

УДК: 544.165

DOI: 10.33184/bulletin-bsu-2020.4.4

**QSPR-МОДЕЛИРОВАНИЕ АНТИОКСИДТЕЛЬНОЙ АКТИВНОСТИ
ПОТЕНЦИАЛЬНЫХ И ПРОМЫШЛЕННО ИСПОЛЗУЕМЫХ
СТАБИЛИЗАТОРОВ ИЗ КЛАССА ЗАМЕЩЕННЫХ АЛКИЛФЕНОЛОВ**© Ю. З. Мартынова¹, В. Р. Хайруллина^{1*}, А. Я. Герчиков¹,
Ф. С. Зарудий², А. Г. Мустафин¹¹Бакирский государственный университет
Россия, Республика Башкортостан, 450076 г. Уфа, ул. Заки Валиди, 32.²Бакирский государственный медицинский университет,
Россия, Республика Башкортостан, 450000, г. Уфа, ул. Ленина, 3,

*Email: Veronika1979@yandex.ru

С использованием программы GUSAR 2019 выполнен количественный анализ взаимосвязи между структурой и антиокислительной активностью для 113 азот-, сера- и селенсодержащих алкилфенолов в интервале значений $\lg k_7 = 2.48 \div 5.30$. Построено 12 статистически значимых валидных консенсус-моделей. Все эти консенсус-модели предназначены для прогноза параметра $\lg k_7$ в ряду алкилфенолов. Умеренная прогностическая способность моделей подтверждена в результате сопоставления численных значений экспериментальных и расчетных значений параметра $\lg k_7$ для структур четырех тестовых выборок, содержащих структурные аналоги моделируемых соединений.

Ключевые слова: алкилфенолы, полифункциональные антиокислители, моделирование QSPR, дескрипторы, коэффициент детерминации.

Введение

Алкилфенолы занимают ведущее место среди антиокислителей и термостабилизаторов, используемых для предотвращения термоокислительной деструкции и старения полимерных материалов и смазочных масел. Вместе с тем, в последние годы более востребованными в полимерной промышленности являются полифункциональные антиокислители, содержащие в своей структуре фенольный фрагмент, который способен эффективно замедлять окисление благодаря дезактивации пероксильных радикалов, и функциональную группу, способную разрушать гидропероксиды (функциональные группы на основе атом азота, серы, селена, фосфора).

В научной литературе представлено обширное количество исследований, посвященных экспериментальному изучению антирадикальных и антиокислительных свойств пространственно затрудненных фенолов и полифункциональных термостабилизаторов на их основе в различных модельных системах инициированного окисления органических субстратов [1–3]. Вместе с тем ведется поиск новых антиокислительных присадок с улучшенными эксплуатационными характеристиками и более высокой эффективностью защитного действия.

Разработка материалов, в т.ч. и антиокислителей, базируется на двух стратегиях: 1) стратегии синтеза органических соединений и их дальнейшего исследования, исключительно ориентируясь на интуитивный подход и информацию «структура-активность», полученную в узких сериях близких

по строению соединений; 2) стратегия разработки новых структур антиокислителей, основанная на моделировании связи «структура-активность» в больших сериях структурно разнородных соединений.

Первая стратегия является малоперспективной, т.к. не позволяет создавать принципиально новые продукты. Реализация второй стратегии гораздо более перспективна, т.к. позволяет разрабатывать уникальные соединения, отличающиеся широким структурным разнообразием, но в большинстве случаев имеет мало шансов на практическую реализацию, т.к. большинство современных программ моделирования количественной связи «структура-активность» способны строить модели количественного прогноза активности только на основе структурно близких соединений.

Для преодоления этой проблемы в Институте биомедицинской химии им. Ореховича была разработана программа GUSAR 2019, принцип расчета дескрипторов в которой базируется не только исходя из наличия определенных функциональных групп и циклических фрагментов в молекулах, но и свойств отдельных атомов. До настоящего времени эта программа не применялась для моделирования физико-химических свойств органических молекул, включая моделирование количественной связи между строением и антиокислительной активностью синтетических антиокислителей.

В этой связи настоящее исследование ориентировано на моделирование количественной взаимосвязи «структура – антиокислительная активность (АОА)» в ряду азот-, сера- и селенсодержа-

щих алкилфенолов с общими структурными формулами I–IV (рис. 1) с использованием программы GUSAR 2019, построение моделей QSPR, позволяющих количественно предсказывать АОА для разнообразных по строению молекул, а также оценку валидности этих моделей.

Методика вычислительного эксперимента

Для моделирования связи «структура – АОА» использовали программу GUSAR 2019 [4–9]. В моделировании участвовало 113 соединений, общие структурные формулы которых изображены на рис. 1. Модели QSPR строили в четыре этапа.

