

В.В. Важев¹, М.А. Губенко², Н.В. Важева², Э.М. Ергалиева²

¹Костанайский социально-технический университет им. З. Алдамжар;

²Костанайский государственный педагогический институт
(E-mail: vladimir.vazhev@gmail.com)

Прогнозирование противоопухолевой активности химических соединений методом QSAR

В статье исследована возможность применения QSAR метода для одновременного моделирования противоопухолевой активности больших наборов химических соединений в отношении клеток человеческой меланомы SK-MEL-2. Для набора 250 органических веществ разных классов достигнут коэффициент корреляции между экспериментальными и вычисленными значениями $\lg(LC_{50})$ $R = 0,956$.

Ключевые слова: QSAR, противоопухолевая активность, меланома, органические вещества.

На сегодняшний день эффективность системной лекарственной терапии в клинической онкологии, позволяющей улучшить показатели отдаленной выживаемости больных раком, остается проблемной. Это связано, прежде всего, с тем, что лишь немногие из противоопухолевых агентов зарегистрированы в качестве лекарственных средств и вошли в стандартные схемы лечения. Неудачи в терапии онкологических заболеваний, как правило, связаны с отсутствием у новых препаратов преимуществ перед существующими, низкой селективностью действия или трудоемкостью их создания. Традиционно существует и отставание применения научных достижений в клинической практике. Поиск и создание нового лекарственного препарата сопряжены как с большими материальными затратами, так и с риском получения отрицательных результатов из-за возможного выявления побочных эффектов и токсичности у испытуемых соединений. Таким образом, разработка новых подходов к созданию и экспериментальному изучению соединений, проявляющих противоопухолевую активность, является, несомненно, существенным вкладом в решение актуальной проблемы фундаментальной и практической науки в области медицинской химии, онкологии и биохимии.

Изучению количественных соотношений структура — активность в рамках QSAR направления (Quantitative Structure — Activity Relationship) посвящены сотни работ по моделированию и прогнозированию противоопухолевой активности химических соединений. В исследованиях использован весь спектр известных дескрипторов и алгоритмов решений. Качество QSAR моделей принято оценивать коэффициентом корреляции R (или R^2) между экспериментальными и вычисленными значениями свойства, а также соответствующим стандартным отклонением s .

В работе [1] представлены данные по моделированию противоопухолевой активности 52 производных тилофрина (13aS)-2,3,6,7-тетраметокси-9,11,12,13,13a,14-гексагидродибензо[F,H]пирроло-[1,2-b]изохинолин), для лучшей модели показатель $R^2 > 0,57$.

При изучении 19 производных диспироиндолов в отношении различных опухолевых клеточных линий человека (HepG2 «печени», HELA «шейки матки» и PC3 «простаты») с использованием трех дескрипторов Рихамом и другими [2] в лучшей модели получено $R^2 = 0,903$.

А. Катрицким и другими [3] выполнены моделирование и прогнозирование более эффективных противораковых агентов в отношении лейкемии с использованием дескрипторов CODESSA PRO. На тренировочной выборке из 26 веществ достигнут $R^2 = 0,771$, для контрольной из 8 веществ — $R^2 = 0,781$.

Рядом авторов [4] синтезирован и исследован небольшой набор из 18 производных изохинолин-3-карбоновой кислоты, ингибирующих рост опухолей. Лучшая модель характеризовалась $R = 0,989$.

Группой Рао и других [5] для 37 производных 1,2,4,5-тетразина с противоопухолевой активностью в отношении А-549 клеток были получены модели с хорошими прогнозирующими способностями, $R^2 = 0,916$.

В работе [6] исследована серия производных 2-гидразино-4-(4-морфолинил)тиено[3,2-d]пиримидина, $R^2 = 0,947$.

При моделировании противоопухолевого потенциала набора из 65 производных 1H-имидазо-[4,5-b]пиридина достигнуто $R^2 = 0,98$ [7].

