogen atom on the nitrogen. Two phenyl rings are attached to the piperidine ring at the 2 and 6 positions. Each phenyl ring has a methoxy group (-O-CH<sub>3</sub>) at the 2-position. The piperidine ring has asterisks at the 2, 3, 4, and 5 positions. The phenyl rings have asterisks at the 1, 3, and 4 positions and double asterisks at the 2 and 6 positions.</p>	+	+	-	21,3
14	21	<p>Chemical structure of a cyclooxygenase-2 inhibitor. It features a central piperidine ring with a hydrogen atom on the nitrogen. Two phenyl rings are attached to the piperidine ring at the 2 and 6 positions. Each phenyl ring has a chlorine atom (-Cl) at the 2-position. The piperidine ring has asterisks at the 2, 3, 4, and 5 positions. The phenyl rings have asterisks at the 1, 3, and 4 positions and double asterisks at the 2 and 6 positions.</p>	+	+	-	5,0

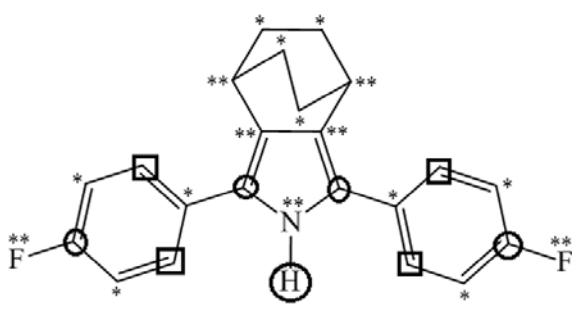
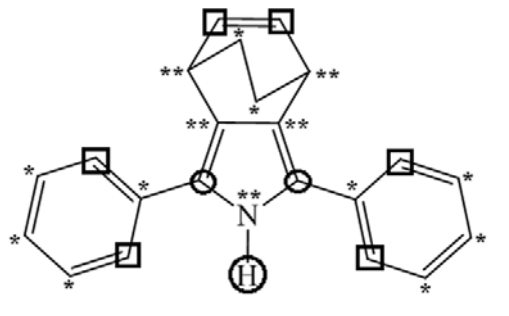
Продолжение табл. 1

1	2	3	4	5	6	7
15	22		+	-	+	42,0
16	24		+	+	-	3,1
17	25		+	-	+	14,5
18	27		+	+	-	0,7
19	28		+	+	-	2,9

Продолжение табл. 1

1	2	3	4	5	6	7
20	31		+	+	-	2,6
21	32		+	-	+	1,1
22	33		+	-	+	4,5
23	34		+	+	-	700,0
24	36		+	+	-	1,6



1	2	3	4	5	6	7
25	37		+	+	–	0,6
26	38		+	+	–	35,6

Примечание. В формулах квадратами выделены атомы, не влияющие на активность ингибиторов ЦОГ-2; окружностями отмечены атомы, повышающие активность ингибиторов ЦОГ-2; звездочкой отмечены атомы, снижающие активность, количество звездочек соответствует степени негативного влияния рассматриваемого атома в целевое свойство.

\* Экспериментальные данные получены Портевин с соавт. [24] методом связывания по степени снижения каталитической активности ЦОГ-2 резидентных макрофагов у самцов белых мышей.

QSAR-моделей данные  $IC_{50}$  в нмоль/л переводили в  $IC_{50}$  в моль/л, которые затем были трансформированы в величины  $pIC_{50}$  по формуле:

$$pIC_{50} = -\log_{10}(IC_{50}).$$

Диапазон активности, в котором находятся соединения обучающих выборок OB1, OB2 и тестовой выборки ТВ, в единицах  $pIC_{50}$  составил от  $-9,22$  до  $-6,15$  для OB1, от  $-9,22$  до  $-6,15$  для OB2 и от  $-8,96$  до  $-6,30$  для ТВ соответственно.

**Построение QSAR-моделей.** Построение QSAR-моделей количественного прогноза ингибирующей активности в отношении ЦОГ-2 проводили на основе MNA- и QNA-дескрипторов. Итоговые регрессионные модели представляли собой консенсус QSAR-моделей, построенных независимо друг от друга по каждому из этих типов дескрипторов и характеризующихся высокой прогностической способностью.

