

**МІНІСТЕРСТВО ОХОРОНИ ЗДОРОВ'Я
НАЦІОНАЛЬНИЙ МЕДИЧНИЙ УНІВЕРСИТЕТ
імені О.О. БОГОМОЛЬЦЯ**

**ФАРМАЦЕВТИЧНИЙ ФАКУЛЬТЕТ
КАФЕДРА АНАЛІТИЧНОЇ, ФІЗИЧНОЇ ТА КОЛОЇДНОЇ ХІМІЇ**

ВИПУСКНА КВАЛІФІКАЦІЙНА РОБОТА

**на тему «QSPR ПРОГНОЗУВАННЯ ЛІПОФІЛЬНОСТІ ОРГАНІЧНИХ
ЛІКАРСЬКИХ РЕЧОВИН»**

Виконала: здобувачка вищої освіти 5 курсу
групи 10802

напряму підготовки 22 «Охорона здоров'я»
спеціальності 226 «Фармація, промислова
фармація»

освітньої програми «Фармація»

Пробита Аліна Андріївна

Керівник: кандидат хімічних наук, доцент
Пушкарьова Я.М.

Рецензент: кандидат медичних наук, доцент
Половинка В.О.

КИЇВ-2025

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ	3
ВСТУП	4
ОСНОВНА ЧАСТИНА.....	7
1. ЛІТЕРАТУРНИЙ ОГЛЯД	7
1.1. Роль молекулярних дескрипторів у прогнозуванні ліпофільності	7
1.2. Методи валідації та оцінки якості QSPR моделей	9
Висновки до розділу 1	10
2. ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА	12
2.1. Формування набору даних для моделювання	12
2.2. Реалізація методології QSPR	25
2.2.1. Оцінка інформативних дескрипторів для побудови QSPR моделі..	25
2.2.2. QSPR модель на основі двох дескрипторів.....	30
2.2.2. QSPR модель на основі трьох дескрипторів	35
Висновки до розділу 2	40
ВИСНОВКИ.....	41
СПИСОК ЛІТЕРАТУРИ.....	42
ДОДАТОК.....	45
SUMMARY	46

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

QSPR	Quantitative Structure-Property Relationship (метод кількісного зв'язку між структурою сполуки та її фізико-хімічними властивостями)
$\log P$	ліпофільність (коефіцієнт розподілу між водою і неполярним розчинником)
MW	молекулярна маса
HBD	кількість донорів водневих зв'язків
HBA	кількість акцепторів водневих зв'язків
RotB	кількість обертових зв'язків
PSA	площа полярної поверхні
A	кислотність водневих зв'язків
B	основність водневих зв'язків
S	дипольність/поляризованість
E	надмірна молекулярна рефракція
R^2	коефіцієнт детермінації
R^2_{adj}	скоригований коефіцієнт детермінації
R	коефіцієнт кореляції

ВСТУП

Актуальність теми. Ліпофільність органічних сполук є одним із ключових факторів, що впливають на ефективність лікарських препаратів. Вона визначає їхню здатність проникати через біологічні мембрани, що безпосередньо пов'язано з рівнем абсорбції та біодоступності активних речовин.

Ця характеристика також впливає на процеси метаболізму та виведення сполук, визначаючи їхній розподіл у різних середовищах організму. Наприклад, висока ліпофільність може сприяти накопиченню речовин у жирових тканинах, що впливає на тривалість їхньої дії.

Крім того, знання ліпофільних властивостей необхідне при розробці нових лікарських засобів, оскільки воно допомагає підібрати оптимальну лікарську форму та спосіб доставки. Це особливо актуально для препаратів, які повинні проникати через гематоенцефалічний бар'єр або діяти у конкретних тканинах.

Також слід враховувати екологічний аспект: сполуки з високою ліпофільністю можуть накопичуватися в навколишньому середовищі, що підвищує ризики для екосистем. Тому оцінка цього параметра є важливою не лише у фармакології, а й у контексті екологічної безпеки.

Загалом, ліпофільність є критичним показником, що впливає на весь життєвий цикл лікарської речовини – від розробки та випробувань до її метаболізму в організмі та подальшого впливу на довкілля.

Враховуючи значення ліпофільності для науки та медицини, було розроблено чимало методів для її оцінки. Проте експериментальні процедури, спрямовані на визначення цього параметра, часто є ресурсозатратними та складними. У зв'язку з цим математичне моделювання на основі QSPR (Quantitative Structure Property Relationship) стало потужним альтернативним підходом, що дозволяє прогнозувати ліпофільність сполук, використовуючи лише структурні дескриптори. Це значно спрощує процес

оцінки властивостей нових речовин та прискорює розробку лікарських засобів.

Мета і завдання дослідження. Мета дослідження – розробка QSPR моделі для прогнозування ліпофільності органічних лікарських сполук на основі теоретично розрахованих молекулярних дескрипторів.

Основні завдання:

- формування бази даних, що містить експериментальні значення ліпофільності органічних сполук та відповідні молекулярні дескриптори;
- аналіз впливу структурних характеристик молекул на ліпофільність з метою відбору найбільш інформативних дескрипторів;
- побудова математичної моделі, що дозволяє прогнозувати ліпофільність сполук на основі їхньої молекулярної структури;
- оцінка точності та прогностичної здатності отриманої QSPR моделі за допомогою статистичних критеріїв.

Методи дослідження. У дослідженні застосовували такі методи:

- множинна лінійна регресія – для побудови QSPR моделі прогнозування ліпофільності органічних лікарських сполук;
- кореляційний аналіз – для оцінки взаємозв'язку між молекулярними дескрипторами та експериментальними значеннями ліпофільності;
- статистичний аналіз – для визначення точності та надійності моделі за допомогою таких показників, як коефіцієнт детермінації, скоригований коефіцієнт детермінації та аналіз залишків.

Програмне забезпечення – Matlab R2024b (trial license), Microsoft Excel 2010.

Новизна та значення одержаних результатів. У дослідженні була розроблена ефективна трипараметрична QSPR модель для прогнозування ліпофільності органічних сполук, яка враховує молекулярні дескриптори, здатні визначати взаємозв'язок між структурними характеристиками молекул та їхніми ліпофільними властивостями. Оцінка впливу різних типів молекулярних дескрипторів на точність моделі дозволила визначити

найвагоміші параметри для прогнозування ліпофільності, серед яких – логарифм коефіцієнта розподілу, топологічна площа полярної поверхні та молекулярна маса.

Розроблена модель здатна значно скоротити час і витрати, пов'язані з експериментальним визначенням ліпофільності органічних сполук, що є важливим для фармацевтичної та хімічної промисловості. Вона дозволяє зменшити потребу в дорогих і тривалих лабораторних випробуваннях, що значно підвищує ефективність досліджень.

Ця модель також може стати цінним інструментом для попереднього відбору кандидатів на основі ліпофільності ще на ранніх етапах досліджень. Використання QSPR моделі дозволяє спрогнозувати молекули, які мають оптимальні властивості для цільових застосувань, зокрема в розробці нових лікарських засобів та лікувальних формах. Це дає змогу заздалегідь оцінити їхній потенціал, мінімізуючи ресурси, витрачені на експерименти, і прискорюючи процес створення ефективних фармацевтичних продуктів.

Апробація результатів дослідження. Результати цієї роботи були представлені на VIII International Scientific and Theoretical Conference «Scientific forum: theory and practice of research», 21.03.2025, San Francisco, USA.

Публікації. Опубліковано тези доповідей: Пробита А.А., Пушкарьова Я.М. QSPR прогнозування ліпофільності органічних лікарських речовин. Матеріали VIII International Scientific and Theoretical Conference «Scientific forum: theory and practice of research», 21.03.2025, San Francisco, USA. С. 207-208.

Структура роботи: 47 сторінок, 2 розділи, 1 додаток, список використаних джерел містить 22 публікації.

ОСНОВНА ЧАСТИНА

1. ЛІТЕРАТУРНИЙ ОГЛЯД

1.1. Роль молекулярних дескрипторів у прогнозуванні ліпофільності

Прогнозування ліпофільності органічних сполук є складним завданням, яке потребує врахування різноманітних молекулярних властивостей. Молекулярні дескриптори – це кількісні параметри, що описують структуру молекул і їхні властивості на різних рівнях організації (атом, група атомів, молекула). У контексті ліпофільності вони є основою для створення математичних моделей, які дозволяють передбачити здатність сполук розчинятися в неполярних розчинниках, таких як октанол [1-3].

Загалом, молекулярні дескриптори для прогнозування ліпофільності можна поділити на кілька категорій [4-8]:

1. Фізико-хімічні дескриптори

Це параметри, що описують фізичні властивості молекул, такі як молекулярна маса, топологічна площа полярної поверхні, логарифм коефіцієнта розподілу ($\log P$), поляризованість, дипольний момент. Зокрема, логарифм коефіцієнта розподілу між октанолом і водою ($\log P$) є одним з найважливіших дескрипторів для прогнозування ліпофільності, оскільки він безпосередньо визначає здатність молекули проходити через біологічні мембрани.

2. Структурні дескриптори

Ці дескриптори характеризують розподіл атомів та груп атомів у молекулі, їхні взаємодії та просторову організацію. Прикладом таких дескрипторів є різні індекси зв'язності, що описують молекулярну топологію. Вони можуть допомогти оцінити, як структурні особливості молекул

(наприклад, розмір та форма молекули) впливають на їхню здатність взаємодіяти з неполярними середовищами.

3. Електронні дескриптори

До цієї категорії належать дескриптори, що описують розподіл електронної густини в молекулі, включаючи електронегативність, енергії найнижчої незаселеної орбіталі (LUMO) і найвищої заселеної орбіталі (HOMO), а також електростатичний потенціал. Електронні характеристики молекул можуть впливати на їхню здатність взаємодіяти з полярними та неполярними розчинниками, що робить ці дескриптори важливими для точності моделей ліпофільності.