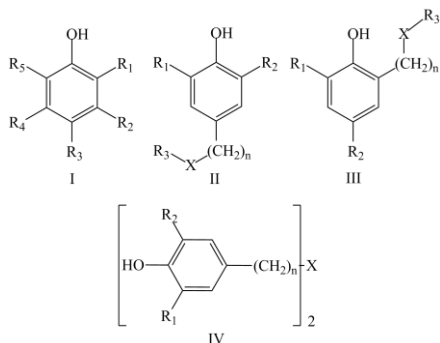


Рисунок 1. Общие структурные формулы антиоксидантов, на основе которых формировали набор данных МД1, где $n = 0\div 4$, $X = S, N$.

1. Сбор, анализ литературных данных и формирование на их основе обучающих и тестовых выборок. Для построения моделей QSPR M1–M12 использовали исходный набор структур антиоксидантов, включающий 113 соединений. Каждому из этих соединений ставилось в соответствие определенное значение константы скорости ингибирования k_7 (рис. 1). В наборе данных МД1 соединения ранжировали в соответствии с возрастанием параметра k_7 . Все используемые в расчетах экспериментальные значения параметра k_7 для моделируемых соединений представлены в научной литературе [10–12] и количественно характеризуют удельную скорость реакции их взаимодействия с кумилпероксильными радикалами при 60 °С в присутствии азодиизобутиронитрила, инициирующего ра-

дикально-цепное окисление кумола. На основе исходного набора данных было сформировано 4 обучающих выборки (ОБ1–ОБ4) и 4 тестовых выборки (ТВ1–ТВ4) в соответствии со схемой, изображенной на рис. 2. Модели QSPR строили на основе параметра $\lg k_7$. Модели QSPR ориентированы на количественный прогноз константы скорости ингибирования $\lg k_7$.

В табл. 1 и 2 представлены характеристики всех обучающих и тестовых выборок соответственно. Обучающая выборка ОБ1 предназначена для построения моделей QSPR M1–M3, для оценки прогностической способности этих моделей служила тестовая выборка ТВ1. Обе эти выборки получены в результате разбивки набора данных МД1 в соотношении 75%:25%. Обучающие выборки ОБ2, ОБ3, ОБ4 предназначены для построения моделей QSPR M4–M6, M7–M9 и M10–M12 соответственно. Для оценки валидности построенные модели M4–M12 тестировали на соединениях тестовых выборок ТВ2, ТВ3, ТВ4. Эти обучающие и тестовые выборки сформированы в результате разбивки набора данных МД1 в соотношениях 80%:20%, 83%:17% и 86%:14% соответственно.

2. Построение моделей QSPR. Модели QSPR M1–M12 строили в программе GUSAR 2019. Эта процедура реализуется автоматически после отбора оптимального набора дескрипторов с использованием метода самосогласованной регрессии SCR, а также ее сочетания с радиальными базисными функциями RBF. Типы дескрипторов, которые были использованы для построения консенсус-моделей M1–M12, представлены в табл. 3.

В программе GUSAR 2019 дескрипторы рассчитывались на основе структурных формул антиоксидантов автоматически. Идеология их расчета детально описана в работах [4–9; 13–16]. Для более объективного учета всех структурных особенностей моделируемых антиоксидантов в выбранной программе генерировали наборы частных регрессионных зависимостей, включающие 20–320 регрессионных уравнений. В дальнейшем на основе заложенного в программе GUSAR алгоритма автоматически подобные модели объединялись в консенсус-модели [4–9; 13–16].

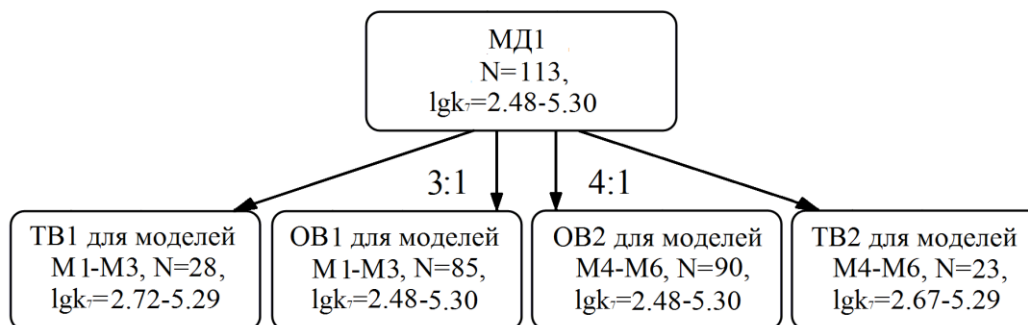


Рис. 2. Принцип формирования обучающих и тестовых выборок для последующего моделирования.

