

<https://doi.org/10.15407/dopovidi2024.01.050>

УДК 544.(16; 332.031):546

Л.М. Огніченко, <https://orcid.org/0000-0002-3344-6119>

А.Г. Артеменко, <https://orcid.org/0000-0002-9592-1384>

М.Є. Кічьова, <https://orcid.org/0000-0002-6294-3261>

С.І. Стельмах, <https://orcid.org/0009-0005-2023-3121>

В.Ф. Зінченко, <https://orcid.org/0000-0003-4778-495X>

В.Є. Кузьмін, <https://orcid.org/0000-0002-2753-0453>

Фізико-хімічний інститут ім. О.В. Богатського НАН України, Одеса, Україна

E-mail: ogni@ukr.net, theorchem@gmail.com

QSPR-модель для прогнозування стандартної ентальпії утворення складних оксигеновмісних неорганічних сполук

Досліджено вибірку з 74 оксигеновмісних неорганічних сполук — складних змішаних оксидів і солей, деякі з яких є перспективними компонентами оптичних матеріалів. Із застосуванням 1D-симплексного підходу побудовано адекватну консенсусну QSPR-PLS-модель для стандартної ентальпії утворення (ΔH_{298}^0) досліджуваних сполук. Під час розрахунку 1D-симплексних дескрипторів (одиниць, двійок, трійок, четвірок) для диференціації вершин (атомів) у симплексах використовували різні характеристики атомів з періодичної системи Д.І. Менделєєва, ступінь окиснення, електронегативність, іонний радіус, характеристики ван-дер-ваальсових взаємодій. Для тестового набору розробленої моделі коефіцієнт детермінації (R_{test}^2) дорівнює 0,94, середня відносна похибка прогнозу становить 10,0 %. У результаті аналізу структурних параметрів, які увійшли до розробленої моделі, виявлено, що найістотніший вплив виявляють орбітальні квантові числа електронів зовнішнього шару атомів (28 %) і ван-дер-ваальсові взаємодії (19 %). Побудовану модель можна використовувати для попередньої оцінки стандартної ентальпії утворення різноманітних оксигеновмісних неорганічних сполук.

Ключові слова: стандартна ентальпія утворення, симплексне представлення молекулярної структури, 1D-симплексний підхід, QSPR-модель, метод часткових найменших квадратів.

Дослідження зв'язку між структурою і властивостями різноманітних речовин є однією з центральних фундаментальних проблем хімії в цілому. В останні десятиліття створення нових речовин (матеріалів, реагентів, ліків тощо) практично завжди пов'язано з дослі-

Ц и т у в а н н я: Огніченко Л.М., Артеменко А.Г., Кічьова М.Є., Стельмах С.І., Зінченко В.Ф., Кузьмін В.Є. QSPR-модель для прогнозування стандартної ентальпії утворення складних оксигеновмісних неорганічних сполук. *Допов. Нац. акад. наук Укр.* 2024. № 1. С. 50—57. <https://doi.org/10.15407/dopovidi2024.01.050>

© Видавець ВД «Академперіодика» НАН України, 2024. Стаття опублікована за умовами відкритого доступу за ліцензією CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

дженнями із застосуванням моделей QSAR / QSPR (quantitative structure-property relationship / quantitative structure-activity relationship). На сьогодні сформувалася методологія вирішення таких завдань, основні принципи і правила її реалізації [1]. Так склалося, що переважна більшість таких досліджень стосується насамперед органічних сполук. Зокрема, велика частина QSAR-моделей спрямована на створення нових лікарських препаратів [2].

Очевидно, що органічні сполуки мають структурну специфіку, яка відрізняє їх від неорганічних сполук. Так, органічні сполуки мають обмежений елементний склад (атоми-органогени), натомість неорганічні сполуки базуються практично на всіх елементах періодичної системи Д.І. Менделєєва. Крім того, більшість класичних неорганічних сполук є менш різноманітними з точки зору топології атомного кістяка (молекулярного графу). Ці фундаментальні відмінності, природно, мають визначати специфіку моделювання структури неорганічних сполук.

