

О.Т. Девіняк², Д.Я. Гаврилюк¹, Б.С. Зіменковський¹, Р.Б. Лесик¹

QSAR АНАЛІЗ 2(4)-ТІАЗОЛІДИНОНІВ ІЗ ПІРАЗОЛІНОВИМ ФРАГМЕНТОМ В МОЛЕКУЛАХ, ЩО ПРОЯВЛЯЮТЬ ПРОТИПУХЛИННУ АКТИВНІСТЬ ЩОДО КЛІТИН НЕДРІБНОКЛІТИННОГО РАКУ ЛЕГЕНЬ *IN VITRO*

Львівський національний медичний університет імені Данила Галицького¹,
м. Львів, Україна

Державний вищий навчальний заклад «Ужгородський національний університет»²,
м. Ужгород, Україна

e-mail: dr_r_lesyk@org.lviv.net

Резюме: Здійснено QSAR-аналіз протиракової активності 2(4)-піразолінзаміщених 4(2)-тіазолідинонів на клітинах недрібноклітинного раку легень, що може бути теоретичним підґрунтям для раціонального дизайну «лікоподібних молекул» з протираковою дією. Одержано однопараметричну модель, точність, надійність та прогностична здатність якої підтверджена статистичними критеріями. Встановлено, що інгібування росту клітин недрібноклітинного раку легень сполуками досліджуваного ряду зумовлена характерним поєднанням атомів у молекулярній структурі.

Ключові слова: 2(4)-тіазолідинони, піразоліни, QSAR аналіз, протиракова активність.

Вступ. Одним з ключових проектів, котрі реалізуються на кафедрі фармацевтичної, органічної і біоорганічної хімії ЛНМУ імені Данила Галицького є пошук нових антимітотичних агентів серед похідних 4-тіазолідинонів^{2,16}. Результати експериментальних досліджень дозволили ідентифікувати значний протираковий потенціал 5-ариліден-2(4)-імінотіазолідинонів-4(2)^{4,7}, 4-тіазолідинон-3-алканкарбонових кислот^{14,15}, конденсованих систем з тіазологіопірановим фрагментом⁶, неконденсованих похідних з 4-тіазолідиноновим та піразоліновим¹⁹ чи бензтіазольним фрагментами¹⁸, тощо. На основі *in silico* досліджень сформовано ряд критеріїв для спрямованого синтезу *de novo* потенційних «лікоподібних» молекул з групи гетерилзаміщених 4-тіазолідинонів³, зокрема з піразоліновим фрагментом¹.

Мета дослідження. Продовжуючи розвиток зазначеної тематики, у даній статті здійснено моделювання кількісного зв'язку між структурою 2(4)-піразолінзаміщених 4(2)-тіазолідинонів та інгібуванням росту клітин недрібноклітинного раку легень з метою виявити ключові властивості молекул, що зумовлюють протиракову активність, та створити модель для подальшого спрямованого синтезу біологічно активних речовин.

Матеріали та методи дослідження. Для проведення QSAR-досліджень відібрано 54 структури з групи 2(4)-піразолінзаміщених 4(2)-тіазолідинонів, протиракова актив-

ність *in vitro* яких досліджувалась за програмою *DTP* в Національному інституті раку (США) і виражена у формі відсотків росту різних культур ракових клітин^{8,10,12,17}. За результатами прескринінгу ряд сполук відібрані для детальних досліджень, які полягали у вивченні активності речовин на 60 лініях ракових клітин у мінімум 5 концентраціях при 10-кратному розведенні. Порівняння результатів претестування із даними ґрунтового скринінгу при концентрації 10⁻⁵ моль/л надає інформацію про точність проведення експерименту, що використано для визначення граничної точності майбутньої моделі та для вибору залежної змінної. З цією метою обчислювались корінь середньоквадратичної похибки (*RMSE*) та пояснювана розбіжність $Q^2 = 1 - \frac{\sum_{i=1:n} (Y_{pretest_i} - Y_{test_i})^2}{\sum_{i=1:n} (Y_{pretest_i} - Y_{pretest_i})^2}$ як для різних лі-

ній ракових клітин, так і для усереднених значень активності для кількох ліній недрібноклітинного раку легень. Клітинні лінії, які досліджувались не для всіх сполук, виключені з подальшого аналізу. Молекулярна геометрія сполук оптимізована в силовому полі *MMF94x* за допомогою програмного пакету *MOE*⁹. Більше 1500 дескрипторів молекулярної будови обчислені за допомогою онлайн-сервісу *E-Dragon*²². Надалі всі процедури здійснювались з використанням середовища математичної розробки *MATLAB*²⁰. Після видалення константних та попарно коре-