Таблица 1

Распределение значений параметра $\lg k_7$ для соединений в обучающих выборках OB1–OB4

Комментарий	Условные обозначения характеристик выборок	Обучающие выборки			
		OB1	OB2	OB3	OB4
Модели, построенные на основе выборки OB _i	–	M1-M3	M4-M6	M7-M9	M10-M12
Количество антиоксидантов	N	85	90	94	97
Среднее значение параметра $\lg k_7$ по выборке	$\overline{\lg k_7}$			2.824	
Диапазон экспериментальных значений $\lg k_7$	$\Delta \lg k_7$			3.890	
Распределение значений $\lg k_7$	$0.10 \times \Delta \lg k_7$			0.282	
	$0.15 \times \Delta \lg k_7$			0.424	
	$0.20 \times \Delta \lg k_7$			0.565	
	$0.25 \times \Delta \lg k_7$			0.706	

Таблица 2

Распределение значений параметра $\lg k_7$ для соединений тестовых выборок TB1–TB4

Комментарий	Условные обозначения характеристик выборок	Тестовые выборки			
		TB1	TB2	TB3	TB4
Шифры моделей, для проверки которых предназначена выборка TB _i	–	M1-M3	M4-M6	M7-M9	M10-M12
Количество антиоксидантов	N	28	23	19	16
Среднее значение параметра $\lg k_7$ по выборке	$\overline{\lg k_7}$	4.560	4.5367	4.5170	4.5801
Диапазон экспериментальных значений $\lg k_7$	$\Delta \lg k_7$	2.5658	2.6198	2.6176	1.7827
Распределение значений $\lg k_7$ для TB _i вокруг средней TB _i (в %)	$\overline{\lg k_7} \pm 0.5_{(TB_i)}, \%$	67.857	60.870	63.1580	68.750
	$\overline{\lg k_7} \pm 1.0_{(TB_i)}, \%$	96.429	95.652	94.737	93.750
	$\overline{\lg k_7} \pm 1.5_{(TB_i)}, \%$	96.429	95.652	94.737	100.000
	$\overline{\lg k_7} \pm 2.0_{(TB_i)}, \%$	100.000	100.000	100.000	100.000
	$\overline{\lg k_7} \pm 0.5_{(OB_i)}, \%$	35.714	34.783	36.842	37.500
Распределение значений $\lg k_7$ TB _i вокруг средней OB _i (в %)	$\overline{\lg k_7} \pm 1.0_{(OB_i)}, \%$	60.714	60.870	57.895	62.500
	$\overline{\lg k_7} \pm 1.5_{(OB_i)}, \%$	100.000	100.000	100.000	100.000
	$\overline{\lg k_7} \pm 2.0_{(OB_i)}, \%$	100.000	100.000	100.000	100.000

Таблица 3

Типы дескрипторов, которые были использованы для построения консенсус-моделей M1–M12 [4–9, 13–6]

Шифр консенсус модели	N_M^*	Названия используемых дескрипторов
M1, M4, M7, M10	20	QNA-дескрипторы, топологическая длина, топологический объем, липофильность
M2, M5, M8, M11	20	MNA-дескрипторы, топологическая длина
M3, M6, M9, M12	320	QNA-дескрипторы, MNA-дескрипторы, топологическая длина, топологический объем, липофильность

* N_M – число частных регрессионных моделей, вошедших в данную консенсус-модель/

3. Оценка описательной способности моделей QSPR. Для оценки описательной способности моделей QSPR M1–M12 использовали разные типы коэффициентов детерминации, включая R^2_{train} , R^2_0 , $\overline{R^2}_{m\ train}$, CCC_{train} . Они рассчитывались в программе Xternal Validation Plus 1.2 [19] при сопоставлении экспериментальных и расчетных значений

параметра $\lg k_7$ для соединений обучающих выборок OB1–OB4 [19].

4. Оценка прогностической способности моделей QSPR. При оценке прогностической способности моделей ориентировались на следующие параметры 1) на численные значения разных типов коэффициентов детерминации (R^2_{test} , $R^2_0\ test$, $Q^2(F1)$,

$Q^2_{(F2)}$, $\overline{R^2}_{m\ test}$, CCC_{test}); 2) на численные величин ошибок прогноза значений $\lg k_7$ (среднеквадратичная ошибка RMSEP, средняя абсолютная ошибка MAE, стандартное отклонение SD) [17–18]. Все эти статистические параметры рассчитывались в программе Xternal Validation Plus 1.2 автоматически при сопоставлении экспериментальных и расчетных значений параметра $\lg k_7$ для соединений тестовых выборок TB1–TB4 [19]. В этой же программе оценивали возможность наличия систематических ошибок моделирования.