4. Геометричні дескриптори

Ці параметри описують форму та розмір молекул. Вони можуть бути корисними для оцінки доступності молекули для взаємодії з неполярними середовищами. Прикладом таких дескрипторів є об'єм молекули, її молекулярна площа або радіус Гіршфельда, що відображає мінімальну відстань між молекулою та її найближчим оточенням.

Завдяки використанню цих молекулярних дескрипторів, QSPR моделі можуть точно передбачити ліпофільність сполук, даючи змогу спрогнозувати їхні властивості на ранніх етапах розробки. Вибір відповідних дескрипторів є ключовим етапом у побудові точної та надійної моделі, оскільки невідповідні або надмірні дескриптори можуть призвести до зниження точності прогнозів [9, 10].

Усі ці категорії дескрипторів вимагають ретельного аналізу, щоб визначити, які з них найбільше впливають на ліпофільність органічних сполук. Найбільш ефективні QSPR моделі використовують поєднання кількох дескрипторів для досягнення високої точності та універсальності прогнозів.

1.2. Методи валідації та оцінки якості QSPR моделей

Одним із ключових етапів у створенні надійних QSPR моделей є їх валідація та оцінка точності. Незважаючи на те, що такі моделі дозволяють передбачати властивості органічних сполук, їх ефективність значною мірою залежить від того, наскільки точно модель може передбачати поведінку невідомих сполук, які не входили в тренувальний набір даних. Важливою частиною процесу є виявлення та уникнення таких проблем, як надмірне підгонювання (*overfitting*), коли модель дає хороші результати на тренувальних даних, але не здатна забезпечити точні прогнози на нових [11-13].

Основні методи валідації та оцінки якості QSPR моделей включають [14-18]:

1. Крос-валідація

Один з найбільш поширених методів валідації, який полягає в поділі всього набору даних на кілька частин (зазвичай 5 або 10) та поетапному використанні кожної частини для тестування, а решти – для навчання. Це дозволяє отримати більш надійні оцінки ефективності моделі та зменшити ризик надмірного підгонювання до конкретного набору даних. Одним з найбільш розповсюджених підходів є *k-fold* крос-валідація.

2. Зовнішня валідація

У цьому методі модель тестується на незалежних даних, які не використовувалися під час побудови моделі. Зовнішня валідація дозволяє оцінити, наскільки добре модель працює на нових даних, що є важливим для забезпечення її застосовності в реальних умовах.

3. Оцінка залишкових помилок

Для оцінки точності моделі зазвичай використовують кілька статистичних показників, таких як середньоквадратичну помилку, середню абсолютну помилку, середнє відхилення від реальних значень. Ці показники

дозволяють оцінити, наскільки близькими є передбачені значення до реальних.

4. Коефіцієнт детермінації (R^2)

Це показник, який визначає, наскільки добре модель пояснює варіацію у даних. Значення R^2 ближче до 1 вказує на високу точність прогнозу. Проте для оцінки якості моделі важливо враховувати також коригований коефіцієнт детермінації (adjusted), який враховує кількість вхідних змінних.

5. Тест на стабільність моделі

Стабільність моделі можна оцінити шляхом її застосування до нових підмножин даних або в умовах зміни вихідних параметрів. Це допомагає визначити, наскільки модель є стійкою та може працювати в умовах варіацій.

6. Перевірка на перенавчання (Overfitting)

Для уникнення проблеми надмірного підгоняння часто використовують методи регуляризації, які обмежують складність моделі. Для оцінки, чи не є модель надмірно підлаштованою під тренувальний набір даних, використовують такі підходи, як тестування на незалежному наборі даних або порівняння результатів на навчальній та тестовій вибірках.

Забезпечення правильної валідації моделі дозволяє не лише оцінити її точність, але й підвищити довіру до прогнозів, що є особливо важливим для фармацевтичних та хімічних досліджень, де точність є критично важливою. На додаток до статистичних показників, методи валідації також дозволяють перевірити, чи відповідає модель фізичним та хімічним закономірностям, що дає змогу уникнути неприродних або неправдоподібних результатів.

Висновки до розділу 1

Молекулярні дескриптори є ключовими для побудови точних QSPR моделей, що дозволяють прогнозувати ліпофільність органічних сполук. Різні методи побудови моделей, зокрема множинна регресія та машинне навчання, використовуються для точного прогнозування на основі молекулярних характеристик. Валідація моделей є критично важливою для

перевірки їхньої точності та надійності через крос-валідацію та інші статистичні методи.

QSPR моделі значно знижують час і витрати на розробку нових молекул у фармацевтиці та хімії. Для подальшого вдосконалення моделей потрібно активно використовувати нові методи та технології, зокрема машинне навчання та великі дані.

2. ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА

2.1. Формування набору даних для моделювання

Набір даних для моделювання включав 76 органічних лікарських речовин:

– Ацетамінофен – знеболювальний та жарознижувальний препарат. Використовується для лікування помірного болю (головний біль, біль у спині, біль при застуді) та для зниження високої температури.

– Ацетогексамід – оральний гіпоглікемічний препарат, який використовується для лікування діабету 2 типу. Підвищує секрецію інсуліну з підшлункової залози, що допомагає знижувати рівень глюкози в крові.

– Алосетрон – препарат для лікування синдрому роздратованого кишечника у жінок із діареєю. Зменшує рухливість кишечника і допомагає полегшити симптоми діареї.

– Альпренолол – бета-блокатор, який використовується для лікування гіпертонії та певних видів аритмій. Знижує частоту серцебиття та кров'яний тиск, блокуючи бета-адренорецептори.

– Амінофеназон – анальгезуючий та протизапальний засіб, який раніше використовувався для зменшення болю та запалення, але зараз рідко застосовується через проблеми з безпекою.

– Амлодипін – блокатор кальцієвих каналів, який використовується для лікування гіпертонії та стенокардії. Розслабляє кровоносні судини, знижуючи артеріальний тиск і полегшуючи кровообіг.

– Амоксапін – антидепресант, що використовується для лікування депресії. Має також седативний ефект і деякі антипсихотичні властивості, подібні до препарату локсапін.

– Анастрозол – інгібітор ароматази, що застосовується для лікування раку молочної залози у жінок після менопаузи. Знижує рівень естрогену, що може уповільнити ріст деяких видів раку молочної залози.

– Азатиоприн – імуносупресивний препарат, який використовується для запобігання відторгненню органів після трансплантації та для лікування аутоімунних захворювань, таких як виразковий коліт та системний червоний вовчак.

– Бепридил – блокатор кальцієвих каналів, який використовується для лікування гіпертонії та стенокардії, а також для корекції деяких видів аритмій. Має антиаритмічні властивості.

– Бетаметазон – потужний глюкокортикостероїд, який застосовується для лікування запальних захворювань, таких як артрит, астма, алергії, а також для зменшення запалення після травм чи операцій.

– Бетаксоллол – бета-блокатор, що використовується для лікування глаукоми (зниження внутрішньоочного тиску) та для зниження артеріального тиску при гіпертонії. Впливає на бета-адренорецептори в серці та очах.

– Біпериден – антихолінергічний препарат, який застосовується для лікування паркінсонізму та акатизії, пов'язаної з прийомом антипсихотичних препаратів. Покращує рухові функції шляхом зниження рівня ацетилхоліну.

– Бромазепам – лікарський засіб з групи бензодіазепінів, який використовується для лікування тривожних розладів, панічних атак та інсомнії. Має заспокійливу та протитривожну дію.

– Бупропіон – антидепресант, який також застосовується для допомоги при відмові від куріння (як засіб для припинення куріння). Покращує настрій і знижує симптоми депресії через вплив на нейромедіатори, такі як норадреналін та допамін.

– Буторфанол – протизапальний і анальгезуючий засіб, що належить до опіоїдних анальгетиків. Використовується для полегшення болю середньої і сильної інтенсивності, зокрема після хірургічних операцій.

– Капецитабін – хіміотерапевтичний препарат, що застосовується для лікування раку молочної залози, раку товстої кишки та інших видів раку. Це про-ліки, яке активується в організмі до токсичних метаболітів, що знищують ракові клітини.

– Карбідоба – препарат, що використовується разом з леводопою для лікування хвороби Паркінсона. Карбідоба блокує фермент, який розщеплює леводопу до того, як вона досягне мозку, тим самим підвищуючи її ефективність.

– Церивастатин – препарат, що належить до групи статинів і використовується для зниження рівня холестерину в крові. Застосовується для профілактики серцево-судинних захворювань та інсультів.

– Хлорамбуцил – хіміотерапевтичний препарат, який використовується для лікування деяких видів раку, таких як лімфома Ходжкіна, а також для лікування автоімунних захворювань, таких як системний червоний вовчак.

– Хлорфенірамін – антигістамінний препарат, що використовується для лікування алергічних захворювань, таких як риніт, кон'юнктивіт, алергія на пилок, укуси комах і алергічні реакції шкіри. Зменшує симптоми, пов'язані з алергією (свербіж, набряк, висип).

– Хлорпромазин – антипсихотичний препарат, що використовується для лікування психозів, таких як шизофренія, та для контролю симптомів тривоги, агресії та бурхливої поведінки. Він також застосовується для лікування нудоти та блювання, пов'язаних з хіміотерапією.

– Цилазаприл – інгібітор ангіотензинперетворюючого ферменту, що використовується для лікування гіпертонії та серцевої недостатності. Допомагає знижувати артеріальний тиск і покращувати серцеву функцію.

– Клемастин – антигістамінний препарат, який використовується для лікування алергічних реакцій, таких як сезонний алергічний риніт, а також для зменшення симптомів укусів комах або алергічних висипів.