Розроблення і реалізація нових QSPR підходів для дослідження неорганічних сполук є достатньо актуальною проблемою. Актуальність і перспективність таких досліджень обумовлені широкими можливостями цієї наукової галузі в отриманні фундаментальних та прикладних результатів за рахунок синергетичного ефекту поєднання хімічних знань, сучасних інформаційних технологій та потужної обчислювальної техніки.

Раніше нами було розроблено 1D-симплексний підхід для опису класичних неорганічних сполук [3—5]. Показано, що незважаючи на досить примітивне подання хімічних сполук в 1D-QSAR/QSPR підході (фактично використовується інформація тільки щодо атомного складу сполуки) такий підхід може бути корисним для позаекспериментального скринінгу, наприклад, таких важливих властивостей компонентів потенціальних оптичних матеріалів, як показники заломлення і температури плавлення [5], а також для дослідження деяких видів біологічної активності [6].

Кількість інформації щодо будови сполуки в рамках 1D-QSPR підходу можна значно збільшити, застосовуючи різні комбінаторні процедури до брутто-формули (формули Хілла) сполуки. Як дескриптори використовують кількості різноманітних комбінацій атомів (від одиниць до четвірок), що входять до брутто-формули сполуки. При цьому атоми характеризуються мітками, які відображують їхнє положення в періодичній системі елементів, електронегативність та інші характеристики, що легко інтерпретуються. Відмінна особливість даного підходу полягає в тому, що він дає можливість у фрагментних дескрипторах мітки (символи) конкретних атомів (елементів) замінювати на мітки, що характеризують належність до певного інтервалу зміни заданої властивості атомів [3, 4], тобто емпіричну формулу будь-якої сполуки можна замінити на її аналог $A_a B_b C_c D_d$, де символи атомів замінено на мітки, що характеризують інтервали заданої властивості відповідних атомів. Для такої формули можна оцінити кількість різноманітних фрагментів — одиниць, двійок, трійок тощо.

Метою дослідження було із застосуванням 1D-симплексного підходу побудувати QSPR-модель для такої фундаментальної властивості, як стандартна ентальпія утворення (ΔH_{298}^0), яка є мірою термодинамічної стійкості сполук (за стандартних умов), кількісним вираженням їх енерговмісту.

Матеріали і методи. Об'єктом дослідження була вибірка з 74 оксигеновмісних неорганічних сполук — складних змішаних оксидів і солей, деякі з яких є перспективними компонентами оптичних матеріалів [7]. У таблиці наведено дані щодо стандартної ентальпії утворення досліджуваних сполук, зібрані з різних літературних джерел [8—10].

