

Кратко о структурной базе данных

Двумерная структура в запросе

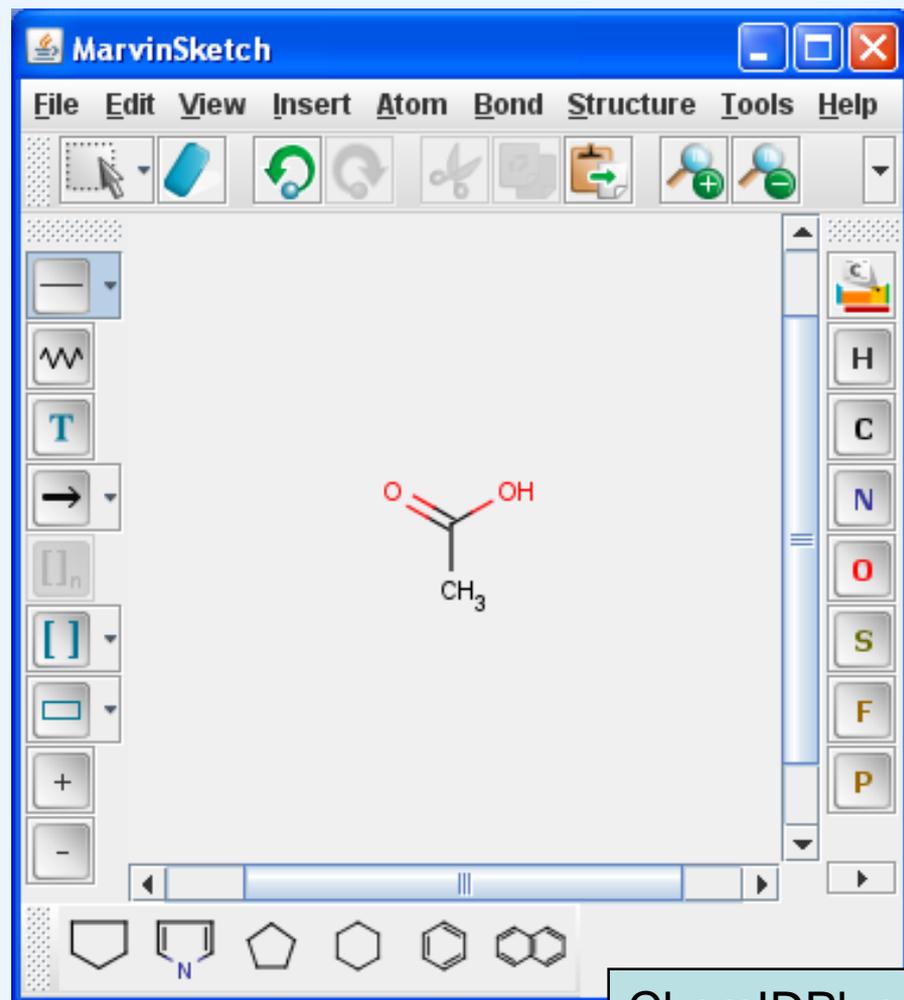
Формирование структуры на поисковом бланке (апплет).
Загрузка файла пользователя (обычно MOL).

Search for Species Data by Structure or Substructure

There are three structure search options available:

1. Use applet to draw a structure.
This option requires a Java capable browser.
2. Submit a mol (MDL) file containing the structure. This option requires a browser which can upload files.
3. Specify structure properties and subgroup form.