– Кломіпрамін – антидепресант, який використовується для лікування obsesивно-компульсивного розладу, депресії, тривоги та фобій. Має серотонінергічну активність, що допомагає покращити настрій і зменшити тривожність.

– Клоназепам – лікарський засіб із групи бензодіазепінів, що застосовується для лікування тривожних розладів, панічних атак та епілепсії. Має заспокійливий та протитривожний ефект.

– Клоразепат – протитривожний та седативний препарат, що належить до класу бензодіазепінів. Застосовується для лікування тривожних розладів, судом, безсоння та для зниження рівня агресивної поведінки.

– Циклобензаприн – м'язовий релаксant, який використовується для зменшення болю та спазмів м'язів, зокрема при болях в спині або шиї, а також при м'язових розтягненнях і травмах.

– Ципрогептадин – антигістамінний препарат, що застосовується для лікування алергічних реакцій (наприклад, при алергічному риніті) і для стимулювання апетиту в пацієнтів, що страждають на недоїдання або втрату апетиту через захворювання.

– Дексаметазон – глюкокортикостероїд, який використовується для лікування запальних захворювань, таких як астма, алергії, артрит, а також при захворюваннях, пов'язаних з імунною системою. Має сильну протизапальну та імуносупресивну дію.

– Діазепам – препарат із групи бензодіазепінів, який застосовується для лікування тривожних розладів, панічних атак, безсоння, а також як засіб для релаксації м'язів та при епілепсії. Має заспокійливу, протисудомну та анксиолітичну дію.

– Еналаприл – інгібітор ангіотензинперетворюючого ферменту, що застосовується для лікування гіпертонії та серцевої недостатності. Допомагає знижувати артеріальний тиск і покращує функцію серця, зменшуючи навантаження на нього.

– Естрадіол – основний естроген, що використовується в гормональній терапії для жінок, що переживають менопаузу. Застосовується для зменшення симптомів менопаузи, таких як приливи, а також для лікування остеопорозу.

– Флувоксамін – антидепресант із групи інгібіторів зворотного захоплення серотоніну, який використовується для лікування депресії, obsесивно-компульсивного розладу, тривожних та панічних розладів.

– Летрозол – інгібітор ароматази, який застосовується для лікування раку молочної залози у жінок після менопаузи. Знижує рівень естрогену, що уповільнює ріст естрогенозалежних ракових клітин.

– Омепразол – інгібітор протонної помпи, що застосовується для лікування захворювань шлунково-кишкового тракту, таких як гастрит, виразка шлунка та рефлюкс-езофагіт. Знижує вироблення шлункової кислоти.

– Прохлорперазин – антипсихотичний препарат, який застосовується для лікування психозів, таких як шизофренія, а також для полегшення нудоти та блювання, зокрема після хірургічних операцій або хіміотерапії.

– Тіоридазин – антипсихотичний засіб, який використовується для лікування шизофренії та інших психічних розладів. Зменшує симптоми психозів, такі як галюцинації та марення.

– Астемізол – антигістамінний препарат, який використовувався для лікування алергічних реакцій, таких як сезонний алергічний риніт. Зараз зазвичай не застосовується через потенційний ризик серцевих аритмій.

– Бевантолол – бета-блокатор, який застосовується для лікування гіпертонії та деяких серцевих захворювань, таких як стенокардія. Знижує артеріальний тиск і частоту серцевих скорочень, що полегшує роботу серця.

– Бікалутамід – антиандрогенний препарат, який використовується для лікування раку передміхурової залози. Блокуючи дію тестостерону, він зменшує ріст ракових клітин, чутливих до цього гормону.

– Босентан – антагоніст ендотеліну, що застосовується для лікування легеневої артеріальної гіпертензії. Допомагає знижувати кров'яний тиск у легневих судинах, покращуючи серцеву функцію та витривалість пацієнтів.

– Буспірон – анксиолітичний препарат, який застосовується для лікування генералізованих тривожних розладів. Відрізняється від

бензодіазепінів тим, що не має седативного ефекту і менше викликає звикання.

– Карбамазепін – протисудомний засіб, що використовується для лікування епілепсії, нейропатичного болю та біполярних розладів. Регулює електричну активність у мозку, зменшуючи кількість нападів.

– Карведилол – бета-адреноблокатор, що використовується для лікування гіпертонії, серцевої недостатності та ішемічної хвороби серця. Знижує артеріальний тиск, частоту серцевих скорочень і покращує серцеву функцію.

– Цилостазол – антиагрегант, що використовується для лікування периферичної артеріальної хвороби. Покращує кровообіг, зменшуючи симптоми, такі як болі в ногах при ходьбі.

– Цінакалцет – препарат, що застосовується для лікування гіперкальціємії у пацієнтів з парацитовидною дисфункцією. Знижує рівень кальцію в крові, зменшуючи ризик розвитку ускладнень, таких як камені в нирках.

– Цисаприд – прокінетичний засіб, що раніше використовувався для лікування диспепсії та гастроезофагеального рефлюксу. Підвищує рухливість шлунково-кишкового тракту, покращуючи травлення. Зараз рідше використовується через побічні ефекти.

– Циталопрам – антидепресант із групи інгібіторів зворотного захоплення серотоніну, що застосовується для лікування депресії, тривожних розладів та obsesивно-компульсивного розладу. Підвищує рівень серотоніну в мозку, покращуючи настрій.

– Клотримазол – протигрибковий препарат, який використовується для лікування грибкових інфекцій шкіри, слизових оболонок, а також для лікування вагінальних кандидозів. Має фунгіцидну дію, блокуючи ріст грибків.

– Клозапін – антипсихотичний препарат, який використовується для лікування шизофренії, особливо у пацієнтів, у яких інші препарати не дали

ефекту. Має високу ефективність, але потребує моніторингу за кількістю лейкоцитів через ризик агранулоцитозу.

– Дасатініб – препарат для лікування хронічного мієлоїдного лейкозу, який діє на специфічні мутовані форми ферменту тирозинкінази, що сприяють розвитку раку.

– Дисульфірам – препарат, який використовується для лікування алкоголізму. Його дія полягає в тому, що він спричиняє неприємні симптоми (наприклад, нудоту) у разі вживання алкоголю, що сприяє відмові від алкоголю.

– Домперидон – прокінетик, що застосовується для лікування симптомів диспепсії та нудоти, зокрема після хіміотерапії. Покращує моторику шлунка та кишечника.

– Донепезил – ліки для лікування хвороби Альцгеймера, що покращують когнітивні функції та уповільнюють прогресування захворювання. Інгібітор ацетилхолінестерази.

– Надолол – бета-блокатор, що застосовується для лікування гіпертонії та стенокардії. Знижує артеріальний тиск і частоту серцевих скорочень, покращуючи загальний стан пацієнтів із серцевими захворюваннями.

– Наратриптан – протимігреневий засіб, що застосовується для лікування мігрені та інших видів головного болю. Звужує судини голови та зменшує біль.

– Фенілпропаноламін – симпатоміметик, що використовувався як деконгестант для полегшення закладеності носа. Тепер рідше застосовується через побічні ефекти на серце.

– Піперазин – протигельмінтний засіб, що застосовується для лікування аскаридозу та ентеробіозу. Вбиває гельмінтів у кишечнику.

– Пірензепін – антихолінергічний препарат, який застосовується для лікування виразки шлунка та гастриту. Знижує виділення шлункової кислоти.

– Правастатин – статин, який застосовується для зниження рівня холестерину в крові та профілактики серцево-судинних захворювань. Блокує фермент, який бере участь у виробництві холестерину в печінці.

– Ранітидин – антацид, що використовується для лікування виразкової хвороби шлунка та рефлюксної хвороби стравоходу. Блокує вироблення шлункової кислоти.

– Сульпірид – антипсихотичний засіб, що застосовується для лікування шизофренії та депресивних розладів. Має антидепресивну та анксиолітичну дію.

– Ацетазоламід – препарат для лікування глаукоми та набряків, зокрема при серцевій недостатності. Інгібує карбонатдегідразу, що допомагає знижувати тиск у оці та запобігати утворенню рідини.

– Сульфадіазин – антибіотик сульфаніламідів, що застосовується для лікування інфекцій, викликаних бактеріями, таких як пневмонія, менінгіт і інфекції сечовивідних шляхів.

– Ацикловір – противірусний препарат, що використовується для лікування герпесу, вітряної віспи, оперізувального лишая та інших вірусних інфекцій, викликаних вірусом герпесу. Блокує реплікацію вірусу.

– Амісульприд – антипсихотичний препарат, що застосовується для лікування шизофренії та депресії. Знижує симптоми психозів і допомагає покращити емоційний стан пацієнтів.

– Цефпрозил – це антибіотик з групи цефалоспоринів 2-го покоління. Він має широкий спектр дії, ефективний проти багатьох грампозитивних і грамнегативних бактерій. Використовується для лікування інфекцій дихальних шляхів, сечовивідних шляхів, шкіри та м'яких тканин, а також для профілактики інфекцій після хірургічних втручань.

– Хлортіазид – діуретик (сечогінний препарат), що використовується для лікування гіпертонії та набряків, знижуючи рівень рідини в організмі та зменшуючи навантаження на серце.

– Хлорталідон – діуретик, який застосовується для зниження артеріального тиску та лікування набряків при хворобах серця, печінки чи нирок.

– Циноксацин – антибіотик з групи хінолонів, який використовується для лікування інфекцій сечовивідних шляхів, таких як цистит, пієлонефрит.

– Клоксацилін – антибіотик групи пеніцилінів, застосовується для лікування інфекцій, викликаних стафілококами, таких як шкірні інфекції, пневмонія, остеомієліт.