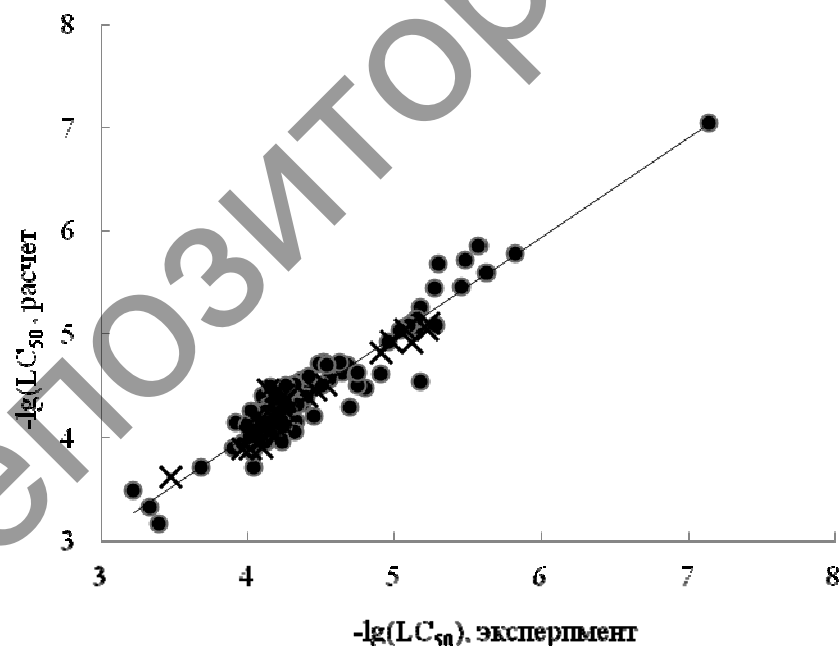
Количественные отношения структура – противораковая активность производных изатина исследованы в работе [8]. Топологические, химические, геометрические и функциональные группы дескрипторов оказались эффективными параметрами на цитотоксическую активность, $R^2 = 0,92$.

Несколько моделей были построены для антипролиферативной активности замещенного фенил-4-(2-оксо-1-имидазолидинил)бензолсульфоната (PIB-SOs) [9]. Лучшие параметрические модели показали $R^2 = 0,94$. Модель производных N-гидрокси-N'-аминогуанидина [10] имела $R^2 = 0,858$.

Рассмотренные примеры прогнозирования (как и другие имеющиеся работы) содержат такие общие черты, как небольшой (несколько десятков веществ) набор из структурно схожих производных основного соединения, малое число привлеченных дескрипторов и невысокие статистические параметры моделей. Такая ситуация вполне естественна, если учесть многофакторность и сложность объекта исследований. Очевидно, для повышения качества прогнозирования следует привлекать наибольшее число дескрипторов, «схватывающих» все структурные нюансы соединений. Однако на этом пути существует проблема отбора дескрипторов, которая не имеет пока удовлетворительного решения.

В наших работах предложен [11] алгоритм, реализованный в программе PROGROC, который снимет проблему отбора дескрипторов и позволит привлечь любое их количество.

В данной работе нами предпринято исследование возможности одновременного моделирования противоопухолевой активности больших наборов химических соединений разных классов. Объектами исследования выбраны линии клеток человеческой меланомы SK-MEL-2 и значения логарифма LC_{50} (Lethal concentration 50 %) — концентрации химического вещества, которая вызывает гибель 50 % испытываемой популяции раковых клеток на протяжении определенного временного промежутка. Для качественного прогнозирования необходимо иметь достоверную базу с данными, как по канцерогенной, так и по противоопухолевой активности. Одной из таких баз является база данных Института США (NCI), занимающегося проблемами рака [12]. Был сформирован набор из 250 химических соединений, имеющих наиболее достоверные значения LC_{50} и обладающих молекулярной массой от 112 до 440. Были использованы все 739 дескрипторов, генерируемых программой Dragon для данного набора веществ. Тренировочная выборка состояла из 200 веществ, контрольная — из 50. Результаты прогнозирования приведены на рисунке 1.



● — тренировочная выборка; × — контрольная выборка

Рисунок 1. Корреляционная зависимость между экспериментальными и вычисленными значениями $-\lg(LC_{50})$

График, представленный на рисунке 1, отражает сильную корреляционную зависимость между экспериментальными и вычисленными значениями ($R = 0,956$ для всего набора), что является

свидетельством правильности использованного метода расчета. Высокие показатели корреляции подтверждаются видом гистограммы (рис. 2). Частота распределения ошибок близка к нормальной, что говорит об адекватности модели.