Результаты и их обсуждение

В результате моделирования QSPR, заложенного в программе GUSAR 2019, в общей сложности нами было построено 12 консенсус-моделей QSPR, предназначенных для количественного прогноза параметра $\lg k_7$ для азот-, сера- и селенсодержащих алкилфенолов с общими структурными формулами I–IV. Описательная и прогностическая способность моделей M1–M12 оценена на основе анализа численных значений различных статистических критериев, расчет которых осуществлялся автоматически в программах GUSAR 2019 и Xter-

nalValidationPlus 1.2 [19]. В частности, статистические параметры консенсус-моделей QSPR M1–M12, оцененные с использованием программы GUSAR 2019, представлены в *табл. 4*. Все численные значения статистических параметров, соответствующие каждой консенсус-модели M_i , которые содержатся в *табл. 4*, являются усредненными значениями, которые получены с учетом всех частных регрессионных моделей N_M (*табл. 3*).

Для проверки устойчивости моделей M1–M12 применялся скользящий контроль, при реализации которого из каждой обучающей выборки OB_i рандомизированно 20 раз исключалось 20% соединений. Об устойчивости построенных нами консенсус-моделей свидетельствует незначительная разница численных значений коэффициентов детерминации R^2 и Q^2 , усредненные значения которых для каждой модели представлены в *табл. 4*. Обработка экспериментальных и расчетных значений параметра $\lg k_7$ в программе XternalValidationPlus 1.2 показала отсутствие систематической ошибки. Этот результат свидетельствует во-первых, об универсальности дескрипторов, использованных для построения моделей, во-вторых, о корректности построения регрессионных зависимостей в программе GUSAR 2019.

Таблица 4

Статистические параметры и оценка показателя точности предсказаний консенсус-моделей QSPR M1–M12; $\square \lg k_7 (OB_i) = 2.82$

Шифр обучающей выборки	Шифр модели	N	N_{PM}	$\overline{R^2}_{OB}$	\overline{F}	\overline{SD}	Q^2_{OB}	$\overline{R^2}_{TB}$	W
Модели на основе QNA-дескрипторов и дескрипторов целевой молекулы									
OB1	M1	85	20	0.983	5.615	0.344	0.607	0.6159	17
OB2	M4	90	20	1.000	5.025	0.301	0.705	0.5685	23
OB3	M7	94	20	0.984	4.000	0.323	0.644	0.6758	26
OB4	M10	97	20	1.000	4.781	0.327	0.636	0.6621	22
Модели на основе MNA-дескрипторов и дескрипторов целевой молекулы									
OB1	M2	85	20	0.984	2.637	0.386	0.510	0.5796	21
OB2	M5	90	20	1.000	2.248	0.382	0.514	0.7572	23
OB3	M8	94	20	0.983	2.123	0.389	0.487	0.7998	24
OB4	M11	97	20	1.000	3.067	0.357	0.580	0.4143	22
Модели на основе QNA- и MNA-дескрипторов и дескрипторов целевой молекулы									
OB1	M3	85	320	0.984	3.932	0.342	0.630	0.671	18
OB2	M6	90	320	0.998	4.023	0.330	0.670	0.670	19
OB3	M9	94	320	0.986	3.510	0.338	0.630	0.795	21
OB4	M12	97	320	0.999	4.740	0.322	0.671	0.513	1

N – число антиоксидантов в обучающей выборке; R^2_{OB} – коэффициент детерминации, рассчитанный на основе антиоксидантов обучающей выборки; R^2_{TB} – коэффициент детерминации рассчитанный на основе антиоксидантов тестовой выборки; Q^2_{OB} – коэффициент детерминации, рассчитанный на основе антиоксидантов обучающей выборки при скользящем контроле с исключением 20% соединений; F – критерий Фишера; SD – стандартное отклонение; W – число дескрипторов в итоговом регрессионном уравнении.

Таблица 5

Критерии для оценки прогностической способности моделей QSPR M1–M12

Качество модели	Критерии модели с высокой прогностической способностью	Критерии модели с умеренной прогностической способностью	Критерии модели с низкой прогностической способностью
Критерий на основе коэффициентов детерминации R^2	$R^2 \rightarrow R^2_0 \rightarrow 1$, $\overline{R^2}_m > 0.5$, $\Delta R^2_m > 0.2$, $CCC \rightarrow 1$	$R^2 \rightarrow R^2_0 \rightarrow 0.7$, $\overline{R^2}_m > 0.5$, $\Delta R^2_m < 0.2$, $CCC \rightarrow 0.8$	$R^2 \rightarrow R^2_0 \rightarrow 0.6$, $\overline{R^2}_m > 0.5$, $\Delta R^2_m < 0.2$, $CCC \rightarrow 0.7$
Критерий на основе ошибки прогноза MAE	$MAE \leq 0.1 \Delta \lg k_7$, $MAE + 3SD \leq 0.2 \Delta \lg k_7$	$MAE \leq 0.15 \Delta \lg k_7$, $MAE + 3SD \leq 0.2 \Delta \lg k_7$	$MAE > 0.15 \Delta \lg k_7$, $MAE + 3SD \geq 0.25 \Delta \lg k_7$