льованих ($R > 0.95$) дескрипторів здійснювався розподіл сполук на навчальну (для побудови та відлагодження моделі, 67%) та тестову (для здійснення зовнішньої валідації, 33%) вибірки. З метою об'єктивного відображення загальної сукупності структур і навчальною, і тестовою вибіркою здійснено кластерний аналіз на основі Евклідової відстані між молекулярними структурами, кожна з яких представлена точкою у N -вимірному просторі, координати якої задаються молекулярними дескрипторами. Кількість кластерів визначено як $\sqrt{n} \approx 7$. Відбір сполуки у тестову вибірку здійснювався таким чином: 1) з кожного кластера 33% сполук, вибраних випадковим чином, потрапляли у тестову вибірку, 2) якщо кластер містить 2 і менше сполуки, то жодна з них не потрапляє в тестову вибірку. Решта структур склали навчальну вибірку. З метою зменшення розміру аналізованої матриці дескрипторів, процедуру видалення константних та попарно корельованих дескрипторів для сполук, що ввійшли до навчальної вибірки, проведено повторно.

Побудова лінійних моделей здійснювалась методом покрокової регресії при послідовному впровадженні кожного з дескрипторів як стартового у моделі. Після цього проводився пошук «сполук-вискочок» (*outliers*) та визна-

чення домену застосовності шляхом побудови графіка із значеннями впливу як абсциса та значеннями стандартної похибки як ордината (графік *Вільямса*)²¹. Загалом визначалися наступні статистичні параметри моделі: розмір навчальної вибірки (N), квадрат коефіцієнта кореляції (R^2), приведений квадрат коефіцієнта кореляції (R^2_{adj}), критерій *Фішера* (F), статистична значимість (p), похибка регресії (корінь середньоквадратичної похибки, $RMSE$). Валідація моделі здійснювалась як за допомогою тестової вибірки (зовнішня), так і перехресна (внутрішня). Внутрішня перехресна валідація проводилась методами *leave-one-out* та *leave-many-out* (25%)¹³. Оцінка кожного процесу валідації здійснювалась шляхом обчислення пояснюваної розбіжності Q^2 та похибки валідації (корінь середньоквадратичної похибки передбачення, $RMSEP$).

Результати дослідження та їх обговорення. Внаслідок відсутності даних про відсотки інгібування клітин лінії HOP-62 для 8 сполук, дані щодо цієї лінії виключені з розгляду. Проведений аналіз точності експериментальних активностей виявив значні відхилення в значеннях відсотків інгібування росту при проведенні повторних досліджень (табл. 1).

Таблиця 1. Параметри, що характеризують точність (повторюваність) протипухлинної активності визначеної за програмою *DTP*

Клітинна лінія	RMSE	Q ²
A549/ATCC	35.58	0.452
EKVX	33.98	0.292
HOP-92	45.73	-0.177
NCI-H226	33.17	0.784
NCI-H23	31.62	0.649
NCI-H322M	26.93	0.611
NCI-H460	27.11	0.772
NCI-H522	53.72	-0.944
Середнє значення	22.66	0.772
Середнє значення без врахування лінії NCI-H522	22.37	0.806
Середнє значення без врахування ліній NCI-H522 та HOP-92	23.79	0.785
Середнє значення без врахування ліній NCI-H522, HOP-92 та EKVX	22.96	0.819
Середнє значення без врахування ліній NCI-H522, HOP-92, EKVX та A549/ATCC	22.28	0.836

Негативні значення пояснюваної розбіжності для ліній NCI-H522 та HOP-92 означають те, що сумарна розбіжність результатів при повторенні експерименту перевищує сумарне відхилення від середньої активності для даного ряду сполук. Загалом 4 із 8 ліній показують збіжність $Q^2 < 0.5$, тобто є малопридатними для моделювання. Для того, щоб знайти найбільш придатне для моделювання скалярне представлення протипухлинної активності сполуки на лініях недрібноклітинного раку легень, обчислено ряд середніх активностей, поступово не враховуючи най-

менш прогнозовані лінії. Середні значення активності характеризуються значно меншою похибкою та вищим відсотком пояснюваної розбіжності, що є очевидним наслідком теорії про центральну границю. Причому спостерігається незначне покращення статистичних параметрів при видаленні слабо прогнозованих ліній. Таким чином, в якості залежної змінної для створення QSAR моделі вибрано усереднену активність за чотирма лініями: NCI-H226, NCI-H23, NCI-H322M і NCI-H460, яку надалі позначено *GP* (*Growth Percent*). Також, згідно отриманих даних, по-

хибка моделі повинна бути не меншою від граничної $RMSE=22.28$. Отримання більш точної моделі свідчить про «перенавчання» (*overfitting*), і така модель потребує спрощення.