**Оценка качества и прогностической способности QSAR-моделей.** Для оценки прогностической способности моделей M2, M4, M6, пост-

роенных на основе обучающей выборки OB2, использовали внутреннюю тестовую выборку (ТВ). Прогностическую способность моделей M1, M3 и M5 оценивали только по результатам прогноза численных значений активности обучающей выборки OB1. В качестве внутренней валидации использовался скользящий контроль со случайным двадцатикратным исключением 20% обучающей выборки. Остальные параметры GUSAR использовались по умолчанию. Всего создавалось 360 моделей (по 180 для каждого типа дескрипторов).

**Оценка вклада атомов в целевую активность.** Вклад атомов в активность ингибиторов ЦОГ-2 оценивали по консенсус-модели M5, содержащей 26 ингибиторов ЦОГ-2. Как отмечалось выше, эта процедура реализуется в программе GUSAR автоматически при построении QSAR-моделей на основе QNA-дескрипторов и консенсус-моделей. Для удобства восприятия иллюстративного материала результаты анализа вклада атомов в целевую активность, представленные в QSAR-моделях в различной цветовой

гамме, переводили в графические символы. Квадратами отмечали атомы, не влияющие на активность ингибиторов ЦОГ-2, окружностями – атомы, повышающие активность ингибиторов ЦОГ-2, звездочками выделяли атомы, снижающие активность. Количество звездочек соответствовало степени отрицательного вклада рассматриваемого атома в целевое свойство.

### РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

С использованием консенсусного подхода (когда результаты прогноза усредняются по нескольким моделям), реализованного в программе GUSAR, проведено моделирование количественной взаимосвязи между строением и эффективностью ингибирования каталитической активности ЦОГ-2 производными 4,5,6,7-тетрагидро-2H-изоиндола, 2,3-дигидро-1H-пирролизина и бензотиофена, содержащимися в обучающих выборках OB1 и OB2. В итоге, в зависимости от типа использованных при расчетах дескрипторов, для каждой из обучающих выборок получено по три консенсус-модели (табл. 2). Проведена оценка прогностической способности итоговых регрессионных уравнений консенсус-моделей M1–M6 на структурах обучающих выборок OB1 и OB2 соответственно при скользящем контроле с исключением 20% соединений. Дополнительно выполнена независимая проверка прогностической способности моделей M2, M4 и M6 путем тестирования по ней шести производных изоиндола, содержащихся в

тестовой выборке ТВ и не вошедших в обучающую выборку OB1. Статистические параметры консенсус-моделей, а также характеристики точности предсказанных по ним значений  $pIC_{50}$  для ингибиторов ЦОГ-2 представлены в табл. 2.

Как следует из результатов, представленных в табл. 2, для построения QSAR-моделей с приемлемыми статистическими характеристиками ( $R^2 > 0,6$ ,  $Q^2 > 0,5$ ) [15–18] можно использовать консенсус-модели, объединяющие QSAR-модели, построенные как на одном определенном типе, так и на обоих типах дескрипторов (QNA или MNA). При этом незначительная разница между статистическими характеристиками моделей прогноза параметров  $pIC_{50}$  для OB1 и OB2, построенных с использованием разных типов дескрипторов, свидетельствует об устойчивости создания данных моделей (табл. 2). Создание QSAR-моделей на основе MNA-дескрипторов позволяет получить консенсус-модели M1 и M2 с высокими значениями  $R^2$  и  $Q^2$ . Однако предсказательная способность модели M2 на соединениях тестовой выборки TM, которые по структуре значительно схожи с соединениями обучающей выборки OB2, характеризуется невысокой точностью.

Консенсус-модель M4, построенная на основе QNA-дескрипторов, характеризуется более низким качеством прогноза соединений обучающей выборки OB1 (величины  $R_{OB}^2$  и  $Q^2$ ) по сравнению с моделью M1. Вместе с тем она демонстрирует высокую предсказательную способность на структурах тестовой выборки ТВ ( $R_{ТВ}^2$ ) (табл. 2).