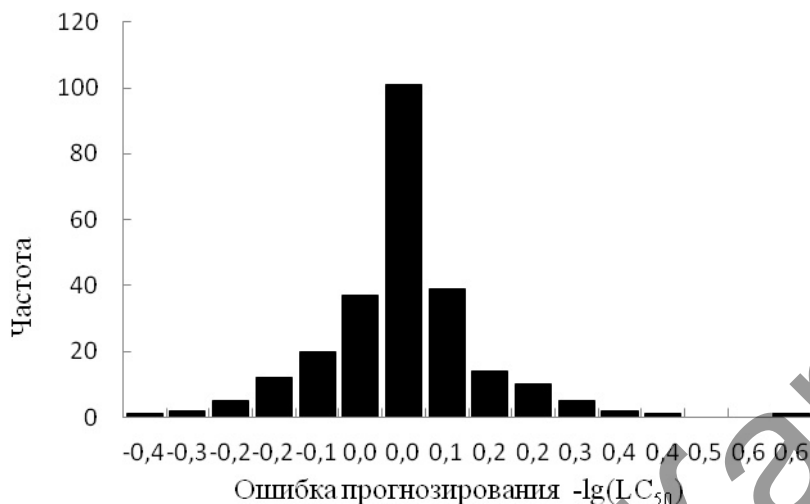


Рисунок 2. Гистограмма распределения частот ошибок прогнозирования LC₅₀

Статистическая характеристика результатов прогнозирования для тренировочной и контрольной выборок представлена в таблице 1. В практике прогнозирования биологической активности химических соединений коэффициенты корреляции выше 0,9 считаются весьма высокими.

Т а б л и ц а 1

Показатели корреляции между экспериментальными и вычисленными значениями $-lg(LC_{50})$

Показатели корреляции	Весь набор	Тренировочная выборка	Контрольная выборка
<i>R</i>	0,956	0,962	0,954
<i>s</i>	0,12	0,11	0,13
Количество веществ	250	200	50

В таблице 2 представлена часть контрольной выборки, которая иллюстрирует:

- порядки значений $-lg(LC_{50})$;
- точность прогнозирования;
- разнообразие классов химических соединений.

Т а б л и ц а 2

Экспериментальные [12] и рассчитанные значения $-lg(LC_{50})$ для клеток меланомы SK-MEL-2 (контрольная выборка, часть)

№	Соединение	Эксп.	Расч.	Абс. разн.	Отн. разн., %
1	2	3	4	5	6
1	4-[(4-Аминофенил)(4-имино-2,5-циклогексадиен-1-илиден)метил]-2-метиланилин	4,36	4,55	0,19	4,36
2	(2Z)-2,3-Дихлор-4-[(3-хлорфенил)амино]-4-оксо-2-бутеновая кислота	4,01	3,90	0,11	2,74
3	2-[(4-Нитрофенил)амино]нафто[2,3-d][1,3]тиазол-4,9-дион	4,12	4,26	0,14	3,40
4	2,2-Дибутил-1,3,2-оксагистаннолан	5,48	5,72	0,24	4,38
5	4Н-Фуоро[3,2-г]хромена-4,7,9-трион	4,20	4,35	0,15	3,57
6	{Бензил[2-гидрокси-2-(1-фенилвинил) циклобутил]амино}ацетонитрил	4,09	4,28	0,19	4,65
7	1-[(2-Амино-6-фенил-4-пиримидинил)метил]циклогексанола	4,02	4,02	0,00	0,00