Стандартна ентальпія утворення (ΔH^0_{298}) оксигеновмісних сполук

№	Сполука	$-\Delta H^0_{298}$, кДж/моль		№	Сполука	$-\Delta H^0_{298}$, кДж/моль	
		Експ.	Прогноз.			Експ.	Прогноз.
1	Ba ₃ (PO ₄) ₂	3960	3927	38	Li ₂ WO ₄	1604	1437
2	BaCrO ₄	1426	1364	39	LiBO ₂	1018	1070
3	BaSO ₄	1458	1369	40	Mg ₃ (PO ₄) ₂	3790	3605
4	Ca(BO ₂) ₂	2031	1718	41	MgSO ₄	1280	1363
5	Ca(PO ₃) ₂	2462	2163	42	MnSiO ₃	1320	1069
6	Ca ₂ P ₂ O ₇	3333	2952	43	MnSO ₄	1066	1066
7	Ca ₂ SiO ₄	2308	2051	44	Na ₂ B ₄ O ₇	3291	3251
8	Ca ₃ (BO ₃) ₂	3429	4005	45	Na ₂ Cr ₂ O ₇	1967	2018
9	Ca ₃ (PO ₄) ₂	4113	4590	46	Na ₂ CrO ₄	1336	1467
10	CaCrO ₄	1379	1290	47	Na ₂ S ₂ O ₇	1938	2208
11	CaSiO ₃	1630	1498	48	Na ₂ Si ₂ O ₅	2474	2561
12	CaSO ₄	1435	1613	49	Na ₂ SiO ₃	1588	1524
13	CaWO ₄	1684	1277	50	Na ₂ SO ₄	1388	1494
14	CdCrO ₄	957	1183	51	Na ₂ WO ₄	1547	1418
15	CdSiO ₃	1190	1121	52	Na ₃ PO ₄	1917	1718
16	CdSO ₄	934	1040	53	Na ₃ VO ₄	1757	1783
17	CoCrO ₄	938	947	54	Na ₄ P ₂ O ₇	3166	3668
18	CoSO ₄	889	1129	55	Na ₄ SiO ₄	2106	1994
19	Cs ₂ CrO ₄	1429	1437	56	Na ₄ V ₂ O ₇	2920	2723
20	Cs ₂ SO ₄	1443	1494	57	NaBO ₂	976	1025
21	CuCrO ₄	815	1052	58	NaPO ₃	1220	1171
22	CuSO ₄	771	954	59	NaVO ₃	1148	1124
23	Fe ₂ SiO ₄	1478	1279	60	NiCrO ₄	935	1129
24	FeCrO ₄	969	923	61	NiSO ₄	874	945
25	FeSiO ₃	1195	1057	62	PbCrO ₄	915	1025
26	FeSO ₄	928	1002	63	PbSiO ₃	1137	1428
27	K ₂ B ₄ O ₇	3341	3112	64	PbSO ₄	920	1052
28	K ₂ Cr ₂ O ₇	2068	1862	65	PbWO ₄	1121	1098
29	K ₂ CrO ₄	1398	1576	66	Rb ₂ CrO ₄	1411	1432
30	K ₂ SiO ₃	1542	1643	67	Rb ₂ SO ₄	1436	1425
31	K ₂ SO ₄	1438	1655	68	Sb ₂ (SO ₄) ₃	2393	1923
32	K ₂ WO ₄	1580	1591	69	SnSO ₄	887	1141
33	KBO ₂	982	1127	70	SrCrO ₄	1435	1227
34	KPO ₃	1246	1299	71	SrSO ₄	1467	1423
35	Li ₂ B ₄ O ₇	3377	3193	72	ZnCrO ₄	992	1052
36	Li ₂ SiO ₃	1653	1418	73	ZnSiO ₃	1264	1052
37	Li ₂ SO ₄	1436	1562	74	ZnSO ₄	980	1129

Для кожної молекули були розраховані 1D-симплексні дескриптори (одиноці, двійки, трійки, четвірки), атоми в яких диференційовані за такими характеристиками з періодичної системи Д.І. Менделєєва: номер групи — NG (максимальне значення головного квантового числа), ступінь окиснення — StOk (від -2 до $+6$), заряд ядра — Zар (діапазон зміни значень на інтервали: $A < 25 \leq B < 45 \leq C < 65 \leq D$), належність до s -, p -, d -, f - елементів — SPDF (орбітальні квантові числа електронів зовнішнього шару атомів: 0—3).

Вершини симплексів були також диференційовані за електронегативністю (En) атомів (інтервали: $A < 1,7 \leq B < 2,7 \leq C < 3,0 \leq D$), іонним радіусом за Бокієм (IR): $A < 0,03 \leq B < 0,06 \leq C < 0,08 \leq D < 0,10 \leq E < 0,12 \leq F < 0,14 \leq G$, характеристиками ван-дер-ваальсових взаємодій (ВДВ) із параметрів універсального силового поля [11]. Ван-дер-ваальсові взаємодії між незв'язаними атомами в даному випадку описано формулою

$$E_{VDW} = \left[D_{ij} \cdot \left(\frac{6}{\zeta - 6} \right) \cdot e^{\zeta} \right] \cdot e^{-\zeta(x/x_{ij})} - \left[D_{ij} \cdot \left(\frac{\zeta}{\zeta - 6} \right) \cdot x_{ij}^6 \right] / x^6,$$