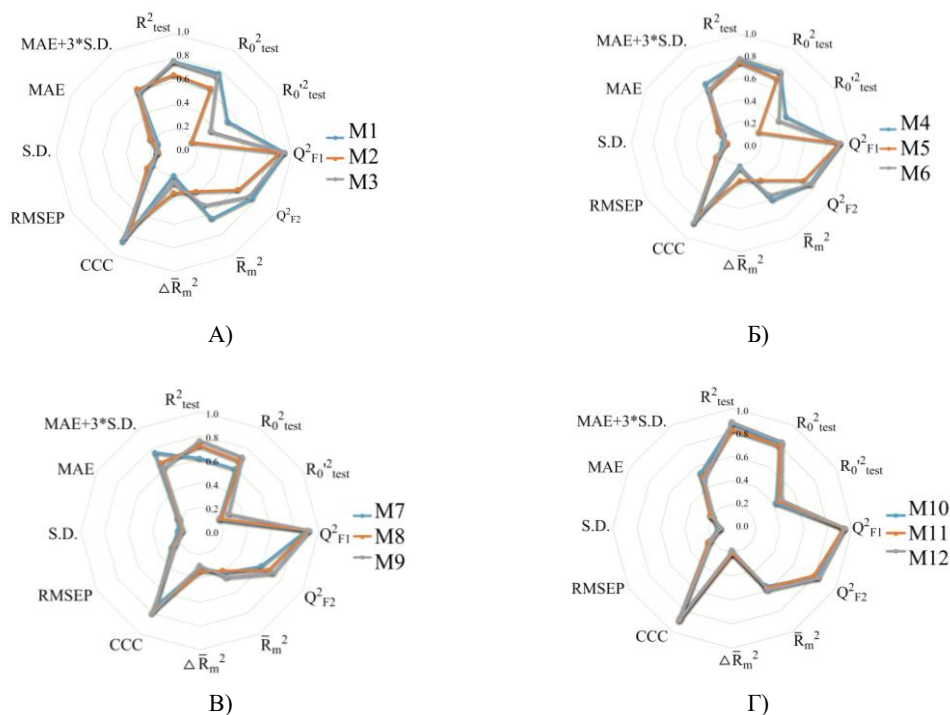


Рис. 3. Статистические параметры для производных замещенных фенолов, содержащихся в тестовых выборках TB1 (А) и TB2 (Б), TB3 (В), TB4 (Г) по консенсус-моделям QSPR M1–M12, построенных с использованием метода отбора дескрипторов SCR-RBF.

На рис. 3 изображены диаграммы, которые построены на основе коэффициентов детерминации (R^2_0 , $Q^2_{(F1)}$, $Q^2_{(F2)}$, R_m^2 , CCC) и средней абсолютной ошибки MAE, рассчитанных в программе XternalValidationPlus при сопоставлении экспериментальных и расчетных значений параметра $\lg k_7$ для антиоксидантов, содержащихся в тестовых выборках TB1–TB4.

В табл. 5 представлены пороговые значения статистических параметров, на основе которых нами проводилась оценка описательной и прогностической способности моделей M1–M12. Эти критерии носят рекомендательный характер и описаны в ряде работ профессора Kunal Roy с соавторами [20–21].

Анализ численных значений статистических параметров, изображенных на диаграммах в соответствии с критериями, представленными в табл. 5, свидетельствует о довольно высокой описательной и умеренной прогностической способности построенных нами консенсус-моделей.

Как видно из диаграмм (рис. 3), разница между численными значениями разных коэффициентов детерминации фактически является незначительной. Пороговые значения этих критериев отвечают моделям с умеренной прогностической способностью. Значения ошибки MAE и рассчитанного на ее основе критерия MAE + 3SD, находятся в интервале, который тоже соответствует моделям с умеренной прогностической способностью.

Все эти количественные характеристики свидетельствуют о том, что построенные нами модели QSPR M1–M12 могут быть использованы для поиска и разработки новых потенциальных полифункциональных антиоксидантов, в том числе и стабилизаторов к полимерным продуктам, на основе азот-, сера- и селенсодержащих алкилфенолов.

Таким образом, программа GUSAR 2019, которая изначально предназначалась для моделирования биологической активности, может быть использована и для моделирования антиокислительной активности в виде физико-химического параметра $\lg k_7$, что значительно расширяет возможности моделирования на ее основе и открывает новые перспективы ее использования в фундаментальных и прикладных исследованиях.

Выводы

1. С использованием программы GUSAR 2019 проведено QSPR-моделирование связи «структура – антиокислительная активность» в ряду 113 азот-, сера- и селенсодержащих алкилфенолов, в результате которого построено 12 консенсус-моделей QSPR. Все эти модели могут быть использованы для прогноза численных значений параметра $\lg k_7$ в ряду структурных аналогов моделируемых соединений.

2. С использованием структур синтетических антиоксидантов, входящих в обучающие и тестовые выборки, показана высокая описательная и

умеренная прогностическая способность QSPR для всех двенадцати консенсус-моделей.