– Фуросемід – потужний діуретик, що використовується для лікування набряків при серцевій недостатності, хронічних захворюваннях нирок і печінки, а також при гіпертонії.

– Рокситроміцин – антибіотик макролідної групи, застосовується для лікування респіраторних інфекцій, інфекцій шкіри та м'яких тканин, а також для лікування венеричних хвороб.

– Сульфаметизол – антибіотик сульфаніламідів, використовується для лікування інфекцій сечовивідних шляхів та деяких бактеріальних інфекцій.

– Сульфізоксазол – антибіотик сульфаніламідної групи, який застосовується для лікування інфекцій, зокрема сечовивідних шляхів і деяких бактеріальних інфекцій шкіри.

Для характеристики представлених органічних лікарських речовин використали набір із 9 електронних, фізико-хімічних та топологічних дескрипторів:

- молекулярна маса (MW),
- кількість донорів водневих зв'язків (HBD),
- кількість акцепторів водневих зв'язків (HBA),
- кількість обертових зв'язків (RotB) – кількість зв'язків, навколо яких молекула може обертатися,
- площа полярної поверхні (PSA),
- кислотність водневих зв'язків (A),
- основність водневих зв'язків (B),

- дипольність/поляризованість (S),
- надмірна молекулярна рефракція (E).

Значення фізико-хімічних та топологічних дескрипторів, а також експериментальні значення ліпофільності були отримані з бази даних DrugBank [19]. Значення електронних дескрипторів взяті з наукової літератури [20].

Досліджені у роботі лікарські речовини випадковим чином поділили на навчальну (85%) та тестову вибірки (15%). Навчальні дані застосовуються для побудови та оптимізації моделі, тоді як тестові дані використовуються для перевірки її ефективності на нових, не використовуваних під час навчання даних.

У таблиці 2.1 наведені значення 9 дескрипторів, експериментальні значення ліпофільності та розподіл речовин по навчальній та тестовій вибіркам.

Таблиця 2.1. Значення 9 дескрипторів, експериментальні значення ліпофільності та розподіл речовин по навчальній та тестовій вибіркам

№	Лікарська речовина	Вибірка	log P	Дескриптор								
				A	B	S	E	RotB	HBA	HBD	MW	PSA
1	Парацетамол	навчальна	0,46	0,95	0,80	1,63	3,24	2	2	2	151,17	55,41
2	Ацетоекзамід	навчальна	2,44	0,59	1,46	2,79	4,32	6	6	2	324,40	103,22
3	Алоцетрон	навчальна	2,00	0,35	1,38	2,64	3,24	2	5	1	294,36	44,23
4	Альпренолол	навчальна	3,10	0,29	1,36	1,12	0,50	8	2	2	249,36	46,24
5	Амінофеназон	навчальна	1,00	0,00	1,79	1,88	0,81	2	4	0	231,30	26,79
6	Амлодипін	навчальна	3,00	0,36	2,19	2,26	2,65	10	7	2	408,89	105,47
7	Амоксапін	навчальна	3,40	0,16	1,43	1,68	0,99	1	3	1	313,79	34,47
8	Анастрозол	навчальна	1,58	0,00	1,00	2,38	2,06	4	5	0	293,37	61,36
9	Азатіоприн	навчальна	0,10	0,35	1,56	2,86	1,02	3	6	1	277,27	99,66
10	Бепридил	навчальна	5,20	0,00	1,32	1,81	2,52	10	3	0	366,55	10,84
11	Бетаметазон	навчальна	1,14	0,80	1,97	2,95	1,98	2	5	3	392,47	104,22
12	Бетаксоллол	навчальна	2,81	0,29	1,53	1,31	1,10	11	3	2	307,44	55,03
13	Біпериден	навчальна	4,25	0,31	1,17	1,32	1,19	5	2	1	311,47	23,73
14	Бромазепам	навчальна	2,05	0,47	1,27	1,93	1,35	1	4	1	316,16	52,40
15	Бупропіон	навчальна	3,60	0,13	0,94	1,32	3,44	4	2	1	239,75	32,84
16	Буторфанол	навчальна	3,30	0,73	1,32	1,42	2,30	2	2	2	327,47	46,61
17	Капецитабін	навчальна	0,40	0,60	2,40	2,41	1,87	8	9	3	359,36	124,33
18	Карбідоба	навчальна	-1,90	1,69	1,77	1,79	1,03	4	4	5	226,23	132,39
19	Церивастатин	навчальна	3,40	1,20	1,80	2,25	1,45	11	6	3	459,56	104,98
20	Хлорамбуцил	навчальна	1,70	0,57	0,80	1,60	1,62	9	3	1	304,22	41,70
21	Хлорфенірамін	навчальна	3,38	0,00	1,02	1,49	2,33	5	2	0	274,80	11,42
22	Хлорпромазин	навчальна	5,41	0,00	0,99	1,45	3,63	4	2	0	318,87	1,75
23	Цилазаприл	навчальна	0,80	0,71	2,51	2,70	0,73	9	9	3	417,51	106,36

24	Клемастин	навчальна	5,20	0,00	0,97	1,55	1,98	6	2	0	343,90	9,97
25	Кломіпрамін	навчальна	5,19	0,00	0,89	1,66	2,09	4	2	0	314,86	1,75
26	Клоназепам	навчальна	2,41	0,47	1,09	2,25	1,22	2	3	1	315,72	86,30
27	Клоразепат	навчальна	3,00	1,04	1,34	2,14	1,19	2	6	4	314,73	82,98
28	Циклобензаприн	навчальна	5,20	0,00	0,83	1,41	1,74	3	1	0	275,40	1,18
29	Ципрогептадин	навчальна	4,69	0,00	0,83	1,85	2,67	0	1	0	287,41	1,18
30	Дексаметазон	навчальна	1,83	0,80	1,97	2,95	1,61	2	5	3	392,47	104,22
31	Діазепам	навчальна	2,82	0,00	1,04	1,72	1,04	1	3	0	284,75	28,76
32	Еналаприл	навчальна	0,07	0,71	1,92	2,61	1,72	11	7	2	376,46	101,91
33	Естрадіол	навчальна	4,01	0,81	0,95	2,30	2,66	0	1	2	272,39	45,43
34	Флувоксамін	навчальна	3,20	0,23	1,14	0,95	1,41	10	2	1	318,34	58,37
35	Летрозол	навчальна	2,50	0,00	0,97	2,92	2,21	3	5	0	285,31	61,36
36	Омепразол	навчальна	2,23	0,35	2,05	3,18	3,07	5	4	1	345,42	71,04
37	Проклоперазин	навчальна	4,88	0,00	1,47	2,11	2,23	4	3	0	373,95	2,92
38	Тіорідазин	навчальна	5,90	0,00	1,13	1,93	2,05	4	2	0	370,58	1,75
39	Астемізол	навчальна	5,80	0,13	1,64	2,70	2,48	8	4	1	458,58	35,37
40	Бевантолол	навчальна	3,00	0,29	1,82	2,14	2,26	10	2	2	381,91	33,17
41	Бікалутамід	навчальна	2,50	0,71	1,63	3,05	1,79	7	6	2	430,38	110,56
42	Босентан	навчальна	3,70	0,60	2,48	3,50	2,41	11	8	2	551,63	142,96
43	Буспірон	навчальна	2,63	0,00	2,16	2,18	2,26	6	7	0	385,51	60,25
44	Карбамазепін	навчальна	2,77	0,39	0,92	2,06	1,84	1	3	1	236,28	46,81
45	Карведилол	навчальна	3,80	0,62	2,09	3,00	2,08	10	3	3	406,49	78,42
46	Цилостазол	навчальна	2,30	0,41	1,63	2,37	0,91	7	6	1	369,47	81,66
47	Цінакалцет	навчальна	6,50	0,13	0,63	1,37	0,50	7	1	1	357,42	14,57
48	Цизаприд	навчальна	3,30	0,50	2,17	3,15	1,54	10	5	2	465,96	88,69
49	Циталопрам	навчальна	3,76	0,00	1,08	2,25	2,19	5	3	0	324,40	28,67
50	Клотримазол	навчальна	6,10	0,00	0,78	2,37	1,18	4	2	0	344,85	10,81
51	Клозапін	навчальна	3,23	0,20	1,65	1,66	2,17	1	4	1	326,83	25,63

52	Дасатиніб	навчальна	1,80	0,76	2,50	3,47	1,51	8	9	3	488,02	101,85
53	Дисульфірам	навчальна	2,88	0,00	1,16	1,62	2,04	9	2	0	296,54	2,35
54	Домперидон	навчальна	3,90	0,72	1,83	3,13	2,30	5	7	2	425,92	66,80
55	Донепезил	навчальна	4,70	0,00	1,50	2,17	0,91	6	2	0	379,50	37,66
56	Надолол	навчальна	0,81	0,83	1,90	1,56	2,97	6	4	4	309,41	91,35
57	Наратріптан	навчальна	1,60	0,68	1,62	2,14	3,77	5	5	2	335,47	67,24
58	Фенілпропаноламін	навчальна	0,67	0,46	1,22	1,09	3,49	2	2	2	151,21	50,53
59	Піперазин	навчальна	-1,50	0,29	0,89	0,63	2,66	0	2	2	86,14	29,15
60	Пірензепін	навчальна	0,60	0,42	2,42	3,13	2,38	3	7	1	351,41	63,98
61	Фуросемід	навчальна	2,03	1,25	1,50	2,37	2,36	5	6	3	330,75	130,02
62	Ранітидин	навчальна	0,20	0,27	1,97	1,74	0,81	10	3	2	314,41	84,19
63	Рокситроміцин	навчальна	1,70	1,05	5,12	2,90	2,64	13	15	5	837,07	224,19
64	Ацетазоламід	навчальна	-0,26	0,85	1,50	2,55	1,45	3	7	2	222,25	121,43
65	Сульфісоксазол	навчальна	1,01	0,59	1,31	2,44	3,06	3	4	2	267,31	102,81
66	Ацикловір	тестова	-1,76	0,65	2,18	1,95	2,16	4	8	3	225,21	112,32
67	Амісульприд	тестова	1,06	0,50	2,18	3,16	2,11	8	6	2	369,49	107,55
68	Цефпрозил	тестова	0,60	1,55	2,89	3,22	1,48	6	7	4	389,43	143,96
69	Хлортіазид	тестова	-0,24	0,64	1,66	2,74	1,85	1	7	2	295,72	126,91
70	Хлорталідон	тестова	0,85	1,01	1,98	3,05	2,26	2	6	3	338,77	120,90
71	Циноксацин	тестова	1,50	0,57	1,55	2,05	2,67	2	5	1	262,22	91,03
72	Клоксацилін	тестова	2,48	0,84	2,32	3,27	0,74	5	6	2	435,89	116,46
73	Сульфаметизол	тестова	0,54	0,59	1,26	2,71	1,01	3	6	2	270,33	102,86
74	Правастатин	тестова	0,59	1,51	1,89	2,11	2,64	11	7	4	424,54	135,57
75	Сульпірид	тестова	0,57	0,72	2,15	1,78	1,21	7	6	2	341,43	108,62
76	Сульфадіазин	тестова	-0,09	0,59	1,40	2,58	2,91	3	6	2	250,28	99,95