Для максимального використання наявної інформації та отримання адекватної оцінки прогностичної здатності моделі необхідним є рівномірний як за структурами, так і за активностями розподіл сполук на навчальну та тестову вибірки. Рівномірний розподіл за молекулярною будовою досягнуто за допомогою кластерного аналізу, при цьому розмір навчальної вибірки склав 36, а тестової 18 сполук. Рівномірність розподілу за активностями перевірено шляхом обчислення середніх значень та стандартних відхилень, що відповідно становили 76.55 і 46.86 для навчальної вибірки та 78.16 і 50.76 для тестової вибірки. Схожість середніх значень та стандартних відхилень у вибірках свідчить про оптимальність проведеного розподілу.

Кількість дескрипторів, що залишилися після видалення константних та попарно корельованих, склала 619. Враховуючи невисоку точність вихідних даних та дотримуючись принципу «леза Оккама», апріорно встанов-

лено обмеження щодо кількості незалежних змінних – модель повинна містити не більше двох дескрипторів. Таке обмеження дозволило без використання стохастичних алгоритмів вибору дескрипторів перевірити всі ($C_2^{619} = 191271$) можливі комбінації.

У результаті проведених досліджень одержано двопараметричну лінійну модель (1) із дескрипторами SRW09 і nRCONR2 ($R^2=0.7661$, $RMSE=23.670$), однак наступний аналіз домену застосовності та пошук екстремальних значень за допомогою графіка Вільямса (рис. 1) виявив значне відхилення сполуки 23 із домену застосовності. Дескриптор nRCONR2 (що вказує на наявність карбоксиамідного фрагменту) потрапив до моделі (1) тільки через те, що пояснював відсутність активності цієї сполуки наявністю в ній вказаної групи. Однак, не можна достовірно стверджувати про вплив карбоксиамідної групи на зниження активності, оскільки вона присутня тільки в одній молекулі. Крім того, похибки сполук 5 і 6 перевищували 3 $RMSE$, тому ці сполуки визначено як «вискочки» (*outliers*).

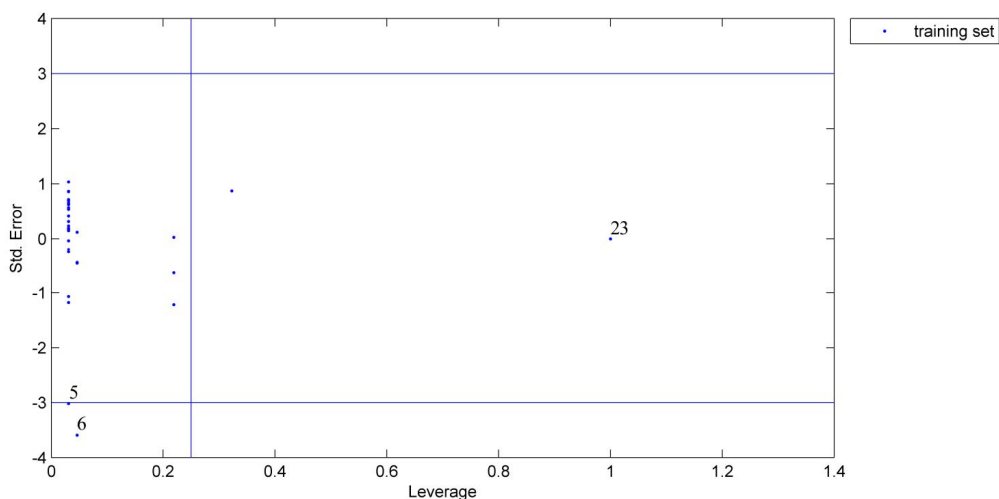


Рис. 1. Графік Вільямса для навчальної вибірки. Аналіз домену застосовності

Після видалення вказаних вище трьох сполук із навчальної вибірки одержано наступну спрощену однопараметричну модель (2),

$$GP = 296.74 (\pm 23.815) - 0.1097 (\pm 0.0118) * 'SRW09'; \quad (2)$$

$$N = 33, R^2 = 0.9211, R^2_{adj} = 0.9186, RMSE = 13.546, p < 10^{-15}, F = 361.95.$$

результати валідації якої наведено в таблиці 2:

Таблиця 2.

