

Міністерство освіти і науки України
Харківський національний університет імені В.Н. Каразіна
Кафедра прикладної хімії
Кафедра хімічного матеріалознавства

“ЗАТВЕРДЖУЮ”

Проректор з науково-педагогічної
роботи

“ _____ ” _____ 20 __ р.

Робоча програма навчальної дисципліни
МОЛЕКУЛЯРНИЙ ДИЗАЙН НОВИХ БІОЛОГІЧНО АКТИВНИХ
РЕЧОВИН

рівень вищої освіти	<u>магістр</u>
галузь знань	<u>10 Природничі науки</u>
спеціальність	<u>102 хімія</u>
освітня програма	<u>освітньо-професійна програма «Фармацевтична хімія»</u>
вид дисципліни	<u>за вибором</u>
факультет	<u>хімічний</u>

2019 / 2020 навчальні роки

Програму рекомендовано до затвердження вченою радою хімічного факультету

від “ ____ ” _____ 2019 року, протокол № ____

РОЗРОБНИКИ ПРОГРАМИ: **Іванов Володимир Венедиктович**, д. х. н., професор,
професор кафедри хімічного матеріалознавства, **Ткаченко Володимир Володимирович**,
к.х.н., старший викладач кафедри прикладної хімії.

Програму схвалено на засіданні кафедри _____ хімічного матеріалознавства _____

Протокол від “ ____ ” _____ 2019 року № ____

Завідувач кафедри _____ хімічного матеріалознавства _____

_____ Коробов О.І.
(підпис) (прізвище та ініціали)

Програму схвалено на засіданні кафедри _____ прикладної хімії _____

Протокол від “ ____ ” _____ 2019 року № ____

Завідувач кафедри _____ прикладної хімії _____

_____ Чебанов В.А.
(підпис) (прізвище та ініціали)

Програму погоджено методичною комісією

_____ хімічного факультету _____
назва факультету, для здобувачів вищої освіти якого викладається навчальна дисципліна

Протокол від “ ____ ” _____ 2019 року, протокол № ____

Голова методичної комісії _____ хімічного факультету _____

_____ Єфімов П.В.
(підпис) (прізвище та ініціали)

ВСТУП

Програма навчальної дисципліни «Молекулярний дизайн нових біологічно активних речовин» складена відповідно до освітньо-професійної програми підготовки магістра спеціальності (напряму) 102 Хімія

1. Опис навчальної дисципліни

1.1. Мета викладання навчальної дисципліни.

Надання студентам знань з використання статистичних, хемометричних, та інших математичних (комп'ютерних) методів обробки масивів хімічних (експериментальних) і теоретичних даних. Зокрема надати студентам знання щодо вирішенні задач по встановленню зв'язку структура – біологічна властивість (QSAR).

1.2. Основні завдання вивчення дисципліни:

Основними завданнями вивчення дисципліни є знайомство студентів із комп'ютерними технологіями для розв'язання типових навчальних та наукових задач QSAR; оволодіння доступними програмними засобами для розв'язання прикладних задач по виявленню зв'язку молекулярної структури із біологічною дією; оцінка біоактивності існуючих систем.

1.3. Кількість кредитів: 9

1.4. Загальна кількість годин: 270

1.5. Характеристика навчальної дисципліни	
За вибором	
Денна форма навчання	Заочна (дистанційна) форма навчання
Рік підготовки	
1-й	1-й
Семестр	
2-й	2-й
Лекції	
32 год.	12 год.
Практичні, семінарські заняття	
непередбачені	
Лабораторні заняття	
64 год.	20 год.
Самостійна робота	
174 год.	238 год.
Індивідуальні завдання	
не передбачені	

1.6. Заплановані результати навчання.

Знати: основні математичні концепції і методи, що мають безпосереднє застосування в сучасній хімії при дослідженні біологічних властивостей речовин; основні теоретичні положення QSAR; володіти інформацією про наявне загальне та спеціалізоване програмне забезпечення, яке може бути використано при вирішенні розрахункових задач в цих галузях.