**Таблица 2.** Статистические характеристики и оценка показателей точности предсказаний значений  $pIC_{50}$  для ингибиторов ЦОГ-2 по консенсус-моделям M1–M6. Структуры OB1 и OB2 лежат в диапазоне активности  $pIC_{50} = 6,15–9,22$

Обучающая выборка	Модель	N	$R_{OB}^2$	$Q^2$	$R_{ТВ}^2$	F	S.D.	V
QSAR-модели, построенные на основе MNA-дескрипторов								
OB1	M1	26	0,865	0,818	–	19,626	0,320	5
OB2	M2	20	0,885	0,833	0,421	16,90	0,290	4
QSAR-модели, построенные на основе QNA-дескрипторов								
OB1	M3	26	0,826	0,731	–	14,996	0,357	5
OB2	M4	20	0,779	0,654	0,824	11,293	0,390	4
QSAR-модели, построенные на основе QNA- и MNA-дескрипторов								
OB1	M5	26	0,889	0,837	–	15,301	0,316	5
OB2	M6	20	0,874	0,802	0,706	11,022	0,334	4

Примечание. N – число структур в обучающей выборке;  $R_{OB}^2$  – коэффициент детерминации, рассчитанный для соединений из обучающей выборки;  $R_{ТВ}^2$  – коэффициент детерминации, рассчитанный для соединений из тестовой выборки;  $Q^2$  – коэффициент корреляции, рассчитанный на обучающей выборке при скользящем контроле с исключением по одному; F – критерий Фишера; S.D. – стандартное отклонение; V – число переменных в конечном регрессионном уравнении.

Как видно из данных табл. 2, применение консенсусного подхода к построению регрессионных уравнений, основанного на сочетании MNA- и QNA-дескрипторов, позволило получить модели M5 и M6, характеризующиеся высокими статистическими показателями. В частности модель M6 отличается достаточно высокой прогностической способностью как на соединениях обучающей выборки OB2, так и на близких к нему по структуре соединениях тестовой выборки ТВ. В пользу данного заключения свидетельствуют результаты, предсказанные консенсус-моделью, в сравнении с экспериментальными величинами, которые приведены на рисунке.

Результаты предсказания находятся в хорошем соответствии с литературными данными. Известно, что использование консенсус-модели, которая усредняет результаты предсказания отдельных моделей, позволяет снизить вариабельность результатов предсказания отдельных моделей и таким образом получить более точный результат [25]. Кроме того, применение консенсусного подхода к генерации моделей позволяет получать более объективную информацию о влиянии различных атомов в исследуемый тип активности. В этой связи для дальнейшего анализа были использованы полученные на базе консенсусного подхода модели M5 и M6.

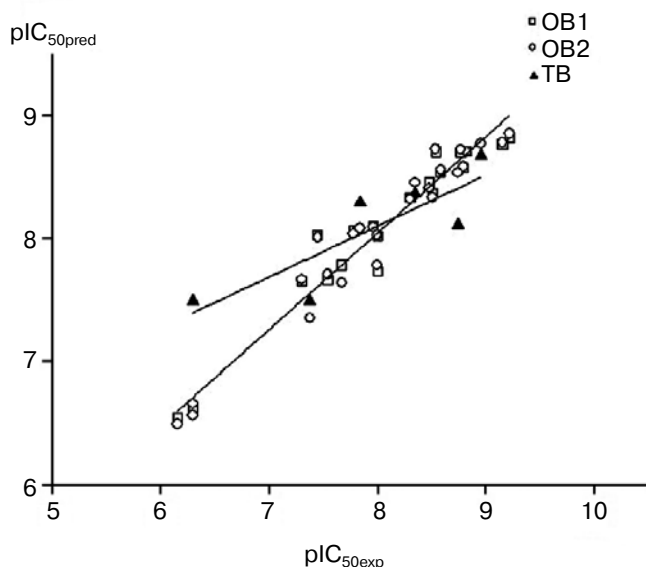
Далее с помощью программы GUSAR был проанализирован вклад атомов и содержащие их функциональные группы в соединениях выборки OB1 в величину ингибирования ими

ЦОГ-2. Анализ выполнен по модели M5, содержащей 26 ингибиторов ЦОГ-2. В табл. 1 показаны вклады атомов в прогнозируемую величину активности.