де ζ — параметр форми; D_{ij} — глибина потенційної ями, ккал/моль; x_{ij} — відстань 6—12. Звідси коефіцієнт ВДВ притягання розраховано за формулою $C_{Attr} = D_{ij} \cdot \left(\frac{\zeta}{\zeta - 6} \right) \cdot x_{ij}^6$, а коефіцієнт ВДВ відштовхування — за формулою $C_{Rep} = D_{ij} \cdot \left(\frac{6}{\zeta - 6} \right) \cdot e^{\zeta}$. Увесь діапазон значень параметрів ван-дер-ваальсових взаємодій було розділено на такі інтервали:

ВДВ притягання (Attr) — $A < 50 \leq B < 100 \leq C < 250 \leq D < 400 \leq E < 650 \leq F < 2000 \leq G$;
ВДВ відштовхування (Rep) — $A < 20000 \leq B < 32000 \leq C < 50000 \leq D < 100000$.

Крім вище зазначених дескрипторів під час побудови QSPR-моделі було використано такі параметри, як кількість атомів катіона, кількість атомів оксигену, інтегральні параметри: молекулярна маса (I.AW), електронегативність (I.En). Таким чином, для кожної молекули було розраховано 697 дескрипторів.

Зв'язок між розрахованими дескрипторами і досліджуваною властивістю встановлювали за допомогою методу часткових найменших квадратів (partial least squares (PLS)) [12], який ґрунтується на перетворенні великої кількості структурних параметрів у невелику кількість прихованих (латентних) змінних (LV). Використання PLS-методу дає змогу визначити кількісні характеристики впливу кожного дескриптора на властивість. Для вибору початкових наборів дескрипторів для розроблення PLS-моделі було вилучено постійні параметри та такі, що взаємно корелюють. Для побудови PLS-моделі також використано процедуру *trend vector* [13], автоматичного відбору змінних (AVS) [14] та генетичний алгоритм [15].

Для оцінки прогностичної здатності моделі виконано п'ятиразову зовнішню перехресну перевірку [14]. Для унеможливлення помилкової кореляції застосовано процедуру *Y-scrambling* — створення моделі методом випадкового змішування значень властивостей [14].

Як статистичні характеристики моделей використано такі параметри: R^2 , Q^2 , R^2_{test} — коефіцієнт детермінації для навчальної вибірки, в умовах крос-валідації та тестової вибірки відповідно; S_{ws} , S_{cv} , S_{test} — стандартна похибка прогнозу для навчальної, крос-валідації та тестової вибірок відповідно.

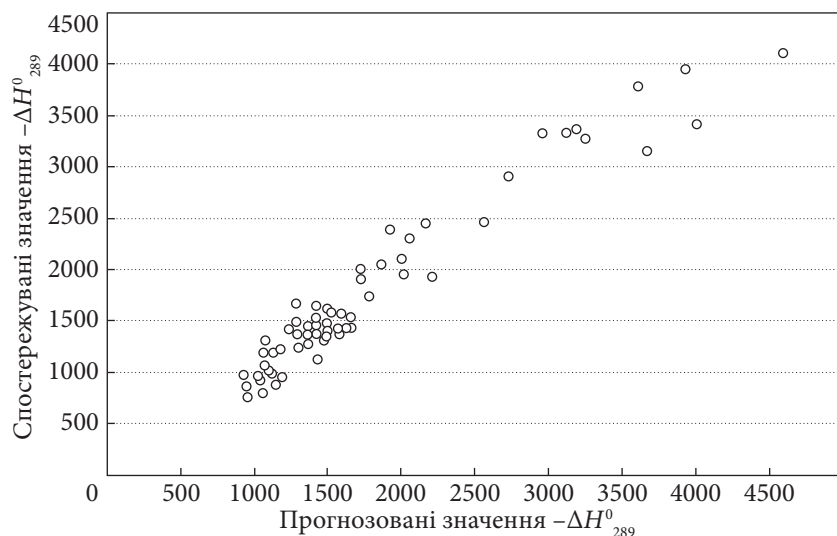


Рис. 1. Співвідношення прогнозованих і спостережуваних значень ΔH_{298}^0 для тестової вибірки (консенсусна модель)