3. На основании статистических параметров, полученных в результате расчетов, было установлено, что наибольшей прогностической способностью при прогнозе значения параметра lgk_7 обладают модели M10, M11 и M12.

Работа выполнена при финансовой поддержке РНФ, проект 19-73-20073.

ЛИТЕРАТУРА

- Nimse S. B., Pal D. Free radicals, natural antioxidants, and their reaction mechanisms // RSC Adv. 2015. Vol. 5. Pp. 27986–28006.
- Khairullina V. R., Gerchikov A. Ya., Akhmetkhanov R. M., Gabitov I. T., Minniyanova R. T., Safarova A. B., Chukchieva I. Yu., Kuchin A. V. Antioxidant Properties of Some 4-, 6-Methyl-substituted Derivatives of Isobornylphenol // Russ. J. Appl. Chem. 2012. Vol. 85. No. 3. Pp. 401–406.
- Bendary E., Francis R. R., Ali H. M. G., Sarwat M. I., El Hady S. Antioxidant and structure-activity relationships (SARs) of some phenolic and anilines compounds // J. Ann. Agric. Sci. 2013. Vol. 58. No. 2. Pp. 173–181.
- Khairullina V. R., Gerchikov A. Ja., Lagunin A. A., Zarudii F. S. QSAR Modelling of Thymidylate Synthase Inhibitors in a Series of Quinazoline Derivatives // Pharm. Chem. J. 2018. Vol. 51. No. 10. Pp. 884–888.
- Khairullina V. R., Gimadieva A. R., Gerchikov A. Ja., Mustafin A. G., Zaarudii F. S. Quantitative structure-activity relationship of the thymidylate synthase inhibitors of *Mus musculus* in the series of quinazolin-4-one and quinazolin-4-imine derivatives // J. Mol. Graphics Modell. 2018. Vol. 85. Pp. 198–211.
- Мартынова Ю. З., Хайруллина В. Р., Гарифуллина Г. Г., мицукова Д. С., Зарудий Ф. С., Мустафин А. Г. QSAR-моделирование связи «структура – антиокислительная активность» в ряду некоторых производных бензопирана и бензофурана // Вестник БашГУ. 2019. Т. 24. №3. С. 573–580.
- Хайруллина В. Р., Герчиков А. Я., Лагунин А. А., Зарудий Ф. С. Количественный анализ взаимосвязи «структура-активность» ингибиторов циклооксигеназы-2 среди производных тетрагидро-2Н-изоиндола // Биохимия. 2015. Т. 80, №1. С. 96–110.
- Хайруллина В. Р., Герчиков А. Я., Зарудий Ф. С. Анализ взаимосвязи «структура-ингибирующая активность циклооксигеназы-2» в ряду производных ди-трет-бутилфенола, тиазолон и оксазолон // Вестник БашГУ. 2014. Т. 19. №2. С. 417–422.
- Martynova Y. Z., Khairullina V. R., Biglova Y. N., Mustafin A. G. Quantitative structure-property relationship modeling of the C60 fullerene derivatives as electron acceptors of polymer solar cells: Elucidating the functional groups critical for device performance // Journal of Molecular Graphics and Modelling. 2019. Vol. 88. Pp. 49–61.
- Фарзалиев В. М., Джафарова Т. Д., Мирзоева М. А., Рзаева И. А., Мамедов Ф. А. Исследование влияния некоторых производных 2-метилтио-4-трет-бутилфенола на окисление кумола // Азербайджанский хим. журнал. 2014. №2. С. 26–30.
- Хольшин С. В., Чеблукова В. П., Ягунов С. Е., Олейник А. С., Кандалинцева Н. В., Просенко А. Е. Синтез и антиоксидантные свойства додецил-(3-(4-гидроксиарил)пропил)селенидов // Вестник Новосибирского гос. пед. ун-та. 2015. Т. 3. №25. С. 112–123.
- Терах Е. И., Просенко А. Е. Реакционная способность тио(амино)производных ω - α -гидроксиарил)алкильного типа в радикальных реакциях // Кинетика и катализ. 2012. Т. 53. №5. С. 563–569.
- Zakharov A. V., Lagunin A. A., Filimonov D. A., Poroikov V. V. Quantitative prediction of antitarget interaction profiles for chemical compounds // Chemical Research in Toxicology. 2012. Vol. 25. No 11. Pp. 2378–2385.
- Таипов И. А., Хайруллина В. Р., Герчиков А. Я., Хома В. К., Заридий Ф. С., Бегель Х. Виртуальный скрининг в ряду эффективных ингибиторов каталитической активности лейкотриен А4-гидролазы // Вестник БашГУ. 2012. Т. 17. №2. С. 886–891.
- Lagunin A. A., Romanova M. A., Zadorozhny A. D., Kurilenko N. S., Shilov B. V., Pogodin P. V., Ivanov S. M., Filimonov D. A., Poroikov V. V. Comparison of Quantitative and Qualitative (Q)SAR Models Created for the Prediction of K_i and IC_{50} Values of Antitarget Inhibitors // Frontiers in Pharmacology. 2018. Vol. 9. P. 1–11.
- Lagunin A. A., Geronikaki A., Eleftheriou P., Pogodin P. V., Zakharov A. V. Rational Use of Heterogeneous Data in Quantitative Structure – Activity Relationship (QSAR) Modeling of Cyclooxygenase/Lipoxygenase Inhibitors // J. Chem. Inf. Model. 2019. Vol. 59. P. 713–730.
- Dearden J. C., Cronin M.T.D., Kaiser K. L. E. How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR) // SAR and QSAR in Environmental Research. 2009. Vol. 20. №3–4. P. 241–266.
- Roy K., Mitra I., Kar S., Ojha P. K., Das R. N., Kabir H. Comparative Studies on Some Metrics for External Validation of QSPR Models // Journal of Chemical Information and Modeling. 2012. Vol. 52. Pp. 396–408.
- Xternal Validation Plus. URL: <https://sites.google.com/site/dtclabxvplus/>
- Roy P. P., Paul S., Mitra I., Roy K. On Two Novel Parameters for Validation of Predictive QSAR Models // Molecules. 2009. Vol. 14. Pp. 1660–1701.
- Roy K., Das R. N., Ambure P., Aher R. B. Be aware of error measures. Further studies on validation of predictive QSAR models // Chemometrics and Intelligent Laboratory Systems. 2016. Vol. 152. Pp. 18–33.