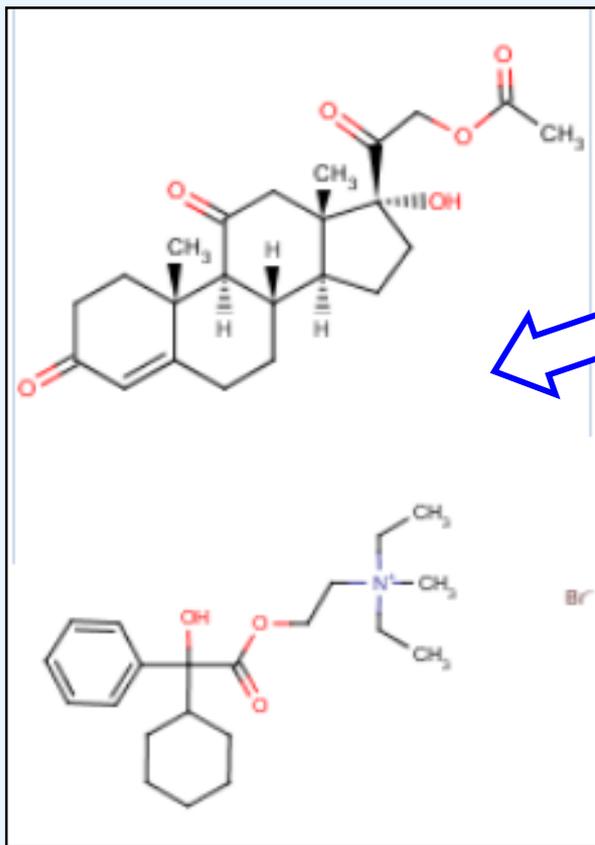
NIST



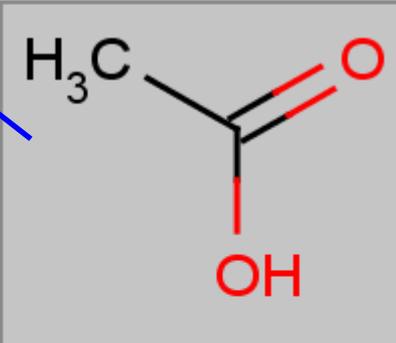
ChemIDPlus

Структура, подструктура (субструктура)

Запрос



View Help

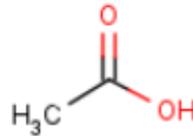


Powered by [ChemAxon Marvin](#)

Structure Search Options ⓘ

- Substructure Search
- Similarity Search 80 %
- Exact (parent only)

1 [Acetic acid, glacial \[USAN:JAN\]](#)
64-19-7



Поиск идентичной структуры и поиск структур, имеющих заданный остов, - профессионалы такие задачи считают тривиальными.

Молекулярное (химическое) подобие

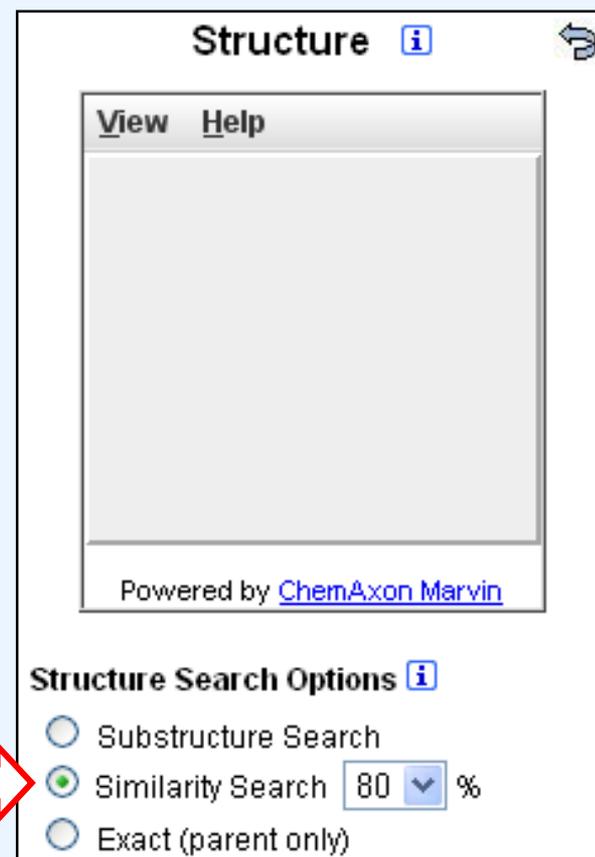
Молекулярное (химическое) подобие Similarity

Молекулярное подобие - это близость, сходство, подобие **структур** химических соединений.