В частности установлено, что электронодонорные и электроноакцепторные заместители неоднозначно влияют на эффективность связывания производных изоиндола с активным центром ЦОГ-2. Можно предположить, что природа данного явления объясняется разным механизмом взаимодействия исследуемых соединений с активным центром ЦОГ-2. К подобному выводу пришли и Силакари с соавт. [26], на основании экспериментальных данных которых было выполнено данное теоретическое исследование. Влияние атомов галогенов на эффективность ингибиторов ЦОГ-2 неоднозначно. Из сравнения структур S11, S17, S22, S27, S28, S33, S34 видно, что атом фтора, связанный с 1,4-дизамещенным бензолом, ослабляет активность содержащих его соединений. Исключением являются соединения с кодами S32 и S37 (табл. 1), в которых замена атома водорода на атом фтора в *para*-положении бензольного фрагмента приводит к прямо противоположному эффекту (табл. 1). Введение в *para*-положение бензольных фрагментов и атомов фтора и сульфонового фрагмента в сочетании с метильной группой приводит к значительному падению активности соединения S34, что четко видно при сравнении ингибирующего эффекта соединений S31 и S34. Эти выводы подтверждаются и результатами визуального анализа.

Расхождение в интерпретации экспериментальных данных с выводами программы GUSAR применительно к соединениям S32 и S37 можно объяснить иным механизмом их взаимодействия с активным центром ЦОГ-2 по сравнению с другими изученными ингибиторами ЦОГ-2. Известно, что зачастую даже незначительная модификация структур биологически активных веществ может приводить к существенному изменению их активности вплоть до изменения механизма их взаимодействия с активным центром исследуемого фермента. Эти явления не могут быть учтены ни одним из классических методов QSAR-моделирования, в т.ч. и в подходах, реализованных в GUSAR.

Сопоставительный анализ структур S18, S20, S34, а также структур S14 и S15 показал, что метильная группа в сочетании с сульфогруппой, атомом кислорода или атомом серы также негативно влияет на целевое свойство (табл. 1). В то же время атом кислорода в составе сульфонной и метоксильной групп будет снижать активность ингибиторов ЦОГ-2, в то время как атом серы в составе тех же групп, а также в тиоме-



Сопоставление экспериментальных ( $pIC_{50exp}$ ) значений  $pIC_{50}$  с предсказанными ( $pIC_{50pred}$ ) по моделям M5 и M6 для ингибиторов ЦОГ-2, содержащихся в обучающих выборках OB1, OB2 и тестовой выборке ТВ

тильной группе в структурах ингибиторов ЦОГ-2 способствует повышению активности (табл. 1). Аналогичный эффект оказывает атом серы в составе гетероциклических ароматических фрагментов в структуре S18 (табл. 1). Моно- и дизамещенный бензол, а также фрагменты 3,5-диметил-4-азатрицикло[5.2.1.0]дека-2,5-диена и 3,5-диметил-4-азатрицикло[5.2.2.0]ундека-2,5-диена оказывают слабое негативное влияние в изучаемый тип активности во всех соединениях. Вторичный атом азота, входящий в гетероциклические фрагменты, также снижает ингибирующую активность содержащих его соединений, в то время как атом водорода в данной функциональной группе действует противоположным образом на активность соединений. Третичный атом азота, входящий в 2,3-дигидро-1Н-пирролизинный фрагмент, снижает эффективность ингибирования ЦОГ-2. Первичная аминогруппа NH<sub>2</sub> негативно влияет на целевое свойство в структуре S14 (табл. 1).

Электронодонорная метильная и электроноакцепторная метоксильная группы, а также имидазольный цикл, находящиеся в *para*-положении бензольных фрагментов, на порядок снижают активность соединений S19 и S20 по сравнению с S11 (табл. 1). Введение электроноакцепторной тиометильной группы в *para*-положение бензольных фрагментов соединения S34 приводит к значительному падению ингибирующей активности практически на два порядка (табл. 1). Сравнение структур соединений S12 и S25 позволяет заключить, что введение метильных заместителей в положения 5 и 6 4,7-дигидро-2Н-изоиндольного фрагмента в 4 раза снижает активность ингибиторов ЦОГ-2 (табл. 1). Таким образом, выводы о влиянии структурных фрагментов на активность ингибиторов ЦОГ-2, полученные на основе визуального анализа с использованием программы GUSAR, не противоречат результатам, основанным на экспериментальных данных.