Вміти: використовувати загальне та спеціалізоване програмне забезпечення та математичні методи, зокрема класифікації та розпізнавання образів,

2. Тематичний план навчальної дисципліни

Розділ 1. Теоретичний матеріал (частина 1)

Комп'ютерні методи представлення та обробки хімічної інформації

Тема 1. Комп'ютерне зображення молекулярних систем. Представлення двовимірних структур. Проблема з формулами. Лінійна нотація. Кодування молекулярного графа. Проблеми представлення структури.

Тема 2. Представлення та обробка стереохімічних даних. Формати файлів. Використання стереохімії у комп'ютерних програмах. Представлення тривимірних структур. Створення 3D структур. Конформаційний аналіз і пошук. Аналіз молекулярної форми. Візуалізація даних.

Тема 3. Представлення хімічних реакцій, бази даних реакцій. Класифікація реакцій. Формати файлів. Хімічні дані. Типи даних. Стандарти форматів обміну спектральними даними. XML і його застосування в хімії.

Тема 4. Хімічні бази даних та системи пошуку. Огляд баз даних і джерел даних. Бібліографічні бази даних. Баз даних хімічних структур. Бази даних хімічних реакцій. Бази даних в біохімії і молекулярної біології. Системи управління лабораторною інформацією. Пошук по двовимірним структурам. Пошук за подібністю.

Тема 5. Високопродуктивна хімія *in silico*. Створення бібліотек. Віртуальний високопродуктивний скринінг.

Тема 6. Фармакофори у пошуку ЛЗ. Створення фармакофорів. Бази даних: підготовка з урахуванням конформерів. Обмеження концепції фармакофорів. Молекулярний докінг. Представлення високомолекулярних рецепторів. Обробка ліганда. Стратегії пошуку у конфігураційному та конформаційному просторі. Скорингові функції. Відбір та оптимізація сполук-лідерів.

Тема 7. Використання віртуального скринінгу для пошуку нових речовин зі смаково-ароматичними властивостями. Віртуальний дизайн смаково-ароматичних речовин та прогнозування їх ольфакторних властивостей.

Розділ 2. Теоретичний матеріал (частина 2)

Квантовохімічні та статистичні методи побудови прогностичних моделей біологічної активності.

Тема 8. Історичний огляд досліджень зв'язку біологічної активності та структури молекул. Методи QSAR (*Quantity Structure-Activity Relationships*). Розрахунки біологічної активності (БА) хімічних речовин як частина проблематики QSAR. Основні етапи оцінки БА. Загальний тезаурус типів БА.

Тема 9. Проблема розробки математичних моделей, що описують зв'язок структури речовини із її властивостями. Основні підходи. Квантова хімія як метод прогнозу та цілеспрямованого створення систем із заданими властивостями.

Тема 10. Систематика квантовохімічних методів. Напівемпіричні методи квантової хімії. Всевалентні наближення – CNDO, MNDO, MINDO, AM1, PM3, RM1, PM6. Їх параметризація. Знайомство з програмними пакетами HyperChem та MOPAC. Розрахунки електронної будови та геометрії біологічно активних сполук.

Тема 11. Логіко-комбінаторний метод у конструюванні лікарських препаратів. Байєсові оцінки ймовірності.

Тема 12. Дескрипторний метод представлення будови молекул (програми DRAGON, ACDLabs, PAdel-DESCRIPTORS, та *інші*). Характеризація молекулярних систем за допомогою дескрипторів і “відбитків пальців”. Топологічні індекси. Тривимірні (геометричні) дескриптори. Представлення молекулярної хіральності.

Тема 13. Аналіз багатовимірних даних. Генетичні алгоритми. Метод головних компонент. Факторний аналіз багатоваріантних даних. Обертання факторів. Регресія на головних компонентах (*Principal Component Regression, PCR*), Неповний метод

найменших квадратів (*Partial Least Squares*, PLS). Порівнювальний аналіз молекулярних полей (CoMFA).

Тема 14. L_1 - та L_2 – регуляризація статистичних розрахунків. Скорочення предикторного (дескрипторного) набору. Метод LASSO (*Least absolute selection and shrinkage operator*).

Тема 15. Метод програмованих нейронних мереж. Типи нейронних мереж. Застосування до проблем класифікації та кластеризації молекулярних систем за біоактивністю. Дискримінаційний аналіз та логістична регресія в проблемі побудови прогностичних моделей біологічної активності.

Розділ 3. Лабораторні заняття (частина 1)

Тема 16. Ознайомлення з форматами хімічних даних. Застосування програм для запису та конверсії хімічних даних.