Результати валідації моделі (2)

Метод валідації	Q^2	$RMSEP$
Внутрішня перехресна, <i>LOO</i>	0.9032	14.543
Внутрішня перехресна, <i>LMO(25%)</i>	0.9142	13.6791
Зовнішня, $N=18$	0.8186	21.621

Видалення сполук поза доменом застосовності та «сполук-вискочок» із навчальної ви-

бірки дає можливість розраховувати на деяке зменшення наведеного вище мінімального

$RMSE$ при оцінці результатів QSAR аналізу. Тим не менше, отримані занадто хороші статистичні параметри моделі у поєднанні із значно нижчими, але зате наближеними до гранично допустимої мінімальної похибки (тобто наближеними до реальної ситуації прогнозу активності) свідчать про деяку «перетренованість». Однак подальше спрощення будь-якої однопараметричної лінійної моделі є неможливим.

Для встановлення ступеня ймовірності випадкового характеру отриманої моделі проведено Y -скрамблінг тест¹¹. Шляхом випадкового перемішування значень залежної змінної з навчальної вибірки отримано 10 тестових рядів. Для кожного з рядів аналогічно здійснювалась покрокова регресія, і фіксувалась похибка регресії ($RMSE_{test}$). Отримані результати (середнє значення похибки рег-

ресії $RMSE_{test} = 41.866$, стандартне відхилення $SD_{test} = 1.7704$) дозволяють заперечити випадковий характер отриманої моделі.

Важливим кроком у моделюванні біологічної активності є інтерпретація отриманої моделі у рамках хіміко-біологічних уявлень. Дескриптор $SRW09$ (*Self-Returning Walk count of 9th order*), що ввійшов до даної моделі, відноситься до топологічних дескрипторів і являє собою сумарну кількість всіх можливих самоповертаючих обходів 9-го порядку²¹. Для обчислення даного дескриптора структура сполуки представляється у вигляді графа. Самоповертаючий обхід 9-го порядку являє собою такий обхід (послідовність з'єднаних між собою вершин), що починається із довільної вершини і закінчується на ній же після 9-го кроку (рис. 2).

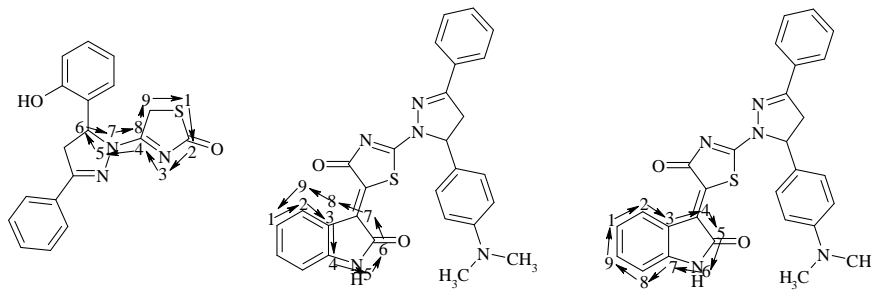


Рис. 2. Приклади самоповертаючих обходів 9-го порядку

Тобто здатність до інгібування росту клітин недрібноклітинного раку легень для деяких похідних 4-тіазолі-дону з піразоліновим фра-

гментом зумовлена саме характерним розташуванням атомів (топологією молекули).

Висновки:

1. Здійснено дослідження кількісного зв'язку між структурою похідних 4-тіазолі-дону з піразоліновим фрагментом та відсотком інгібування росту клітин недрібноклітинного раку легень, що може бути використано для раціонального дизайну «лікоподібних молекул» з протираковою дією.
2. Проведений QSAR-аналіз дозволив одержати однопараметричну модель залежності протиухлинної активності від топологічного дескриптора $SRW09$ (*Self-Re-*

turning Walk count of 9th order). Точність, надійність та прогностична здатність одержаної моделі підтверджена статистичними критеріями.