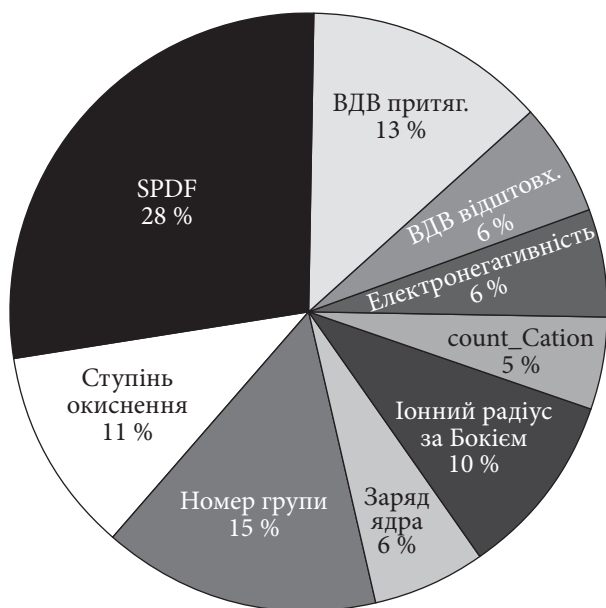


Рис. 2. Вплив деяких структурних чинників на стандартну ентальпію утворення

і тестову (20 % загальної кількості молекул) вибірки. У результаті побудовано адекватні 1D-QSPR-моделі, які, у свою чергу, було використано в основі консенсусної моделі. Прогнозовані значення стандартної ентальпії утворення для молекул, коли вони були введені до тестової вибірки, наведено в таблиці. У рамках консенсусної моделі для тестового набору коефіцієнт детермінації (R_{test}^2) дорівнює 0,94, а середня відносна похибка становить 10,0 %. На рис. 1 для консенсусної моделі проілюстровано співвідношення спостережуваних — прогнозованих значень ΔH_{298}^0 для тестової вибірки.

Результати і їх обговорення. На першому етапі дослідження всі молекули вибірки спочатку були включені в навчання. З набору дескрипторів було видалено малоінформативні. У результаті було побудовано адекватну QSPR-1D-PLS-модель з такими статистичними показниками: $LV = 2$; кількість дескрипторів: $N = 11$; $R^2 = 0,968$; $Q^2(\text{LOO}) = 0,964$; $S_{\text{ws}} = 148,4$; $S_{\text{cv}} = 160,5$. Величини статистичних характеристик моделей, отриманих у разі застосування процедури *Y-scrambling*, значно нижчі, ніж у робочій моделі: $R_{\text{Y-scr}}^2 = 0,056 \pm 0,01$, $Q_{\text{Y-scr}}^2 = 0,016 \pm 0,01$. Це свідчить про невипадковість встановленого зв'язку між структурою сполук і величиною їхньої стандартної ентальпії утворення ΔH_{298}^0 .

Щоб оцінити здатність моделі до прогнозування, було сформовано п'ять наборів, кожний з яких містив навчальну

Для структурних параметрів з найістотнішим впливом, що увійшли до консенсусної моделі, було визначено їх відносний внесок у стандартну ентальпію утворення (ΔH_{298}^0 , %):

Cnk(SPDF)/A, B, B	10,8
Cnk(SPDF)/A, B	9,5
Cnk(NG)/B	7,9
Cnk(SPDF)/A	6,9
Cnk(Zar)/A	6,0
count_Cation (кількість атомів катіона)	4,8
Cnk(Attr)/C, F, F	4,4
Cnk(Rep)/D	4,3

Тут Cnk — 1D-дескриптор, у дужках вказано властивість атомів, за якою атоми було диференційовано, після скісної риски наведено мітки атомів згідно з певним інтервалом значень властивості.