1	2	3	4	5	6
8	3-[2-(4-Бромфенил)-2-оксоэтил]-5,7-дихлор-3-гидрокси-1,3-дигидро-2H-индол-2-он	4,09	4,20	0,11	2,69
9	5-[(E)-(2-Бромфенил)дiazенил]-4-метил-1,3-тиазол-2-амин	4,07	4,03	0,04	0,98
10	4-Бром-3-бутил-5-(дибромметил)-2(5H)-фуранон	4,49	4,71	0,22	4,90
11	17-Оксоандростан-3-ил(2-хлорэтил)нитрозокарбамат	4,07	4,01	0,06	1,47
12	2-Хлор-1-(5-нитро-2-фурил)этанон	4,09	3,95	0,14	3,42
13	3-(Бромметил)хиноксалин-2-ол	4,29	4,17	0,12	2,80
14	4-[(1E)-3-Оксо-3-фенил-1-пропен-1-ил]бензонитрил	4,12	3,95	0,17	4,13
15	Платина(4+)хлорид азанид (1:2:2)	4,04	3,71	0,33	8,17
16	7-[(3-Метоксифенил)сульфонил]-4-нитро-2,1,3-бензоксадиазол 1-оксид	4,02	4,05	0,03	0,75
17	2,2-Дибутил-8-метокси-4H-1,3,2-бензодиоксастаннин-4-он	5,57	5,86	0,29	5,21
18	(2Z)-2,3-Дихлор-4-[(2,6-диметилфенил)амино]-4-оксо-2-бутеновая кислота	4,23	4,34	0,11	2,60
19	2-Изопропил-5,6,7,8-тетрагидро-3,4-фенантрендион	4,32	4,51	0,19	4,40
20	2-(4-{(Z)-[1-(2-Гидроксиэтил)-3,5-диметил-1H-пиразол-4-ил]-NNO-азокси}-3,5-диметил-1H-пиразол-1-ил)этанол	4,11	4,09	0,02	0,49
21	Метил-6-хлор-2-(1,3-диоксо-1,3-дигидро-2H-изоиндол-2-ил)-5-оксогексаноат	4,27	4,29	0,02	0,47
22	4-[5-(4-Нитрофенил)-2-фурил]-N-фенил-1,3-тиазол-2-амин	4,16	4,09	0,07	1,68
23	8-Метокси-1H-бензо[а]карбазол-1,4(1H)-дион	4,23	4,29	0,06	1,42
24	3-(4-Метоксифенил)-3,4-дигидро-2H-хромен-6,7-диил диацетат	3,96	3,94	0,02	0,51
25	3-(3-Азабицикло[3.2.2]нон-3-ил)-1-(2-тиенил)-1-пропанон	4,18	4,13	0,05	1,20
26	Серебро(4-аминофенил)сульфонил-(2H-пиримидин-1-ID-2-ил)аммония	4,22	4,17	0,05	1,18
27	2,2-Дибутил-1,3,2-оксатиастаннолан-5-он	5,30	5,68	0,38	7,17
28	4-[(5,5,8,8-Тетраметил-5,6,7,8-тетрагидро-2-нафталинилкарбамоил)бензойная кислота	4,05	4,15	0,10	2,47
29	1,3-Бис(2-хлорэтил)-1-нитрозомочевина	3,47	3,62	0,15	4,32
30	3-Фенилпирроло[2,1-a]изохинолин-2-карбоксамид	4,08	4,10	0,02	0,49

В целом, по данным таблицы 2, где представлен широкий спектр органических соединений разных классов, относительная ошибка для большей части соединений не превышает 5 %, что в области прогнозирования биологической активности считается превосходным результатом. Исходя из этого, точность прогнозирования следует считать высокой.

Выводы

1. Предпринято исследование возможности одновременного моделирования параметра $\lg(LC_{50})$ противоопухолевой активности больших наборов химических соединений разных классов в отношении клеток человеческой меланомы SK-MEL-2 методом QSAR.

2. Моделирование $\lg(LC_{50})$ выполнено с помощью предложенного нами алгоритма, реализованного в программе PROGROC и позволяющего использовать число дескрипторов, превышающее число химических соединений.

3. Для набора из 250 органических веществ разных классов (тренировочная выборка включала 200 веществ, контрольная — 50) получен коэффициент корреляции между экспериментальными и вычисленными значениями $\lg(LC_{50})$ $R = 0,956$, стандартное отклонение $s = 0,12$.

4. Полученные нами результаты превосходят данные других авторов как по разнообразию структур химических соединений, так и по статистическим параметрам. Предложенные подходы могут быть применены для решения задачи прогнозирования противоопухолевой активности химических соединений.

Список литературы

1 Zhang S., Wei L., Bastow K., Zheng W., Bossi A., Lee K.-H., Tropsha A. Antitumor Agents 252. Application of validated QSAR models to database mining: discovery of novel tylophorine derivatives as potential anticancer agents // Journal of Computer-Aided Molecular Design. — 2007. — Vol. 21, No. 1. — P. 97-112.