Для объяснения неоднозначного влияния электронодонорных и электроноакцепторных групп в активность ингибиторов ЦОГ-2 нами был выполнен анализ численных значений констант *para*-заместителей Брауна [27] для ингибиторов ЦОГ-2 (табл. 3).

Как видно из табл. 3, функциональные группы с выраженными нуклеофильными свойствами снижают активность ингибиторов ЦОГ-2. Кроме того, замещение атомов водорода в *para*-положениях бензольных фрагментов на функциональные группы, не способные к образованию водородных связей с активным центром ЦОГ-2, также негативно влияет на ингибирующую активность по отношению к ЦОГ-2. Предложенные нами выводы относительно влияния нуклеофильности и способности функциональных групп к образованию водородных связей полностью согласуются с выводами, сделанными Силакари с соавт. [26], которые на основе четырех типов дескрипторов (электронные, структурные, термодинамические, пространственные), представленных в табл. 4, выполнили QSAR-моделирование 25 соединений, вошедших в сформированные нами обучающие выборки OB1 и OB2.

Следует отметить, что построенные нами модели прогноза численных значений IC<sub>50</sub> для ингибиторов ЦОГ-2 не уступают по точности шести QSAR-моделям, полученным Силакари с соавт. [26] с использованием генетического алгоритма формирования регрессионных уравнений, о чем свидетельствуют результаты исследований, приведенные в табл. 5.

Таким образом, влияние природы атомов на ингибирующую эффективность производных изоиндола в отношении ЦОГ-2 неоднозначно и зачастую зависит от природы близлежащих структурных фрагментов соединения. Кроме того, при анализе вклада атомов в активность необходимо учитывать, в какую функциональ-

**Таблица 3.** Влияние констант *para*-заместителей Брауна в бензольном кольце соединений S11, S17–S21 на эффективность их ингибирующего действия в отношении ЦОГ-2

№	Шифр соединения в табл. 1	4-R	$\sigma_{para}^+$	IC <sub>50</sub> , нМ
1	11	H	0,000	1,5
2	21	Cl	0,114	5,0
3	17	F	-0,073	1,7
4	19	CH <sub>3</sub>	-0,311	16,7
5	18	SCH <sub>3</sub>	-0,604	500,0
6	20	OCH <sub>3</sub>	-0,778	21,3

**Таблица 4.** Перечень дескрипторов, использованных в подходе Ханша в работе Силакари с соавт. [26]

Тип дескриптора	Символ	Описание дескриптора
Электронный	$A_{pol}$	сумма поляризуемостей атомов
Электронный	НОМО	энергия высшей занятой молекулярной орбитали, эВ
Структурный	HBD	число доноров водородной связи
Термодинамический	$F_{H_2O}$	свободная энергия десольватации в воде, ккал/моль
Пространственный	$S_{xyf}$	проекция поверхности молекулы на плоскость, лежащую внутри прямоугольника, образованного осями OX и OY
Пространственный	$S_{xzf}$	проекция поверхности молекулы на плоскость, лежащую внутри прямоугольника, образованного осями OX и OZ

**Таблица 5.** Сравнение статистических параметров полученных нами моделей M5 и M6 с QSAR-моделями Силакари с соавт. [26]. Структуры выборок OB1 и OB2 лежат в диапазоне активности  $pIC_{50} = -9,22 \div -6,15$ 

Номер модели	N	R <sup>2</sup>	Q <sup>2</sup>	F	S.D.	V	Тип дескрипторов
M5	20	0,873	0,803	11,411	0,335	4	QNA и MNA
M6	26	0,887	0,834	15,451	0,318	5	QNA и MNA
M (1)*	25	0,757	—	32,673	0,581	1	$A_{pol}$
M (2)*	25	0,817	—	29,819	0,516	2	$A_{pol}$ , HBD
M (3)*	25	0,761	—	21,276	0,590	2	$A_{pol}$ , НОМО
M (4)*	25	0,804	—	27,334	0,535	2	$A_{pol}$ , $F_{H_2O}$
M (5)*	25	0,758	—	20,872	0,594	2	$A_{pol}$ , $S_{xyf}$
M (6)*	25	0,764	—	21,530	0,587	2	$A_{pol}$ , $S_{xzf}$