2.2. Реалізація методології QSPR

Для оцінки ефективності математичного моделювання було використано кілька статистичних показників [11, 21]:

- коефіцієнт детермінації R^2 – відображає, який відсоток змін варіації залежної змінної можна пояснити змінами незалежних змінних у моделі;
- скоригований коефіцієнт детермінації R^2_{adj} – це вдосконалена версія коефіцієнта детермінації, що враховує кількість незалежних змінних:

$$R^2_{adj} = 1 - \frac{(1-R^2)(n-1)}{n-k-1}, \quad (2.1)$$

де n – загальна кількість органічних речовин, k – кількість незалежних змінних у моделі (дескрипторів);

- коефіцієнт кореляції R – дає змогу оцінити зв'язок між експериментальними та прогнозованими значеннями ліпофільності органічних сполук.

Також розраховувалися залишки (residuals) – різниця між фактичними і прогнозованими значеннями ліпофільності.

Усі етапи дослідження реалізовані з використанням Matlab R2024b (trial individual license) [22].

2.2.1. Оцінка інформативних дескрипторів для побудови QSPR моделі

Перший етап побудови QSPR моделі включає оцінку дескрипторів на наявність серед них явища мультиколінеарності. Це здійснюється за допомогою коефіцієнтів кореляції.

Таблиця 2.2 містить значення коефіцієнтів кореляції між кожною парою параметрів. Результати свідчать, що існує сильна мультиколінеарність між дескрипторами «кількість донорів водневих зв'язків» та «кислотність водневих зв'язків» ($R = 0,8940$), тобто ці два параметри дублюють

інформацію, і один з них варто виключити із набору для математичного прогнозування ліпофільності.

Рисунки 2.1 та 2.2 свідчать, що ці параметри мають близькі значення коефіцієнта детермінації із експериментальними значеннями ліпофільності, але все ж таки вищий прогностичний вплив на моделювання значень ліпофільності має параметр «кислотність водневих зв'язків».

Отже, робимо висновок, що дескриптор «кількість донорів водневих зв'язків» має бути вилучений із набору характеристик без втрати важливої інформації.

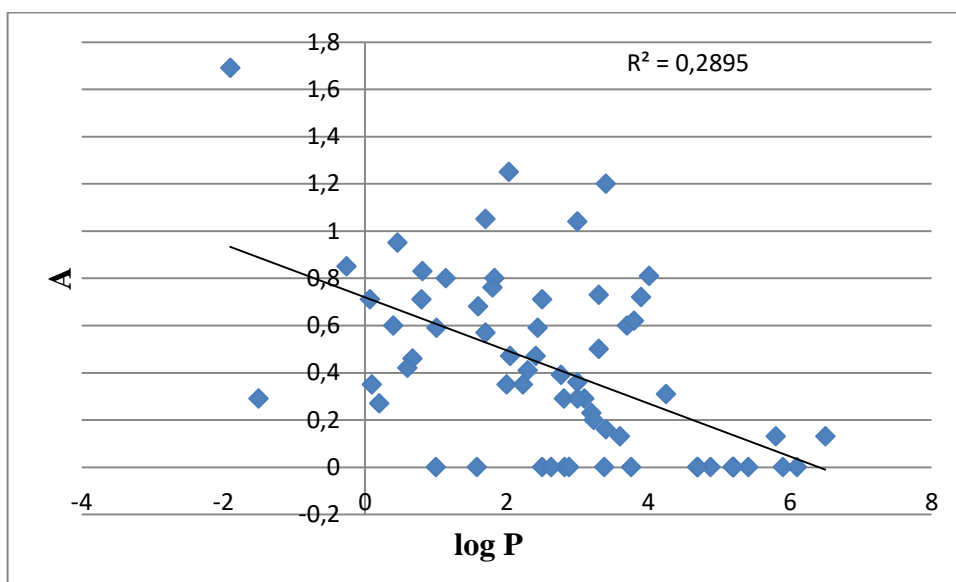


Рисунок 2.1. Залежність значень дескриптора «кислотність водневих зв'язків» від експериментальних значень ліпофільності

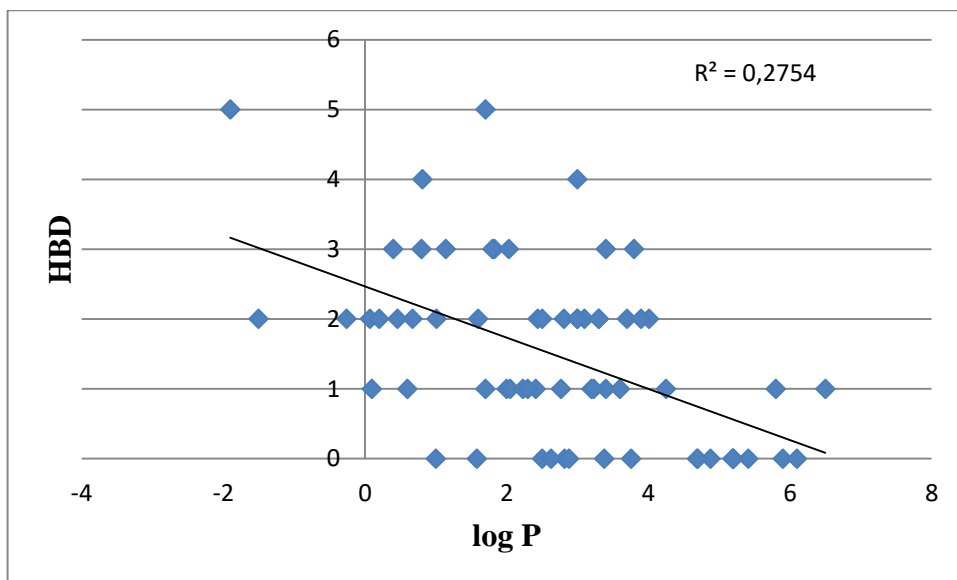


Рисунок 2.2. Залежність значень дескриптора «кількість донорів водневих зв'язків» від експериментальних значень ліпофільності

Таблиця 2.2. Оцінка мультиколінеарності досліджуваних параметрів

Параметр	<i>R</i>								
	A	B	S	E	RotB	HBA	HBD	MW	PSA
A	1,0000	0,4177	0,3161	0,0263	0,1102	0,4800	0,8940	0,1896	0,7599
B	0,4177	1,0000	0,5486	0,0037	0,5068	0,8367	0,5911	0,7702	0,7437
S	0,3161	0,5486	1,0000	0,0654	0,1920	0,6657	0,2580	0,5838	0,5727
E	0,0263	0,0037	0,0654	1,0000	-0,1356	0,0159	0,0047	-0,0510	-0,0123
RotB	0,1102	0,5068	0,1920	-0,1356	1,0000	0,3750	0,2635	0,6132	0,3572
HBA	0,4800	0,8367	0,6657	0,0159	0,3750	1,0000	0,5310	0,6583	0,8197
HBD	0,8940	0,5911	0,2580	0,0047	0,2635	0,5310	1,0000	0,3009	0,7880
MW	0,1896	0,7702	0,5838	-0,0510	0,6132	0,6583	0,3009	1,0000	0,4953
PSA	0,7599	0,7437	0,5727	-0,0123	0,3572	0,8197	0,7880	0,4953	1,0000

Для оцінки значущості дескрипторів для прогнозування ліпофільності розраховали коефіцієнти детермінації між кожним параметром та експериментальними значеннями ліпофільності (таблиця 2.3).

Таблиця 2.3. Оцінка значущості дескрипторів для прогнозування ліпофільності

Параметр	R^2
кислотність водневих зв'язків	0,2895
основність водневих зв'язків	0,0987
дипольність/поляризованість	0,0123
надмірна молекулярна рефракція	0,0136
кількість обертових зв'язків	0,0063
кількість акцепторів водневих зв'язків	0,1663
молекулярна маса	0,0663
площа полярної поверхні	0,3317

На основі аналізу розрахованих коефіцієнти детермінації, можна зробити наступні висновки:

– дескриптори «кількість обертових зв'язків», «дипольність/поляризованість» та «надмірна молекулярна рефракція» не відіграє значної ролі для прогнозування ліпофільності органічних лікарських речовин, тому ними можна знехтувати;

– найбільші відсотки змін варіації залежної змінної (тобто ліпофільності) можна пояснити із застосуванням площі полярної поверхні та кислотності водневих зв'язків.