На основі аналізу структурних параметрів, які увійшли до консенсусної моделі, встановлено, що найбільший вплив виявляють орбітальні квантові числа електронів зовнішнього шару атомів (28 %). Крім цього, істотний вплив мають ван-дер-ваальсові взаємодії (19 %) і головні квантові числа (15 %) атомів елементів, що входять до складу оксидів (рис. 2).

Висновки. За допомогою PLS-методу із застосуванням 1D-дескрипторів побудовано адекватну консенсусну QSPR-модель, для якої R_{test}^2 дорівнює 0,94. Згідно з результатами аналізу структурних параметрів, які увійшли до розробленої моделі, найістотніший вплив виявляють орбітальні квантові числа електронів зовнішнього шару атомів (28 %) і ван-дер-ваальсові взаємодії (19 %). Побудовану модель можна використовувати для попередньої оцінки стандартної ентальпії утворення (ΔH_{298}^0) різноманітних оксигеновмісних неорганічних сполук.

ЦИТОВАНА ЛІТЕРАТУРА

- Cherkasov A., Muratov E.N., Fourches D., Varnek A., Baskin I.I., Cronin M., Dearden J., Gramatica P., Martin Y. C., Todeschini R., Consonni V., Kuz'min V.E., Cramer R., Benigni R., Yang C., Rathman J., Terflath L., Gasteiger J., Richard A., Tropsha A. QSAR modeling: where have you been? Where are you going to? *J. Med. Chem.* 2014. 57, № 12. P. 4977—5010. <https://doi.org/10.1021/jm4004285>
- Wang T., Yuan X.-S., Wu M.-B., Lin J.-P., Yang L.-R. The advancement of multidimensional QSAR for novel drug discovery — where are we headed? *Expert Opin. Drug Discov.* 2017. 12, № 8. P. 769—784. <https://doi.org/10.1080/17460441.2017.1336157>
- Kuz'min V., Artemenko A., Ognichenko L., Hromov A., Kosinskaya A., Stelmakh S., Sessions Z.L., Muratov E.N. Simplex representation of molecular structure as universal QSAR/QSPR tool. *Struct. Chem.* 2021. 32, № 4. P. 1365—1392. <https://doi.org/10.1007/s11224-021-01793-z>
- Кузьмін В., Артеменко А., Огніченко Л., Муратов Е., Хромов А. Симплексное представление молекулярной структуры. Саарбрюкен: Lambert Academic Publishing, 2022. 91 с.
- Kuz'min V.E., Ognichenko L.N., Zinchenko V.F., Artemenko A.G., Shyrykalova A.O., Kozhukhar A.V. QSPR models for prediction of the melting points and refractive indices for inorganic substances: components of the optical film-forming materials. *Int. J. Quant. Struct. Prop. Relatsh.* 2020. 5, № 1. P. 1—21. <https://doi.org/10.4018/IJQSPR.2020010101>
- Sizochenko N., Rasulev B., Gajewicz A., Kuz'min V., Puzyn T., Leszczynski J. From basic physics to mechanisms of toxicity: the “liquid drop” approach applied to develop predictive classification models for toxicity of metal oxide nanoparticles. *Nanoscale.* 2014. 6, № 22. P. 13986—13993. <https://doi.org/10.1039/C4NR03487B>