3. Встановлено, що інгібування росту клітин недрібноклітинного раку легень сполуками досліджуваного ряду зумовлена характерним поєднанням атомів у молекулярній структурі (топологією молекули).

Література:

1. Докінгові дослідження та QSAR аналіз 4-(3,5-діарил-4,5-дигідро-1H-піразол-1-іл)-1,3-тіазол-2(5H)-онів, які проявляють протиракову активність *in vitro* / Д.Я. Гаврилюк, Б.С. Зіменковський, І.В. Драпак [та ін.] // Фармацевтичний журнал. – 2009. – №6. – С. 69–75.
2. Лесик Р.Б. Синтез та біологічна активність конденсованих і неконденсованих гетероциклічних систем

на основі 4-азолідонів: Автореф. дис. на здобуття наук. ступеня д-ра фарм. наук: 15.00.02 «фармацевтична хімія та фармакогнозія» / Р.Б. Лесик. – Львів, 2005. – 39с.

3. Протиухлинна активність та QSAR аналіз похідних роданину з бензтіазольним фрагментом у молекулах / Л.М. Мосула, Б.С. Зіменковський, В.В. Огу-

- рцов [та ін.] // Фармацевтичний журнал. – 2010. – №2. – С. 77–83.
- Синтез 5-ариліденопохідних (2-метил-4-окси-5-ізопропілфенламіно)тіазолонів та вивчення їх протипухлинної активності / І.Ю. Субтельна, Б.С. Зіменковський, Анджей Гзелля [та ін.] // Клінічна фармація, фармакотерапія та медична стандартизація. – 2010. – №1–2. – С. 93–99.
 - Синтез та вивчення протипухлинної активності похідних 2-ариламіно-2-тіазолін-4-онів / Р.Б. Лесик, Б.С. Зіменковський, С.М. Ярмолюк [та ін.] // Фармацевтичний журнал. – 2003. – №1. – С. 51–56.
 - Anticancer thiopyrano[2,3-d]thiazol-2-ones with norbornane moiety. Synthesis, cytotoxicity, physico-chemical properties, and computational studies / R. Lesyk, B. Zimenkovsky, D. Atamanyuk [et al.] // Bioorg. Med. Chem. – 2006. – Vol. 14. – P. 5230–5240.
 - 5-Arylidene-2-amino-4-azolones and evaluation of their anticancer activity / I. Subtel'na, D. Atamanyuk, E. Szymanska [et al.] // Bioorg. Med. Chem. – 2010. – Vol.18, №14. – P. 5089–5101.
 - Boyd M.R. Some Practical Considerations and Applications of the National Cancer Institute In Vitro Anticancer Drug Discovery Screen / M.R. Boyd, K.D. Paull // Drug Development Research. – 1995. – №34. – P. 91–109.
 - Chemical Computing Group Inc. MOE 2007.09 (Molecular Operating Environment software). [Електронний ресурс]. – Режим доступу: <http://www.chemcomp.com>
 - Feasibility of drug screening with panels of human tumor cell lines using a microculture tetrazolium assay / M.C. Alley, D.A. Scudiero, P.A. Monks [et al.] // Cancer Research. – 1988. – Vol.48. – P. 589–601.
 - Gramatica P. Principles of QSAR models validation: internal and external / P. Gramatica // QSAR & Comb. Sci. – 2007. – №5. – P. 694–701.
 - Grever M.R. The National Cancer Institute: cancer drug discovery and development program / M.R. Grever, S.A. Schepartz, B.A. Chabner // Seminars in Oncology. – 1992. – №6. – P. 622–638.
 - Guidance document on the validation of (quantitative) structure-activity relationships [(Q)SAR] models / OECD Environment health and safety publications, series on testing and assessment. [Електронний ресурс]. – 2007. – №69. – 154 p. – Режим доступу: http://www.oecd.org/document/4/0,3343,en_2649_34379_42926724_1_1_1_1,00.html.
 - Kaminsky D.V. Structure – anticancer activity relationships among 4-azolidinone-3-carboxylic acids derivatives / D.V. Kaminsky, R.B. Lesyk // Biopolymers and Cell. – 2010. – №2. – P. 136–145.
 - Kaminsky D. Synthesis and *in vitro* anticancer activity of 2,4-azolidinedione-acetic acids derivatives / D. Kaminsky, B. Zimenkovsky, R. Lesyk // Eur. J. Med. Chem. – 2009. – №9. – P. 3627–3636.
 - Lesyk R. 4-Thiazolidones: Centenarian history, current status and perspectives for modern organic and medicinal chemistry / R. Lesyk, B. Zimenkovsky // Current Organic Chemistry. – 2004. – №16. – P. 1547–1578.
 - Shoemaker R.H. The NCI60 human tumour cell line anticancer drug screen / R.H. Shoemaker // Nature Reviews Cancer. – 2006. – №6. – P. 813–823.
 - Synthesis and anticancer activity evaluation of 4-thiazolidinones containing benzothiazole moiety / D. Havrylyuk, L. Mosula, B. Zimenkovsky [et al.] // Eur. J. Med. Chem. – 2010. – №11. – P. 5012–5021.
 - Synthesis of novel thiazolone-based compounds containing pyrazoline moiety and evaluation of their anticancer activity / D. Havrylyuk, B. Zimenkovsky, O. Vasylenko [et al.] // Eur. J. Med. Chem. – 2009. – №4. – P. 1396–1404.
 - The MathWorks, Inc. MATLAB(R2008b) (The Language of Technical Computing), Version 7.7.0.471. [Електронний ресурс]. – Режим доступу: <http://www.mathworks.com>
 - Todeschini R. Molecular descriptors for chemoinformatics / R. Todeschini, V. Consonni. – Weinheim (Germany), WILEY-VCH, 2009. – 1257 p.
 - Virtual computational chemistry laboratory – design and description / I.V. Tetko, J. Gasteiger, R. Todeschini [et al.] // J. Comput. Aid. Mol. Des. – 2005. – Vol. 19. – P. 453–463.