Тема 17. Представлення та створення тривимірних структур молекул та їх конформерів.

Тема 18. Робота з базами хімічних даних.

Тема 19. Формування бібліотек сполук для скринінгу з використанням типових фільтрів.

Тема 20. Підготовлення фармакофорної моделі на основі структури фермента з лігандом.

Тема 21. Скринінг баз даних з використанням фармакофорних моделей.

Тема 22. Відбір сполук-лідерів та формування списків «хітів».

Розділ 4. Лабораторні заняття (частина 2)

Тема 23. Знайомство з програмним пакетом MORAC. Оптимізація геометрії молекул вєвалентними методами. Порівняння розрахованих даних із експериментальними.

Тема 24. Розрахунок та аналіз дескрипторного набору навчаючої вибірки
Факторний аналіз даних.

Тема 25. Побудова прогностичних моделей біоактивності на основі факторного аналізу.

Тема 26. Регресійна модель біоактивності. Кореляційний аналіз. Кореляційні рівняння в хімії. Крос-валідація як засіб оцінки адекватності та верифікації моделей. Метод найменших квадратів та метод найменших модулів.

Тема 27. Методи PCR та PLS в побудові прогностичних моделей. Метод CoMFA

Тема 28. Кластерний і дискримінантний аналіз. Логістична регресія. Класифікація органічних молекул (наявність/відсутність певного типу активності)

Тема 29. Метод програмованих нейронних мереж в задачах QSAR

3. Структура навчальної дисципліни

Назви розділів	Кількість годин											
	денна форма						заочна форма					
	усього	у тому числі					усього	у тому числі				
		л	п	лаб.	інд.	с. р.		л	п	лаб.	інд.	с. р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Розділ 1. Теоретичний матеріал (частина 1)												
Тема 1	9	2				7		1				10
Тема 2	9	2				7		1				10
Тема 3	9	2				7		1				10
Тема 4	9	2				7		0.5				10
Тема 5	9	2				7		0.5				10
Тема 6	13	6				7		2				10
Тема 7	4	0				4		0				4
Разом за розділом 1	62	16				46	70	6				64
Розділ 2. Теоретичний матеріал (частина 2)												
Тема 8	7	2				5		0.5				8
Тема 9	7	2				5		0.5				8
Тема 10	8	2				6		0.5				8
Тема 11	8	2				6		0.5				8
Тема 12	8	2				6		1				8
Тема 13	8	2				6		1				8
Тема 14	8	2				6		1				8
Тема 15	8	2				6		1				8
Разом за розділом 2	62	16				46	70	6				64
Розділ 3. Лабораторні заняття (частина 1)												
Тема 16	9			4		5				1		7
Тема 17	10			4		6				1		8
Тема 18	10			4		6				1		8
Тема 19	10			4		6				1		8
Тема 20	10			4		6				2		8
Тема 21	10			4		6				2		8
Тема 22	14			8		6				2		8
Разом за розділом 3	73	-		32		41	65			10		55
Розділ 4. Лабораторні заняття (частина 2)												
Тема 23	9			4		5				1		7
Тема 24	11			6		5				1		8
Тема 25	9			4		5				2		8
Тема 26	11			6		5				2		8
Тема 27	9			4		5				2		8
Тема 28	9			4		5				1		8
Тема 29	15			4		11				1		8
Разом за розділом 4	73	-		32		41	65			10		55
Усього годин	270	32		64		174	270	12		20		238

4. Теми семінарських (практичних, лабораторних) занять

№ з/п	Назва теми	Кількість годин	
		денне	заочне
16	Ознайомлення з форматами хімічних даних. Застосування	4	1