Примечание. N – число структур в обучающей выборке; R<sup>2</sup> – коэффициент детерминации, рассчитанный для соединений из обучающей выборки; Q<sup>2</sup> – коэффициент корреляции, рассчитанный на обучающей выборке при скользящем контроле с исключением по одному; F – критерий Фишера; S.D. – стандартное отклонение; V – число переменных в конечном регрессионном уравнении.

\* Данные Силакари с соавт. [26].

ную группу они входят. Полученные результаты могут быть учтены при молекулярном дизайне действующих веществ известных нестероидных противовоспалительных лекарственных препаратов с целью усиления эффективности их ингибирующего действия в отношении ЦОГ-2.

1) С помощью компьютерной программы GUSAR проведен количественный анализ взаимосвязей «структура–активность» ингибиторов ЦОГ-2, являющихся производными тетрагидро-2H-изоиндола, 2,3-дигидро-1H-пирролизина и бензотиофена.

2) На основе MNA- и QNA-дескрипторов, а также их консенсусного сочетания построено шесть QSAR-моделей для прогноза численных

значений  $pIC_{50}$ , характеризующихся хорошей предсказательной способностью на структурах обучающих и тестовой выборок.

3) Выявлены функциональные группы, модулирующие активность ингибиторов ЦОГ-2. Установлено, что *para*-заместители с выраженными нуклеофильными свойствами снижают активность ингибиторов ЦОГ-2.

4) Выполнен анализ влияния природы атомов на эффективность ингибиторов ЦОГ-2. В ходе структурного анализа ингибиторов ЦОГ-2, вошедших в модель M5, выявлены атомы, как снижающие, так и повышающие эффективность их ингибирующего действия в отношении ЦОГ-2. Эти данные могут быть использованы для оптимизации структур биологически актив-

ных веществ с целью усиления эффективности ингибирования ими ЦОГ-2.

5) Показано, что в большинстве случаев зависимости влияния атомов и фрагментов структур на ингибирующую активность соединений, найденные на основе построенных QSAR-моделей и их визуализации в программе GUSAR, совпадают с закономерностями, выявленными на основе экс-

пертной оценки экспериментальных величин исследованных соединений, что подтверждает возможность использования программы GUSAR для оптимизации структур химических соединений.

Работа выполнена при финансовой поддержке Минобрнауки России (проект 299, 2014 г.) и РФФИ (грант 14-04-97035).