Отже, розробку оптимальної QSPR моделі почали саме на основі комбінації площі полярної поверхні та кислотності водневих зв'язків. Інші дескриптори поступово додавалися до цієї двопараметричної моделі.

2.2.2. QSPR модель на основі двох дескрипторів

Математична модель із двома змінними, що описує зв'язок між ліпофільністю лікарських органічних речовин та дескрипторами «площа полярної поверхні» та «кислотність водневих зв'язків» описується рівнянням:

$$\log P = 4,182986 - 1,13952 \cdot x_1 - 0,01637 \cdot x_2, \quad (2.2)$$

де x_1 – кислотність водневих зв'язків, x_2 – площа полярної поверхні.

Статистичні показники для представленої моделі мають не достатньо високі значення, що може свідчити про необхідність включення подальших дескрипторів:

- $R = 0,5963$,
- $R^2 = 0,3556$,
- $R_{adj}^2 = 0,3348$.

Результати застосування двопараметричної QSPR моделі до прогнозування ліпофільності представлено у таблиці 2.4.

Застосування двопараметричної QSPR моделі до прогнозування ліпофільності лікарських речовин тестової вибірки свідчить про неефективність даного регресійного рівняння (рисунок 2.3). Ця модель потребує включення ще додаткових параметрів.

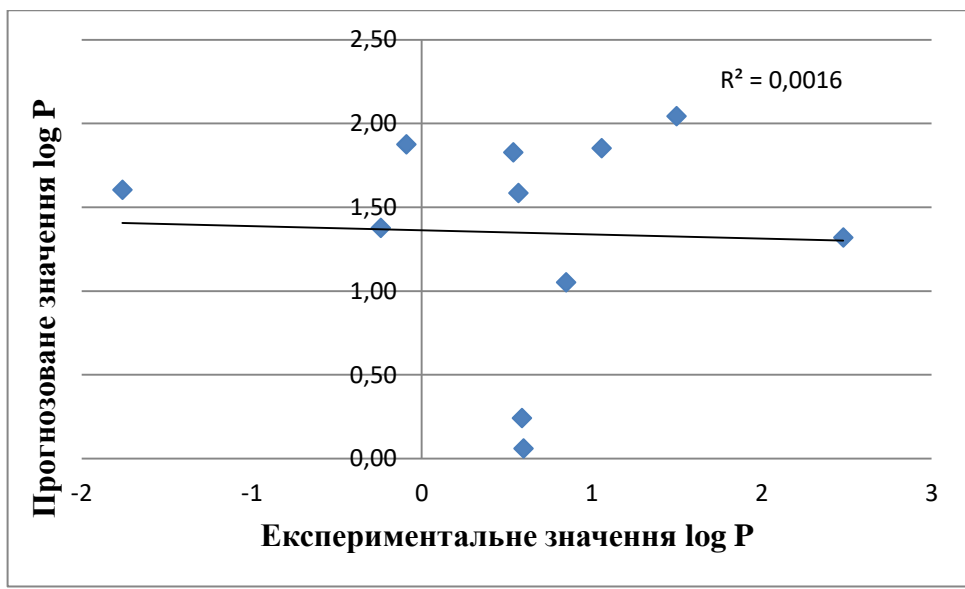


Рисунок 2.3. Залежність прогнозованих значень ліпофільності від експериментальних значень ліпофільності лікарських речовин тестової вибірки у випадку двопараметричної QSPR моделі

Таблиця 2.4. Результати застосування двопараметричної QSPR моделі до прогнозування ліпофільності

№	Лікарська речовина	Вибірка	$\log P$	Прогнозоване значення $\log P$	Залишки
1	Парацетамол	навчальна	0,46	2,19	-1,73
2	Ацетоекзамід	навчальна	2,44	1,82	0,62
3	Алоцетрон	навчальна	2,00	3,06	-1,06
4	Альпренолол	навчальна	3,10	3,10	0,00
5	Амінофеназон	навчальна	1,00	3,74	-2,74
6	Амлодипін	навчальна	3,00	2,05	0,95
7	Амоксапін	навчальна	3,40	3,44	-0,04
8	Анастрозол	навчальна	1,58	3,18	-1,60
9	Азатіоприн	навчальна	0,10	2,15	-2,05
10	Бепридил	навчальна	5,20	4,01	1,19
11	Бетаметазон	навчальна	1,14	1,57	-0,43
12	Бетаксоллол	навчальна	2,81	2,95	-0,14
13	Біпериден	навчальна	4,25	3,44	0,81
14	Бромазепам	навчальна	2,05	2,79	-0,74
15	Бупропіон	навчальна	3,60	3,50	0,10
16	Буторфанол	навчальна	3,30	2,59	0,71
17	Капецитабін	навчальна	0,40	1,46	-1,06
18	Карбідоба	навчальна	-1,90	0,09	-1,99
19	Церивастатин	навчальна	3,40	1,10	2,30
20	Хлорамбуцил	навчальна	1,70	2,85	-1,15
21	Хлорфенірамін	навчальна	3,38	4,00	-0,62
22	Хлорпромазин	навчальна	5,41	4,15	1,26
23	Цилазаприл	навчальна	0,80	1,63	-0,83
24	Клемастин	навчальна	5,20	4,02	1,18

25	Кломіпрамін	навчальна	5,19	4,15	1,04
26	Клоназепам	навчальна	2,41	2,24	0,17
27	Клоразепат	навчальна	3,00	1,64	1,36
28	Циклобензаприн	навчальна	5,20	4,16	1,04
29	Ципрофептадин	навчальна	4,69	4,16	0,53
30	Дексаметазон	навчальна	1,83	1,57	0,26
31	Діазепам	навчальна	2,82	3,71	-0,89
32	Еналаприл	навчальна	0,07	1,71	-1,64
33	Естрадіол	навчальна	4,01	2,52	1,49
34	Флувоксамін	навчальна	3,20	2,97	0,23
35	Летрозол	навчальна	2,50	3,18	-0,68
36	Омепразол	навчальна	2,23	2,62	-0,39
37	Проклоперазин	навчальна	4,88	4,14	0,74
38	Тіорідазин	навчальна	5,90	4,15	1,75
39	Астемізол	навчальна	5,80	3,46	2,34
40	Бевантолол	навчальна	3,00	3,31	-0,31
41	Бікалутамід	навчальна	2,50	1,56	0,94
42	Босентан	навчальна	3,70	1,16	2,54
43	Буспірон	навчальна	2,63	3,20	-0,57
44	Карбамазепін	навчальна	2,77	2,97	-0,20
45	Карведилол	навчальна	3,80	2,19	1,61
46	Цилостазол	навчальна	2,30	2,38	-0,08
47	Цінакалцет	навчальна	6,50	3,80	2,70
48	Цизаприд	навчальна	3,30	2,16	1,14
49	Циталопрам	навчальна	3,76	3,71	0,05
50	Клотримазол	навчальна	6,10	4,01	2,09
51	Клозапін	навчальна	3,23	3,54	-0,31
52	Дасатиніб	навчальна	1,80	1,65	0,15

53	Дисульфірам	навчальна	2,88	4,14	-1,26
54	Домперидон	навчальна	3,90	2,27	1,63
55	Донепезил	навчальна	4,70	3,57	1,13
56	Надолол	навчальна	0,81	1,74	-0,93
57	Наратріптан	навчальна	1,60	2,31	-0,71
58	Фенілпропаноламін	навчальна	0,67	2,83	-2,16
59	Піперазин	навчальна	-1,50	3,38	-4,88
60	Пірензепін	навчальна	0,60	2,66	-2,06
61	Фуросемід	навчальна	2,03	0,63	1,40
62	Ранітидин	навчальна	0,20	2,50	-2,30
63	Рокситроміцин	навчальна	1,70	-0,68	2,38
64	Ацетазоламід	навчальна	-0,26	1,23	-1,49
65	Сульфісоксазол	навчальна	1,01	1,83	-0,82
66	Ацикловір	тестова	-1,76	1,60	-3,36
67	Амісульприд	тестова	1,06	1,85	-0,79
68	Цефпрозил	тестова	0,60	0,06	0,54
69	Хлортіазид	тестова	-0,24	1,38	-1,62
70	Хлорталідон	тестова	0,85	1,05	-0,20
71	Циноксацин	тестова	1,50	2,04	-0,54
72	Клоксацилін	тестова	2,48	1,32	1,16
73	Сульфаметизол	тестова	0,54	1,83	-1,29
74	Правастатин	тестова	0,59	0,24	0,35
75	Сульпірид	тестова	0,57	1,58	-1,01
76	Сульфадіазин	тестова	-0,09	1,87	-1,96

2.2.2. QSPR модель на основі трьох дескрипторів

Математична модель із трьома змінними, що описує зв'язок між ліпофільністю лікарських органічних речовин та дескрипторами «площа полярної поверхні», «кислотність водневих зв'язків» та «кількість акцепторів водневих зв'язків» описується рівнянням:

$$\log P = 4,082266 - 0,96365 \cdot x_1 - 0,0212 \cdot x_2 + 0,077949 \cdot x_3, \quad (2.3)$$

де x_1 – кислотність водневих зв'язків, x_2 – площа полярної поверхні, x_3 – кількість акцепторів водневих зв'язків.

У порівнянні із двопараметричною моделлю (2.2), статистичні показники для представленої моделі (2.3) суттєво не змінилися, що свідчить про неефективність дескриптора «кількість акцепторів водневих зв'язків» для прогнозування ліпофільності лікарських органічних речовин у поєднанні із дескрипторами «площа полярної поверхні» та «кислотність водневих зв'язків»:

- $R = 0,5991$,
- $R^2 = 0,3589$,
- $R_{adj}^2 = 0,3274$.