- Zinchenko V.F. Solid-phase complex compounds and composites of metal oxides, fluorides, and chalcogenides as materials for interference coatings: a review. *Theor. Exp. Chem.* 2021. **57**, № 4. P. 262—271. <https://doi.org/10.1007/s11237-021-09694-2>
- Лидин Р.А., Андреева Л.Л., Молочко В.А. Справочник по неорганической химии: Константы неорганических веществ. Москва: Химия, 1987. 320 с.
- Binnewies M., Milke E. Thermochemical data of elements and compounds. 2nd, rev. and ext. ed. Weinheim: Wiley-VCH, 2002. 928p.
- Ефимов А.И., Белорукова Л.П., Василькова И.В., Чечев В.П. Свойства неорганических соединений. Справочник. Ленинград: Химия, 1983. 392 с.
- Rappe A.K., Casewit C.J., Colwell K.S., Goddard III W.A., Skiff W.M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* 1992. **114**, № 25. P. 10024—10035. <https://doi.org/10.1021/ja00051a040>
- Rännar S., Lindgren F., Geladi P., Wold S. A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. *J. Chemometrics.* 1994. **8**, № 2. P. 111—125. <https://doi.org/10.1002/cem.1180080204>
- Carhart R.E., Smith D.H., Venkataraghavan R. Atom pairs as molecular features in structure—activity studies: definition and applications. *J. Chem. Inf. Comput. Sci.* 1985. **25**, № 2. P. 64—73. <https://doi.org/10.1021/ci00046a002>
- Kuz'min V.E., Artemenko A.G., Muratov E.N., Polischuk P.G., Ognichenko L.N., Liahovsky A.V., Hromov A.I., Varlamova E.V. Virtual screening and molecular design based on hierarchical QSAR technology. *Recent Advances in QSAR Studies*: Puzyn T., Leszczynski J., Cronin M. (eds.). London: Springer, 2010. P. 127—176. https://doi.org/10.1007/978-1-4020-9783-6_5
- Hasegawa K., Miyashita Y., Funatsu K. GA strategy for variable selection in QSAR studies: GA-based PLS analysis of calcium channel antagonists. *J. Chem. Inf. Comput. Sci.* 1997. **37**, № 2. P. 306—310. <https://doi.org/10.1021/ci960047x>

Надійшло до редакції 27.09.2023

REFERENCES

- Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., Dearden, J., Gramatica, P., Martin, Y. C., Todeschini, R., Consonni, V., Kuz'min, V. E., Cramer, R., Benigni, R., Yang, C., Rathman, J., Terfloth, L., Gasteiger, J., Richard, A. & Tropsha, A. (2014). QSAR modeling: where have you been? Where are you going to? *J. Med. Chem.*, 57, No. 12, pp. 4977-5010. <https://doi.org/10.1021/jm4004285>
- Wang, T., Yuan, X.-S., Wu, M.-B., Lin, J.-P. & Yang, L.-R. (2017). The advancement of multidimensional QSAR for novel drug discovery — where are we headed? *Expert Opin. Drug Discov.*, 12, No. 8, pp. 769-784. <https://doi.org/10.1080/17460441.2017.1336157>
- Kuz'min, V., Artemenko, A., Ognichenko, L., Hromov, A., Kosinskaya, A., Stelmakh, S., Sessions, Z. L. & Muratov, E. N. (2021). Simplex representation of molecular structure as universal QSAR/QSPR tool. *Struct. Chem.*, 32, No. 4, pp. 1365-1392. <https://doi.org/10.1007/s11224-021-01793-z>
- Kuz'min, V., Artemenko, A., Ognichenko, L., Muratov, E. & Hromov, A. (2022). Simplex representation of molecular structure. Saarbrücken: Lambert (in Russian).
- Kuz'min, V. E., Ognichenko, L. N., Zinchenko, V. F., Artemenko, A. G., Shyrykalova, A. O. & Kozhukhar, A. V. (2020). QSPR models for prediction of the melting points and refractive indices for inorganic substances: components of the optical film-forming materials. *Int. J. Quant. Struct. Prop. Relatsh.*, 5, No. 1, pp. 1-21. <https://doi.org/10.4018/IJQSPR.2020010101>
- Sizochenko, N., Rasulev, B., Gajewicz, V., Puzyn, T. & Leszczynski, J. (2014). From basic physics to mechanisms of toxicity: the “liquid drop” approach applied to develop predictive classification models for toxicity of metal oxide nanoparticles. *Nanoscale*, 6, No. 22, pp. 13986-13993. <https://doi.org/10.1039/C4NR03487B>
- Zinchenko, V. F. (2021). Solid-phase complex compounds and composites of metal oxides, fluorides, and chalcogenides as materials for interference coatings: a review. *Theor. Exp. Chem.*, 57, No. 4, pp. 262-271. <https://doi.org/10.1007/s11237-021-09694-2>
- Lidin, R. A., Andreeva, L. L. & Molochko, V. A. (1987). Handbook of inorganic chemistry. Constants of inorganic substances. Moscow: Khimiya (in Russian).