## СПИСОК ЛИТЕРАТУРЫ

- Zarghi, A., and Arfaei, S. (2011) Selective COX-2 inhibitors: a review of their structure–activity relationships, *Iran. J. Pharm. Res.*, **10**, 655–683.
- Marnett, L.J. (2009) Mechanisms of cyclooxygenase-2 inhibition and cardiovascular side effects: the plot thickens, *Cancer Prev. Res. (Philadelphia)*, **2**, 288–290.
- Charlier, C., and Michaux, C. (2003) Dual inhibition of cyclooxygenase-2 (COX-2) and 5-lipoxygenase (5-LOX) as a new strategy to provide safer non-steroidal anti-inflammatory drugs, *Eur. J. Med. Chem.*, **38**, 645–659.
- Smith, W.L., DeWitt, D.L., and Garavito, R.M. (2000) Cyclooxygenases: structural, cellular, and molecular biology, *Annu. Rev. Biochem.*, **69**, 145–182.
- Vane, J.R., Bakhle, Y.S., and Botting, R.M. (1998) Cyclooxygenases 1 and 2, *Annu. Rev. Pharmacol. Toxicol.*, **38**, 97–120.
- Cha, Y.I., and DuBois, R.N. (2007) NSAIDs and cancer prevention: targets downstream of COX-2, *Annu. Rev. Med.*, **58**, 239–252.
- Greenhough, A., Smartt, H.J., Moore, A.E., Roberts, R.R., Williams, A.C., and Kaidi, C.P.A. (2009) The COX-2/PGE2 pathway: key roles in the hallmarks of cancer and adaptation to the tumour microenvironment, *Carcinogenesis*, **30**, 377–386.
- Хельгье Х.Д., Зипль В., Роньян Д., Фолькерс Г. (2010) Молекулярное моделирование (теория и практика) (под ред. Палолина В.А., Радченко Е.В.), БИНОМ. Лаборатория знаний, Москва.
- Иванов А.С., Веселовский А.В., Дубанов А.В., Скворцов В.С., Арчаков А.И. (2006) Интегральная платформа «От гена до прототипа лекарства» *in silico* и *in vitro*, *Российский химический журнал*, **L**, 18–35.
- Khan, M.S., and Khan, Z.H. (2003) Molecular modeling for generation of structural and molecular electronic descriptors for QSAR using quantum mechanical semiempirical and *ab initio* methods, *Genome Informatics*, **14**, 486–487.
- Филимонов Д.А., Пороиков В.В. (2006) Прогноз спектра биологической активности органических соединений, *Российский химический журнал*, **50**, 66–75.
- Пороиков В.В., Филимонов Д.А. (2001) Компьютерный прогноз биологической активности химических соединений как основа для поиска и оптимизации базовых структур новых лекарств. В кн. *Азотистые гетероциклы и алкалоиды* (под ред. Карцева В.Г., Толстикова Г.А.), Иридиум-пресс, Москва, **1**, 123–129.
- Лагунин А.А., Глоризова Т.А., Дмитриев А.В., Волгина Н.Е., Пороиков В.В. (2012) Компьютерная оценка взаимодействия фармакологических веществ с Р-гликопротеином, *Бюлл. экп. биол. мед.*, **154**, 520–524.
- Zakharov, A.V., Lagunin, A.A., Filimonov, D.A., and Poroikov, V.V. (2012) Quantitative prediction of antitarget interaction profiles for chemical compounds, *Chem. Res. Toxicol.*, **25**, 2378–2385.
- Lagunin, A., Zakharov, A., Filimonov, D., and Poroikov, V. (2011) QSAR Modelling of rat acute toxicity on the basis of PASS prediction, *Mol. Informatics*, **30**, 241–250.
- Filimonov, D.A., Zakharov, A.V., Lagunin, A.A., and Poroikov, V.V. (2009) QNA based «Star Track» QSAR approach, *SAR QSAR Environ. Res.*, **20**, 679–709.
- Masanda, V.H., Mahajana, D.T., Patil, K.N., Dawale, N.E., Hadda, T.B., Alafeefy, A.A., and Chinchkhede, K.D. (2011) General unrestricted structure activity relationships based evaluation of quinoxaline derivatives as potential influenza NS1A protein inhibitors, *Der Pharma Chemica*, **3**, 517–525.
- Kokurkina, G.V., Dutov, M.D., Shevelev, S.A., Popkov, S.V., Zakharov, A.V., and Poroikov, V.V. (2011) Synthesis, antifungal activity and QSAR study of 2-arylhydroxynitroindoles, *Eur. J. Med. Chem.*, **46**, 4374–4382.
- Таипов И.А., Хайруллина В.Р., Герчиков А.Я., Хома В.К., Зарудий Ф.С., Бегель Х. (2012) Виртуальный скрининг в ряду эффективных ингибиторов каталитической активности лейкотриен-А4-гидролазы, *Вестник Башкирского университета*, **17**, 886–891.
- Тарасов Г.П., Хайруллина В.Р., Герчиков А.Я., Кирлан С.А., Зарудий Ф.С. (2012) Производные 4-амино-N-[2-(диэтиламино)этил]бензамида как потенциально низкотоксичные вещества с выраженным антиаритмическим действием, *Вестник Башкирского университета*, **17**, 1242–1246.
- Хайруллина В.Р., Кирлан С.А., Герчиков А.Я., Зарудий Ф.С., Димогло А.С., Кантор Е.А. (2010) Моделирование структур противовоспалительных соединений с учетом токсичности, *Башкирский химический журнал*, **17**, 76–79.
- MarvinSketch: <http://www.chemaxon.com/download/marvin-suite>.
- Discovery Studio Visualiser: <http://accelrys.com/resource-center/downloads/freeware/index.html>.
- Portevin, B., Tordjman, C., Pastoureau, P., Bonnet, J., and de Nanteuil, G. (2000) 1,3-Diaryl-4,5,6,7-tetrahydro-2H-isoindole derivatives: a new series of potent and selective COX-2 inhibitors in which a sulfonyl group is not a structural requisite, *J. Med. Chem.*, **43**, 4582–4593.
- Lagunin, A., Filimonov, D., Zakharov, A., Xie, W., Huang, Y., Zhu, F., Shen, T., Yao, J., and Poroikov, V. (2009) Computer-aided prediction of rodent carcinogenicity by PASS and CISOC-PSCT, *QSAR and Combinatorial Science*, **28**, 806–810.
- Silakari, P., Shrivastava, S.D., Silakari, G., Kohli, D.V., Rambabu, G., Srivastava, S., Shrivastava, S.K., and Silakari, O. (2008) QSAR analysis of 1,3-diaryl-4,5,6,7-tetrahydro-2H derivatives as selective COX-2 inhibitors, *Eur. J. Med. Chem.*, **43**, 1559–1569.
- Краткий справочник физико-химических величин* (2003) (под ред. Равделя А.А., Пономаревой А.М.), Иван Федоров, СПб., 212–213.