Математична модель із трьома змінними, що описує зв'язок між ліпофільністю лікарських органічних речовин та дескрипторами «площа полярної поверхні», «кислотність водневих зв'язків» та «основність водневих зв'язків» описується рівнянням:

$$\log P = 3,724854 - 0,8047 \cdot x_1 - 0,02473 \cdot x_2 + 0,545547 \cdot x_3, \quad (2.4)$$

де x_1 – кислотність водневих зв'язків, x_2 – площа полярної поверхні, x_3 – основність водневих зв'язків.

Для даної моделі у порівнянні із двопараметричною моделлю (2.2), спостерігається досить незначне підвищення статистичних показників для представленої моделі (2.4):

- $R = 0,6094$ (показник збільшився на 0,0131),
- $R^2 = 0,3714$ (показник збільшився на 0,0158),

– $R_{adj}^2 = 0,3405$ (показник збільшився на 0,0057).

Отже, додавання дескриптора «основність водневих зв'язків» вносить дуже маленький вклад у прогнозування ліпофільності органічних лікарських речовин і цим вкладом можна знехтувати.

Математична модель із трьома змінними, що описує зв'язок між ліпофільністю лікарських органічних речовин та дескрипторами «площа полярної поверхні», «кислотність водневих зв'язків» та «молекулярна маса» описується рівнянням:

$$\log P = 0,641684 + 0,413407 \cdot x_1 - 0,04158 \cdot x_2 + 0,013229 \cdot x_3, \quad (2.5)$$

де x_1 – кислотність водневих зв'язків, x_2 – площа полярної поверхні, x_3 – молекулярна маса.

Для даної моделі у порівнянні із двопараметричною моделлю (2.2), спостерігається суттєве підвищення статистичних показників для представленої моделі (2.5):

– $R = 0,8464$ (показник збільшився на 0,2501),

– $R^2 = 0,7165$ (показник збільшився на 0,3609),

– $R_{adj}^2 = 0,7025$ (показник збільшився на 0,3677).

Отже, додавання дескриптора «молекулярна маса» вносить значний вклад у прогнозування ліпофільності органічних лікарських речовин.

Результати застосування цієї трипараметричної QSPR моделі до прогнозування ліпофільності представлено у таблиці 2.5.

Таблиця 2.5. Результати застосування трипараметричної QSPR моделі до прогнозування ліпофільності

№	Лікарська речовина	Вибірка	Експериментальне значення $\log P$	Прогнозоване значення $\log P$	Залишки
1	Парацетамол	навчальна	0,46	0,73	-0,27
2	Ацетоекзамід	навчальна	2,44	0,88	1,56
3	Алоцетрон	навчальна	2,00	2,84	-0,84
4	Альпренолол	навчальна	3,10	2,14	0,96
5	Амінофеназон	навчальна	1,00	2,59	-1,59
6	Амлодипін	навчальна	3,00	1,81	1,19
7	Амоксапін	навчальна	3,40	3,43	-0,03
8	Анастрозол	навчальна	1,58	1,97	-0,39
9	Азатиоприн	навчальна	0,10	0,31	-0,21
10	Бепридил	навчальна	5,20	5,04	0,16
11	Бетаметазон	навчальна	1,14	1,83	-0,69
12	Бетаксоллол	навчальна	2,81	2,54	0,27
13	Біпериден	навчальна	4,25	3,90	0,35
14	Бромазепам	навчальна	2,05	2,84	-0,79
15	Бупропіон	навчальна	3,60	2,50	1,10
16	Буторфанол	навчальна	3,30	3,34	-0,04
17	Капецитабін	навчальна	0,40	0,47	-0,07
18	Карбідоба	навчальна	-1,90	-1,17	-0,73
19	Церивастатин	навчальна	3,40	2,85	0,55
20	Хлорамбуцил	навчальна	1,70	3,17	-1,47
21	Хлорфенірамін	навчальна	3,38	3,80	-0,42
22	Хлорпромазин	навчальна	5,41	4,79	0,62
23	Цилазаприл	навчальна	0,80	2,04	-1,24
24	Клемастин	навчальна	5,20	4,78	0,42

25	Кломіпрамін	навчальна	5,19	4,73	0,46
26	Клоназепам	навчальна	2,41	1,42	0,99
27	Клоразепат	навчальна	3,00	1,78	1,22
28	Циклобензаприн	навчальна	5,20	4,24	0,96
29	Ципрофептадин	навчальна	4,69	4,39	0,30
30	Дексаметазон	навчальна	1,83	1,83	0,00
31	Діазепам	навчальна	2,82	3,21	-0,39
32	Еналаприл	навчальна	0,07	1,68	-1,61
33	Естрадіол	навчальна	4,01	2,69	1,32
34	Флувоксамін	навчальна	3,20	2,52	0,68
35	Летрозол	навчальна	2,50	1,86	0,64
36	Омепразол	навчальна	2,23	2,40	-0,17
37	Проклоперазин	навчальна	4,88	5,47	-0,59
38	Тіорідазин	навчальна	5,90	5,47	0,43
39	Астемізол	навчальна	5,80	5,29	0,51
40	Бевантолол	навчальна	3,00	4,43	-1,43
41	Бікалутамід	навчальна	2,50	2,03	0,47
42	Босентан	навчальна	3,70	2,24	1,46
43	Буспірон	навчальна	2,63	3,24	-0,61
44	Карбамазепін	навчальна	2,77	1,98	0,79
45	Карведилол	навчальна	3,80	3,01	0,79
46	Цилостазол	навчальна	2,30	2,30	0,00
47	Цінакалцет	навчальна	6,50	4,82	1,68
48	Цизаприд	навчальна	3,30	3,32	-0,02
49	Циталопрам	навчальна	3,76	3,74	0,02
50	Клотримазол	навчальна	6,10	4,75	1,35
51	Клозапін	навчальна	3,23	3,98	-0,75
52	Дасатиніб	навчальна	1,80	3,18	-1,38

53	Дисульфiрам	навчальна	2,88	4,47	-1,59
54	Домперидон	навчальна	3,90	3,80	0,10
55	Донепезил	навчальна	4,70	4,10	0,60
56	Надолол	навчальна	0,81	1,28	-0,47
57	Наратрiптан	навчальна	1,60	2,56	-0,96
58	Фенiлпропаноламiн	навчальна	0,67	0,73	-0,06
59	Пiперазин	навчальна	-1,50	0,69	-2,19
60	Пiрензепiн	навчальна	0,60	2,80	-2,20
61	Фуросемiд	навчальна	2,03	0,13	1,90
62	Ранiтидин	навчальна	0,20	1,41	-1,21
63	Рокситромiцин	навчальна	1,70	2,83	-1,13
64	Ацетазоламiд	навчальна	-0,26	-1,12	0,86
65	Сульфiсоксазол	навчальна	1,01	0,15	0,86
66	Ацикловiр	тестова	-1,76	-0,78	-0,98
67	Амисульприд	тестова	1,06	1,26	-0,20
68	Цефпрозил	тестова	0,60	0,45	0,15
69	Хлортiазид	тестова	-0,24	-0,46	0,22
70	Хлорталiдон	тестова	0,85	0,51	0,34
71	Циноксацин	тестова	1,50	0,56	0,94
72	Клоксацилiн	тестова	2,48	1,91	0,57
73	Сульфаметизол	тестова	0,54	0,18	0,36
74	Правастатин	тестова	0,59	1,25	-0,66
75	Сульпiрид	тестова	0,57	0,94	-0,37
76	Сульфадiазин	тестова	-0,09	0,04	-0,13

Високі значення коефіцієнта детермінації ($R^2 = 0,7428$) та коефіцієнта кореляції ($R = 0,8619$) між експериментальними та прогнозованими значеннями ліпофільності свідчать про застосовуваність математичної моделі (2.5) до нових даних.

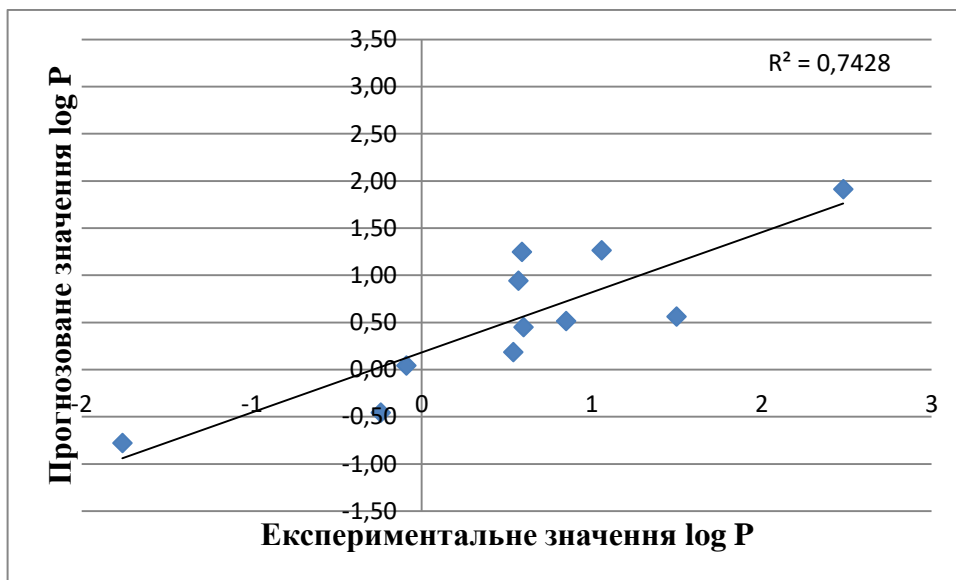


Рисунок 2.4. Залежність прогнозованих значень ліпофільності від експериментальних значень ліпофільності лікарських речовин тестової вибірки у випадку трипараметричної QSPR моделі

Додавання до розробленої трипараметричної моделі дескрипторів «дипольність/поляризованість», «кількість обертових зв'язків» та «надмірна молекулярна рефракція» підтвердило їх неефективність для прогнозування ліпофільності органічних лікарських речовин.