9. Binnewies, M. & Milke, E. (2002). Thermochemical data of elements and compounds second, revised and extended edition. Weinheim: Wiley-VCH.
10. Efimov, A. I., Belorukova, L. P., Vasilkova, I. V. & Chechev, V. P. (1983). Properties of inorganic compounds. Handbook. Leningrad: Khimiya (in Russian).
11. Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard III, W. A. & Skiff, W. M. (1992). UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.*, 114, No. 25, pp. 10024-10035. <https://doi.org/10.1021/ja00051a040>
12. Rännar, S., Lindgren, F., Geladi, P. & Wold, S. (1994). A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. *J. Chemometrics*, 8, No. 2, pp. 111-125. <https://doi.org/10.1002/cem.1180080204>
13. Carhart, R. E., Smith, D. H. & Venkataraghavan, R. (1985). Atom pairs as molecular features in structure—activity studies: definition and applications. *J. Chem. Inf. Comput. Sci.*, 25, No. 2, pp. 64-73. <https://doi.org/10.1021/ci00046a002>
14. Kuz'min, V. E., Artemenko, A. G., Muratov, E. N., Polischuk, P. G., Ognichenko, L. N., Liahovsky, A. V., Hromov, A. I. & Varlamova, E. V. (2010). Virtual screening and molecular design based on hierarchical QSAR technology. In Puzyn, T., Leszczynski, J. & Cronin, M. (Eds.). *Recent advances in QSAR studies* (pp. 127-176). London: Springer. https://doi.org/10.1007/978-1-4020-9783-6_5
15. Hasegawa, K., Miyashita, Y. & Funatsu, K. (1997). GA strategy for variable selection in QSAR studies: GA-based PLS analysis of calcium channel antagonists. *J. Chem. Inf. Comput. Sci.*, 37, No. 2, pp. 306-310. <https://doi.org/10.1021/ci960047x>

Received 27.09.2023

L.M. Ognichenko, <https://orcid.org/0000-0002-3344-6119>

A.G. Artemenko, <https://orcid.org/0000-0002-9592-1384>

M.E. Kichova, <https://orcid.org/0000-0002-6294-3261>

S.I. Stelmakh, <https://orcid.org/0009-0005-2023-3121>

V.F. Zinchenko, <https://orcid.org/0000-0003-4778-495X>

V.E. Kuz'min, <https://orcid.org/0000-0002-2753-0453>

O.V. Bogatsky Physical-Chemical Institute of the NAS of Ukraine, Odesa, Ukraine

E-mail: ogni@ukr.net, theorchem@gmail.com

QSPR MODEL FOR PREDICTING OF THE STANDARD ENTHALPY OF FORMATION OF COMPLEX OXYGEN-CONTAINING INORGANIC COMPOUNDS

This study utilized a dataset comprising 74 oxygen-containing inorganic compounds, including complex mixed oxides and salts—some of which show promise as components of optical materials. Employing the 1D-simplex approach, a robust consensus QSPR model for the standard enthalpy of formation (ΔH_{298}^0) was developed for the investigated compounds. The calculation of 1D simplex descriptors involved differentiating vertices (atoms) in simplexes based on various characteristics, including those from Mendeleev's periodic system, oxidation state, electronegativity, ionic radius, and features of van der Waals interactions. The predictive QSPR model was constructed using the partial least squares (PLS) method. For the test set of the developed 1D-PLS model, the coefficient of determination was found to be 0.94, with an average relative error of 10.0%. An analysis of the influence of structural parameters on the standard enthalpy of formation for the studied compounds revealed that the orbital quantum numbers of electrons in the outer layer of atoms (28 %) and van der Waals interactions (19 %) had the most significant contributions. The developed model proves useful for the preliminary estimation of the standard enthalpy of formation for various oxygen-containing inorganic compounds.

Keywords: *standard enthalpy of formation, simplex approach, 1D-simplex approach, quantitative structure-property relationship, partial least squares.*