Поступила в редакцию 16.09.2020 г.

После доработки – 26.11.2020 г.

DOI: 10.33184/bulletin-bsu-2020.4.4

**QSPR-MODELING OF ANTIOXIDANT ACTIVITY
OF POTENTIAL AND INDUSTRIAL USED STABILIZERS
FROM THE CLASS OF SUBSTITUTED ALKYLPHENOLS**

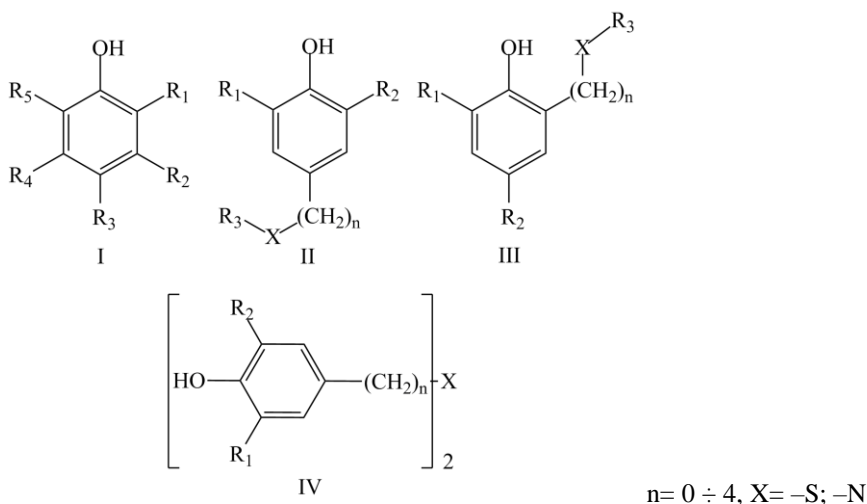
© Yu. Z. Martynova¹, V.R. Khairullina^{1*}, A. Ya. Gerchikov¹,
F. S. Zarudiy², A. G. Mustafin¹

¹Bashkir State University
32 Zaki Validi Street, 450076 Ufa, Republic of Bashkortostan, Russia.

²Bashkir State Medical University
3 Lenin Street, 450000 Ufa, Republic Bashkortostan, Russia.

*Email: veronika1979@yandex.ru

Using the computer program GUSAR 2019 (General Unrestricted Structure Activity Relationships), the quantitative relationship “structure – antioxidant activity (AOA)” was studied in a series of 113 nitrogen-, sulfur-, and selenium-containing alkylphenols with the general structural formulas I–IV.



Twelve statistically significant valid consensus models characterized by high descriptive and moderate predictive accuracy of the $\lg k_7$ parameter for the structures of training and test samples were constructed. AOA of the modeled compounds was in the range of values $\lg k_7 = 2.48 \div 5.30$. The descriptive and predictive ability of these models was evaluated on the structures of four training and four test samples with sliding controls with 20 times randomized exception of 20% of compounds. As statistical criteria for assessing the validity of the models M1–M9, the coefficients of determination (R^2_0 , $Q^2_{(F1)}$, $Q^2_{(F2)}$, $\overline{R^2_{\text{mTSi}}}$, CCC) and the mean absolute error (MAE) were used. For their calculation, the XternalValidationPlus 1.2 program was used. This program showed the absence of a systematic error in the construction of all the regression dependencies M1–M12, which indicates the universality of the descriptors embedded in the GUSAR 2019 program, as well as the correctness of the regression models based on them. Thus, QSAR M1–M9 models can be successfully used for virtual screening of libraries and various databases in order to rationally search for potential antioxidants in them.