	програм для запису та конверсії хімічних даних		
17	Представлення та створення тривимірних структур молекул та їх конформерів	4	1
18	Робота з базами хімічних даних	4	1
19	Формування бібліотек сполук для скринінгу з використанням типових фільтрів	4	1
20	Підготовлення фармакофорної моделі на основі структури фермента з лігандом	4	1
21	Скринінг баз даних з використанням фармакофорних моделей	4	1
22	Відбір сполук-лідерів та формування списків «хітів»	8	1
23	Знайомство з програмним пакетом MORAC. Оптимізація геометрії молекул всевалентними методами. Порівняння розрахованих даних із експериментальними	4	2
24	Розрахунок та аналіз дескрипторного набору навчаючої вибірки Факторний аналіз даних.	6	2
25	Побудова прогностичних моделей біоактивності на основі факторного аналізу.	4	1
26	Регресійна модель біоактивності. Кореляційний аналіз. Кореляційні рівняння в хімії. Крос-валідація як засіб оцінки адекватності та верифікації моделей. Метод найменших квадратів та метод найменших модулів	6	2
27	Методи PCR та PLS в побудові прогностичних моделей. Метод CoMFA	4	2
28	Кластерний і дискримінантний аналіз. Логістична регресія. Класифікація органічних молекул (наявність/відсутність певного типу активності)	4	2
29	Метод програмованих нейронних мереж в задачах QSAR	4	2
	Разом	64	20

5. Завдання для самостійної роботи

Види, зміст самостійної роботи	Кількість годин	
	денне	заочне
Тема 1. Алгоритми пошуку ЛЗ. Планування серії досліджень. Основні засоби хімічної інформатики.	11	13
Тема 2. Створення реакційних файлів для автоматичної побудови набору сполук.	11	15
Тема 3. Застосування OpenBabel та модулів мови Python для обробки експериментальних даних.	11	15
Тема 4. Створення та управління бібліографічними базами даних. Підготовка бази даних сполук для пошуку ЛЗ.	11	15
Тема 5. Застосування генетичних алгоритмів для автоматизованої оптимізації процесу пошуку ЛЗ.	11	15
Тема 6. QSAR та методики цілеспрямованого синтезу.	11	15
Тема 7. Кластерний аналіз хімічних баз даних.	10	15

Тема 8. Критерії відбору смаково-ароматичних речовин. Прогнозування їх фізико-хімічних властивостей за допомогою Perfumery Radar 2.0	11	15
Тема 9. Проблема розробки математичних моделей, що пов'язують структуру речовини із її властивостями. Основні підходи. Квантова хімія як метод прогнозу та цілеспрямованого створення систем із заданими властивостями.	10	15
Тема 10. Систематика квантовохімічних методів. Напівемпіричні та неемпіричні (<i>ab initio</i>) методи. Всевалентні наближення – CNDO, MNDO, MINDO, AM1, PM3, RM1, PM6. Молекулярна механіка. Знайомство з програмними пакетами HyperChem. MOPAC.	12	15
Тема 11. Пакети прикладних програм орієнтовані на проблематику QSAR: Avogadro, HYPERCHEM, DRAGON, ACDLabs, PASS, тощо. Програми візуалізації. Візуалізація біомолекул. Jmol, Pymol, MGLTOOLS	10	15
Тема 12. Метод програмованих нейронних мереж. Генетичні алгоритми. Застосування до проблем класифікації та кластеризації.	12	15
Тема 13. Використання нейронних мереж для прогнозу біоактивності	12	15
Тема 14. Факторний аналіз даних.	11	15
Тема 15. Методи PCR і PLS.	10	15
Тема 16. Скорочення предикторного набору за допомогою методу LASSO. Компактна регресійна модель біоактивності.	10	15
Разом	174	238

6. Індивідуальні завдання

Не передбачено навчальним планом.

7. Методи контролю

Поточний контроль знань (контрольна робота, письмові звіти за результатами лабораторних робіт). Семестровий екзамен (письмова робота).

8. Розподіл балів, які отримують студенти

Поточний контроль, самостійна робота, індивідуальні завдання																Підсумковий семестровий контроль (екзамен)	Сума			
Розділ 1	Розділ 2	Розділ 3						Розділ 4												
Теми 1-7	Теми 8-15	Т 16	Т 17	Т 18	Т 19	Т 20	Т 21	Т 22	Т 23	Т 24	Т 25	Т 26	Т 27	Т 28	Т 29	Т 30	40	100		
К/Р 10 балів	Поточний контроль								Поточний контроль											
	3	3	3	3	3	4	6	3	3	3	3	3	3	3	4					
	Разом: 25								Разом: 25											

1. Студент допускається до підсумкового семестрового контролю (екзамену) за умови виконання та оформлення всіх лабораторних робіт.
2. Якщо студент без поважних причин несвоєчасно виконав та оформив лабораторну роботу чи модульну контрольну роботу, оцінка за роботу знижується на 25%. Терміни оформлення лабораторних і виконання модульних контрольних робіт визначаються викладачами, які ведуть лабораторні заняття.
3. Студент допускається до семестрового підсумкового контролю за умови, якщо за результатами поточного контролю та виконання модульних контрольних робіт він набрав не менше 30 балів.
4. Екзамен вважається зданим, якщо студент одержав за нього оцінку не меншу 10 балів.