Висновки до розділу 2

Із дев'яти досліджених фізико-хімічних, топологічних та електронних дескрипторів інформативними виявилися лише три. Ефективна QSPR модель для прогнозування ліпофільності органічних лікарських речовин має вигляд:

$$\log P = 0,641684 + 0,413407 \cdot x_1 - 0,04158 \cdot x_2 + 0,013229 \cdot x_3,$$

де x_1 – кислотність водневих зв'язків, x_2 – площа полярної поверхні, x_3 – молекулярна маса.

ВИСНОВКИ

1. Аналіз обраних дескрипторів показав, що кислотність водневих зв'язків, площа полярної поверхні та молекулярна маса є ключовими параметрами для прогнозування ліпофільності органічних лікарських речовин.

2. Різниця між прогнозованими та експериментальними значеннями ліпофільності не перевищує одиницю, що свідчить про високу точність застосованої моделі.

3. Розроблена модель може бути успішно використана для первинного скринінгу сполук на ранніх етапах розробки лікарських засобів, що дозволяє ідентифікувати перспективні молекули ще на етапі їхнього проектування.

СПИСОК ЛІТЕРАТУРИ

1. Muhammad, U., Uzairu, A., & Ebuka Arthur, D. (2018). Review on: quantitative structure activity relationship (QSAR) modeling. *J Anal Pharm Res*, 7(2), 240-242.
2. Carracedo-Reboredo, P., Liñares-Blanco, J., Rodríguez-Fernández, N., Cedrón, F., Novoa, F. J., Carballal, A., ... & Fernandez-Lozano, C. (2021). A review on machine learning approaches and trends in drug discovery. *Computational and structural biotechnology journal*, 19, 4538-4558.
3. Kaneko, H. (2023). Molecular descriptors, structure generation, and inverse QSAR/QSPR based on SELFIES. *ACS omega*, 8(24), 21781-21786.
4. Khan, A. U. (2016). Descriptors and their selection methods in QSAR analysis: paradigm for drug design. *Drug discovery today*, 21(8), 1291-1302.
5. Alam, S. M., Nigar, N., Rasheed, M. W., & Amin, L. (2024). Analyzing the role of reducible molecular descriptors and thermodynamic aspects of anti-tuberculosis drugs via QSPR study. *Frontiers in Physics*, 12, 1383216.
6. Chang, E. D., Hogstrand, C., Miller, T. H., Owen, S. F., & Bury, N. R. (2018). The use of molecular descriptors to model pharmaceutical uptake by a fish primary gill cell culture epithelium. *Environmental science & technology*, 53(3), 1576-1584.
7. van den Maagdenberg, H. W., Šícho, M., Araripe, D. A., Luukkonen, S., Schoenmaker, L., Jespers, M., ... & van Westen, G. J. (2024). QSPRpred: a flexible open-source quantitative structure-property relationship modelling tool. *Journal of Cheminformatics*, 16(1), 128.
8. Ouattara, O., & Ziao, N. (2017). Quantum Chemistry Prediction of Molecular Lipophilicity Using Semi-Empirical AM1 and Ab Initio HF/6-311++ G Levels. *Computational Chemistry*, 5(01), 38.
9. Ginex, T., Vazquez, J., Gilbert, E., Herrero, E., & Luque, F. J. (2019). Lipophilicity in drug design: An overview of lipophilicity descriptors in 3D-QSAR studies. *Future medicinal chemistry*, 11(10), 1177-1193.

10. Bahia, M. S., Kaspi, O., Touitou, M., Binayev, I., Dhail, S., Spiegel, J., ... & Senderowitz, H. (2023). A comparison between 2D and 3D descriptors in QSAR modeling based on bio-active conformations. *Molecular Informatics*, *42*(4), 2200186.
11. Roy, K., Kar, S., Das, R. N., Roy, K., Kar, S., & Das, R. N. (2015). Statistical methods in QSAR/QSPR. *A Primer on QSAR/QSPR Modeling: Fundamental Concepts*, 37-59.
12. Driouch, M., Lazrak, J., Bensouda, Z., Elhaloui, A., Sfaira, M., Saffaj, T., & Taleb, M. (2020). Development and validation of QSPR models for corrosion inhibition of carbon steel by some pyridazine derivatives in acidic medium. *Heliyon*, *6*(10).
13. Héberger, K. (2023). Selection of optimal validation methods for quantitative structure–activity relationships and applicability domain. *SAR and QSAR in Environmental Research*, *34*(5), 415-434.
14. Roy, K., Mitra, I., Kar, S., Ojha, P. K., Das, R. N., & Kabir, H. (2012). Comparative studies on some metrics for external validation of QSPR models. *Journal of chemical information and modeling*, *52*(2), 396-408.
15. Fayet, G., & Rotureau, P. (2016). How to use QSPR-type approaches to predict properties in the context of green chemistry. *Biofuels, Bioproducts and Biorefining*, *10*(6), 738-752.
16. Fayet, G., & Rotureau, P. (2016, June). QSPR model for regulatory purpose: from development to integration into the QSAR toolbox. In *15. International Symposium on Loss Prevention and Safety Promotion in the Process Industry* (Vol. 48, pp. 79-84). AIDIC.
17. Liu, P., & Long, W. (2009). Current mathematical methods used in QSAR/QSPR studies. *International Journal of Molecular Sciences*, *10*(5), 1978-1998.
18. Chang, C. M., Banerjee, A., Kumar, V., Roy, K., & Benfenati, E. (2025). The q-RASPR approach for predicting the property and fate of persistent organic pollutants. *Scientific Reports*, *15*(1), 1344.

19. <https://go.drugbank.com/>
20. Ashrafi, A., Teimouri, K., Aghazadeh, F., & Shayanfar, A. (2024). Neural Network Models for Predicting Solubility and Metabolism Class of Drugs in the Biopharmaceutics Drug Disposition Classification System (BDDCS). *European Journal of Drug Metabolism and Pharmacokinetics*, 49(1), 1-6.
21. Nantasenamat, C., Isarankura-Na-Ayudhya, C., Naenna, T., & Prachayasittikul, V. (2009). A practical overview of quantitative structure-activity relationship. *EXCLI J*, 8(7), 74-88.
22. <https://www.mathworks.com/products.html>

ДОДАТОК

IST № 25/2103-001

0,5 ECTS credits

CERTIFICATE OF PARTICIPATION



Alina Probyta

participated in the VIII International Scientific and Theoretical Conference

Scientific forum: theory and practice of research

21.03.2025 | San Francisco, USA

and published scientific paper in the Collection of scientific papers «SCIENTIA»
with DOI 10.36074/scientia-21.03.2025 and ISBN 979-8-89660-286-6 (series)  Bowker

The conference is included in the Academic Resource Index ResearchBib catalog and registered in the UkrISTEI database (Certificate № 119 dated January 6, 2025);



President of the International
Center of Scientific Research
Registration number in the Unified Register of Public Associations: 1499141

MIRIAM GOLDENBLAT
www.scientia.report



SUMMARY

Alina Probyta

QSPR PREDICTION OF LIPOPHILICITY OF ORGANIC DRUGS

Department of Analytical, Physical and Colloid Chemistry

Scientific supervisor: associate professor Yaroslava Pushkarova

Keywords: screening of compounds, drug development, descriptor, lipophilicity.

Introduction. Lipophilicity of organic compounds is one of the key factors affecting the effectiveness of pharmaceutical drugs. It determines their ability to pass through biological membranes, which is directly related to the absorption rate and bioavailability of active substances.

This property also influences the metabolism and elimination processes of compounds, determining their distribution in various body environments.

Moreover, understanding lipophilic properties is essential when developing new drugs as it helps select the optimal dosage form and delivery method. This is especially important for drugs that need to cross the blood-brain barrier or target specific tissues.

Overall, lipophilicity is a critical parameter that affects the entire lifecycle of a pharmaceutical substance – from development and testing to its metabolism in the body and subsequent environmental impact.

Materials and methods. Data set – 76 organic compounds, which are characterized by 9 physical-chemical, electronic and topological descriptors. Methods of investigation: multiple linear regression; correlation analysis; statistical analysis.

Results. Molecular descriptors are key to building accurate QSPR models that allow the prediction of the lipophilicity of organic compounds. Different modeling methods, including multiple regression and machine learning, are used for precise predictions based on molecular characteristics. Model validation is critical for checking their accuracy and reliability through cross-validation and other statistical methods.

QSPR models significantly reduce the time and costs of developing new molecules in pharmaceuticals and chemistry. To further improve the models, it is necessary to actively use new methods and technologies, including machine learning and big data.

Out of the nine studied physicochemical, topological, and electronic descriptors, only three proved to be informative. The effective QSPR model for predicting the lipophilicity of organic pharmaceutical compounds is as follows:

$$\log P = 0,641684 + 0,413407 \cdot x_1 - 0,04158 \cdot x_2 + 0,013229 \cdot x_3,$$

where x_1 – hydrogen bond acidity, x_2 – polar surface area, x_3 – molecular mass.

Conclusions. Analysis of the selected descriptors showed that hydrogen bond acidity, polar surface area, and molecular mass are key parameters for predicting the lipophilicity of organic pharmaceutical compounds.

The difference between the predicted and experimental lipophilicity values does not exceed one unit, indicating high accuracy and reliability of the applied model.

The developed model can be successfully used for the preliminary screening of compounds at early stages of drug development, allowing the identification of promising molecules during their design phase.