- 2 *Georgea R.F., Ismailb N.S.M., Stawinska J., Girgis A.S.* Design, synthesis and QSAR studies of dispiroindole derivatives as new antiproliferative agents // *European Journal of Medicinal Chemistry*. — 2013. — Vol. 68. — P. 339–351.
- 3 *Katritzky A.R., Girgis A.S., Slavov S., Tala S. R., Stoyanova-Slavova I.* QSAR modeling, synthesis and bioassay of diverse leukemia RPMI-8226 cell line active agents // *European Journal of Medicinal Chemistry*. — 2010. — Vol. 45, No. 11. — P. 5183–5199.
- 4 *Zheng M., Yang Y., Zhao M., Zhang X., Wu J., Chen G., Peng L., Wang Y.* A class of novel N-isoquinoline-3-carbonyl-L-amino acid benzylesters: Synthesis, anti-tumor evaluation and 3D QSAR analysis // *European Journal of Medicinal Chemistry*. — 2011. — Vol. 46, No. 5. — P. 1672–1681.
- 5 *Rao G.-W., Wang C., Wang J., Zhao Z.-G., Hu W.-X.* Synthesis, structure analysis, antitumor evaluation and 3D-QSAR studies of 3,6-disubstituted-dihydro-1,2,4,5-tetrazine derivatives // *Bioorganic & Medicinal Chemistry Letters*. — 2013. — Vol. 23, № 23. — P. 6474–6480.
- 6 *Zhu W., Liu Y., Zhai X., Wang X., Zhu Y., Wu D., Zhou H., Gong P., Zhao Y.* Design, synthesis and 3D-QSAR analysis of novel 2-hydrazinyl-4-morpholinothieno[3,2-d]pyrimidine derivatives as potential antitumor agents // *European Journal of Medicinal Chemistry*. — 2012. — Vol. 57. — P. 162–175.
- 7 *Bababdani B.M., Mousavi M.* Gravitational search algorithm: A new feature selection method for QSAR study of anticancer potency of imidazo[4,5-b]pyridine derivatives // *Chemometrics and Intelligent Laboratory Systems*. — 2013. — Vol. 122. — P. 1–11.
- 8 *Sabet R., Mohammadpour M., Sadeghi A., Fassihia A.* QSAR study of isatin analogues as in vitro anti-cancer agents // *European Journal of Medicinal Chemistry*. — 2010. — Vol. 45. — P. 1113–1118.
- 9 *Masand V.H., Mahajan D.T., Alafeefy A.M., Bukhari S.N.A., Elsayedd N.N.* Optimization of antiproliferative activity of substituted phenyl 4-(2-oxoimidazolidin-1-yl)benzenesulfonates: QSAR and CoMFA analyses // *European Journal of Pharmaceutical Sciences*. — 2015. — Vol. 77. — P. 230–237.
- 10 *Basu A., Sinha B.N., Saiko P., Graser G., Szekeres T.* N-Hydroxy-N'-aminoguanidines as anti-cancer lead molecule: QSAR, synthesis and biological evaluation // *Bioorganic & Medicinal Chemistry Letters*. — 2011. — Vol. 21, No. 11. — P. 3324–3328.
- 11 *Vazhev V.* Prediction of olefin IR spectra reasoning from their mass spectra // *Journal of Structural Chemistry*. — 2005. — Vol. 46, № 2. — P. 243–247.
- 12 National Cancer Institute. Developmental Therapeutics Program // [ER]. Access mode: https://dtp.cancer.gov/databases_tools/default.htm

В.В. Важев, М.А. Губенко, Н.В. Важева, Э.М. Ергалиева

Химиялық қосылыстардың ісік ауруларына қарсы белсенділігін QSAR әдісімен болжау

Мақалада SK-MEL-2 адам меланомасы жасушаларына қатысты химиялық қосылыстардың көптеген жиынтықтарының ісік ауруларына қарсы белсенділігін бір мезгілде модельдеу үшін QSAR әдісін қолдану мүмкіндігі зерттелген. Әр түрлі кластардың 250 органикалық заттарынан тұратын жиынтық үшін эксперименттік және есептелген мәндер арасындағы корреляциялық коэффициент $\lg(LC_{50})$ $R = 0,956$ шамасында болды.

V.V. Vazhev, M.A. Gubenko, N.V. Vazheva, E.M. Ergaliyeva

Prediction of the antitumor activity of chemical compounds by QSAR

The possibility of using QSAR methods for simultaneous simulation antitumor activity of large sets of chemical compounds against human melanoma cells SK-MEL-2. For a set of 250 different classes of organic compounds reach the correlation coefficient between the experimental and calculated values of $\lg(LC_{50})$ $R = 0,956$.