Шкала оцінювання

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка
	Для чотирирівневої шкали оцінювання
90 – 100	відмінно
70-89	добре
50-69	задовільно
1-49	незадовільно

Критерії оцінювання навчальних досягнень студентів з дисципліни «Молекулярний дизайн нових біологічно активних речовин»

Знання студентів оцінюється за наступними критеріями:

– «відмінно» – студент міцно засвоїв теоретичний матеріал, здатен вирішувати розрахункові завдання, всебічно орієнтується в усіх основних темах курсу; відповідь на запитання надає переконливі; здатен логічно та обґрунтовано вирішувати поставлені комплексні завдання;

– «добре» – студент добре засвоїв теоретичний матеріал, здатен вирішувати розрахункові завдання та орієнтується в усіх основних темах курсу; надає відповіді на запитання, при цьому можуть бути похибки у логіці викладу та не в повній мірі здатен обґрунтовано вирішувати поставлені комплексні завдання;

– «задовільно» – студент в основному засвоїв теоретичний матеріал, але не в повній мірі здатен використовувати засвоєний матеріал при вирішенні розрахункових завдань, орієнтується не в усіх основних темах курсу; відповідь надає не досить повну; не здатен дати логічну відповідь при вирішенні комплексних завдань;

– «незадовільно» – студент не засвоїв теоретичний матеріал, не орієнтується у більшості основних тем курсу, не здатен виконати розрахункові завдання; не орієнтується в рекомендованій літературі.

9. Рекомендована (базова) література

1. Иванов В.В., Слета Л.А., Расчетные методы прогноза биологической активности органических соединений. Харьков: ЧП Азамаев В.Р., 2003.– 76 с.
2. Young D.C. Computational Chemistry, Wiley Interscience, New York, 2001.– 370 p.
3. Орлов В. Д. Липсон В.В., Иванов В.В., Медицинская химия, Харьков-2018: Эксклюзив.– 552 с.
4. В. В. Иванов, Л. А. Слета. Квантовая химия. Харьков: “Фолио”, 2007.– 443 с.
5. Станкевич М.И., Станкевич И.В., Зефиоров Н.С., Топологические индексы в органической химии // Успехи химии.–1988.–Т. 57, № 3.– С.337-365.
6. Раевский О.А., Дескрипторы молекулярной структуры в компьютерном дизайне биологически активных веществ // Успехи химии.–1999.–Т. 68, №6. – С. 555-575.

7. Факторный, дискриминантный и кластерный анализ, М. "Финансы и статистика", 1989.– 213 с.
8. G. Turrel. *Mathematics for Chemistry and Physics*.- Elsevier 2002.– 424 p.
9. *Справочник по прикладной статистике* в 2-х т., под ред. Э. Ллойда, У. Ледермана, Ю. Н. Тюрина. - М.: Финансы и статистика, 1989.– 1990.
10. Ю.Н. Тюрин, А.А. Макаров. (под ред. Фигурнова). *Анализ данных на компьютере*.– М.: Фин. и стат-ка, 1995. – 384с.
11. Демиденко Е.З. *Линейная и нелинейная регрессии*. - М.: Финансы и статистика, 1981.– 304с.
12. *Neural Networks in QSAR and Drug Design*, ed. J.Devillers, Academic Press, NY, 1996.– 284 p.
13. *Recent Advances in QSAR Studies. Methods and Applications*, Edited by T. Puzyn and J. Leszczynski, Springer, 2010.– 423 p.

10. Посилання на інформаційні ресурси в Інтернеті, відео-лекції, інше методичне забезпечення

1. <http://www.click2drug.org/>
2. <http://autodock.scripps.edu/>
3. <https://www.chemcomp.com/Products.htm>
4. <https://www.schrodinger.com/science>