Keywords: antioxidant activity, antioxidants, QSAR modeling, GUSAR 2013 program, MNA and QNA descriptors.

Published in Russian. Do not hesitate to contact us at bulletin_bsu@mail.ru if you need translation of the article.

REFERENCES

1. Nimse S. B., Pal D. RSC Adv. 2015. Vol. 5. Pp. 27986–28006.
2. Khairullina V. R., Gerchikov A. Ya., Akhmetkhanov R. M., Gabitov I. T., Minniyanova R. T., Safarova A. B., Chukchieva I. Yu., Kuchin A. V. Russ. J. Appl. Chem. 2012. Vol. 85. No. 3. Pp. 401–406.
3. Bendary E., Francis R. R., Ali H. M. G., Sarwat M. I., El Hady S. J. Ann. Agric. Sci. 2013. Vol. 58. No. 2. Pp. 173–181.
4. Khairullina V. R., Gerchikov A. Ja., Lagunin A. A., Zarudii F. S. Pharm. Chem. J. 2018. Vol. 51. No. 10. Pp. 884–888.
5. Khairullina V. R., Gimadieva A. R., Gerchikov A. Ja., Mustafin A. G., Zaarudii F. S. J. Mol. Graphics Modell. 2018. Vol. 85. Pp. 198–211.
6. Martynova Yu. Z., Khairullina V. R., Garifullina G. G., mitsukova D. S., Zarudii F. S., Mustafin A. G. Vestnik BashGU. 2019. Vol. 24. No. 3. Pp. 573–580.
7. Khairullina V. R., Gerchikov A. Ya., Lagunin A. A., Zarudii F. S. Biokhimiya. 2015. Vol. 80, No. 1. Pp. 96–110.
8. Khairullina V. R., Gerchikov A. Ya., Zarudii F. S. Vestnik BashGU. 2014. Vol. 19. No. 2. Pp. 417–422.
9. Martynova Y. Z., Khairullina V. R., Biglova Y. N., Mustafin A. G. Journal of Molecular Graphics and Modelling. 2019. Vol. 88. Pp. 49–61.
10. Farzaliev V. M., Dzhafarova T. D., Mirzoeva M. A., Rzaeva I. A., Mamedov F. A. Azerbaidzhanskii khim. zhurnal. 2014. No. 2. Pp. 26–30.
11. Khol'shin S. V., Cheblukova V. P., Yagunov S. E., Oleinik A. S., Kandalintseva N. V., Prosenko A. E. Vestnik Novosibirskogo gos. ped. un-ta. 2015. Vol. 3. No. 25. Pp. 112–123.
12. Terakh E. I., Prosenko A. E. Kinetika i kataliz. 2012. Vol. 53. No. 5. Pp. 563–569.
13. Zakharov A. V., Lagunin A. A., Filimonov D. A., Poroikov V. V. Chemical Research in Toxicology. 2012. Vol. 25. No. 11. Pp. 2378–2385.
14. Taipov I. A., Khairullina V. R., Gerchikov A. Ya., Khoma V. K., Zaridii F. S., Begel' Kh. Vestnik BashGU. 2012. Vol. 17. No. 2. Pp. 886–891.
15. Lagunin A. A., Romanova M. A., Zadorozhny A. D., Kurilenko N. S., Shilov B. V., Pogodin P. V., Ivanov S. M., Filimonov D. A., Poroikov V. V. Frontiers in Pharmacology. 2018. Vol. 9. Pp. 1–11.
16. Lagunin A. A., Geronikaki A., Eleftheriou P., Pogodin P. V., Zakharov A. V. J. Chem. Inf. Model. 2019. Vol. 59. Pp. 713–730.
17. Dearden J. C., Cronin M.T.D., Kaiser K. L. E. SAR and QSAR in Environmental Research. 2009. Vol. 20. No. 3–4. Pp. 241–266.
18. Roy K., Mitra I., Kar S., Ojha P. K., Das R. N., Kabir H. Journal of Chemical Information and Modeling. 2012. Vol. 52. Pp. 396–408.
19. Xternal Validation Plus. URL: <https://sites.google.com/site/dtclabxvplus/>
20. Roy P. P., Paul S., Mitra I., Roy K. Molecules. 2009. Vol. 14. Pp. 1660–1701.
21. Roy K., Das R. N., Ambure P., Aher R. B. Chemometrics and Intelligent Laboratory Systems. 2016. Vol. 152. Pp. 18–33.

Received 16.09.2020.

Revised 26.11.2020.