QUANTITATIVE ANALYSIS OF STRUCTURE–ACTIVITY  
RELATIONSHIPS OF TETRAHYDRO-2H-ISOINDOL  
CYCLOOXYGENASE-2 INHIBITORS

V. R. Khayrullina<sup>1\*</sup>, A. Ya. Gerchikov<sup>1\*</sup>,  
A. A. Lagunin<sup>2,3</sup>, F. S. Zarudiy<sup>4</sup>

<sup>1</sup> Bashkir State University, Faculty of Chemistry, ul. Zaki Validi 32,  
Ufa 450076, Russia; fax: +7(347)229-9707,  
E-mail: gerchikov@inbox.ru, Veronika1979@yandex.ru

<sup>2</sup> N. I. Pirogov Russian National Research Medical University,  
Medico-Biological Faculty, ul. Ostrovitianova 1, Moscow 117997,  
Russia; fax: +7(495)434-1422

<sup>3</sup> V. N. Orekhovich Institute of Biomedical Chemistry,  
Russian Academy of Medical Sciences, ul. Pogodinskaya 10,  
Moscow 119121, Russia; fax: +7(499)245-0857,  
E-mail: alexey.lagunin@ibmc.msk.ru

<sup>4</sup> Bashkir State Medical University, ul. Lenina 3, Ufa 450000,  
Russia; fax: +7(347)272-3751, E-mail: zarudii.f@yandex.ru

Received May 7, 2014

Revision received July 10, 2014

The structure–activity relationships of cyclooxygenase 2 catalytic activity inhibition for twenty-six 4,5,6,7-tetrahydro-2H-isoindole, 2,3-dihydro-1H-pyrrolizine, and benzothiophene derivatives in the range of concentrations 0.6 to 700 nmol/liter IC<sub>50</sub> values were quantitatively analyzed with the GUSAR program. Six statistically significant consensus QSAR models for prediction of IC<sub>50</sub> values were built based on MNA- and QNA-descriptors and their combinations. These models demonstrated high accuracy of prediction of IC<sub>50</sub> values for both training and test sets. Atoms and structural fragments of the studied structures influencing increase and decrease of COX-2 inhibition were identified by GUSAR visualization of quantitative structure–activity relationships in the created models. This information may be relevant to the molecular design of new COX-2 inhibitors. It is shown that in most cases the detected structural fragments related with inhibitory activity of the studied compounds coincide with the results of expert evaluation of their effects based on experimental data that can be used for optimization of structures to change their biological activity.

**Key words:** cyclooxygenase-2 inhibitors, QSAR-models, GUSAR, QNA- and MNA-descriptors, structure–activity relationship analysis