УДК 615.076:547.789

QSAR АНАЛИЗ 2(4)-ТІАЗОЛІДИНОНОВ С ПИРАЗОЛИНОВЫМ ФРАГМЕНТОМ В МОЛЕКУЛАХ, ЧТО ПРОЯВЛЯЮТ ПРОТИВООПУХОЛЕВУЮ АКТИВНОСТЬ ПО ОТНОШЕНИЮ К КЛЕТКАМ НЕМЕЛКОКЛЕТОЧНОГО РАКА ЛЕГКИХ IN VITRO

О.Т. Девиняк², Д.Я. Гаврилюк¹, Б.С. Зіменковський¹, Р.Б. Лесык¹

Львовский национальный медицинский университет имени Данила Галицкого¹, г. Львов, Украина

Государственное высшее учебное заведение «Ужгородский национальный университет»², г. Ужгород, Украина

Резюме: Осуществлен QSAR анализ противоопухолевой активности 2(4)-пиразолинзамещенных 4(2)-тиазолидинонов на клетках немелкоклеточного рака легких. Точность, надежность и предсказательная способность полученной однопараметрической модели подтверждена статистическими критериями. Установлено, что ингибирование роста клеток немелкоклеточного рака легких соединениями исследуемого ряда обусловлена характерным сочетанием атомов в молекулярной структуре. Текущие результаты могут быть теоретическим основанием для рационального дизайна «лекарствоподобных» веществ с противоопухолевой активностью.

Ключевые слова: 2(4)-тиазолидиноны, пиразолины, QSAR анализ, противоопухолевая активность.

UDC 615.076:547.789

QSAR STUDY OF 2(4)-THIAZOLIDINONES WITH PYRAZOLINE SCAFFOLD POSSESING ANTITUMOR ACTIVITY *IN VITRO* ON NON-SMALL LUNG CANCER CELLS

O.T. Devinyak², D.Ya. Havrylyuk¹, B.S. Zimenkovsky¹, R.B. Lesyk¹

Danylo Halytsky Lviv National Medical University¹, Lviv, Ukraine

State Higher Educational Institution «Uzhgorod National University»², Uzhgorod, Ukraine

Summary: QSAR analysis of antitumor activity of 2(4)-pyrazoline substituted 4(2)-thiazolidinones on non-small lung cancer cells was carried out which can be a theoretical basis for rational design of drug-like molecules with anticancer effect in a number of related systems. Monoparametric model was developed and its goodness-of-fit, robustness and predictivity were confirmed by statistical criteria. It was discovered that special mutual location of atoms in 2(4)-thiazolidinone structure defines its inhibition of the tumor cells growth.

Keywords: 2(4)-thiazolidinones, pyrazolines, QSAR analysis, antitumor activity.

Надійшла до редакції 22.12.2011 р.