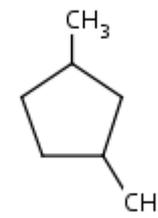
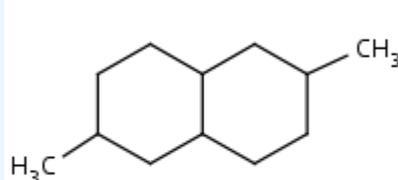
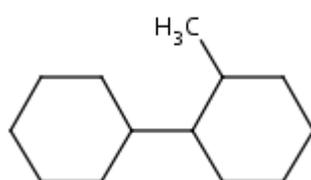
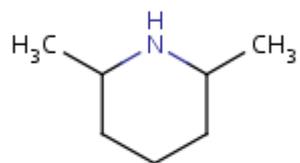
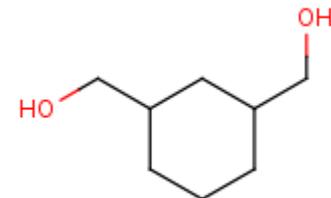
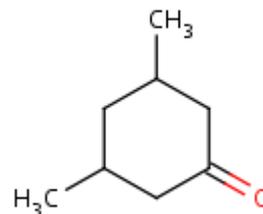
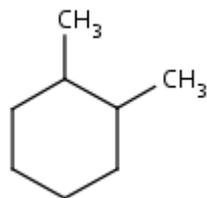
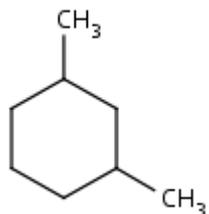
Предполагается, что **подобные соединения обладают близкими химическими свойствами**, в т. ч. подобной биологической активностью.

(Предположение не всегда верно).

поиск подобных структур



Пример: подобные структуры



Как количественно охарактеризовать степень подобия?

Строка битов (Bitstring)

Строка битов – отображение структуры последовательностью бинарных чисел 0 и 1.

Формируют словарь признаков (например, разные группы атомов).

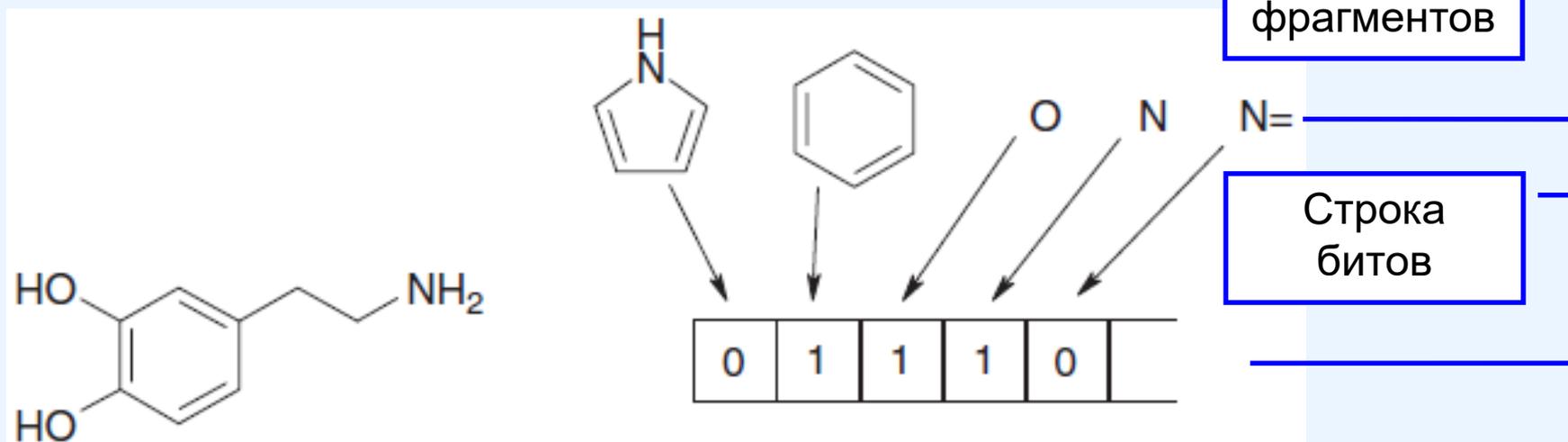
В строке битов на i -й позиции отмечают i -й признак:
1 – признак имеется,
0 – признак отсутствует.

Для каждой структуры генерируют строку битов, например:

011100110001000...

Структурный код (Structural key)

Структурный код – строка битов, в которой отражено наличие или отсутствие определенного фрагмента структуры.



Полученная строка битов – **отпечаток пальцев (fingerprint)** данной *двумерной* структуры.

Количественная оценка подобия

Коэффициент Танимото

Сравнивают отпечатки пальцев двух структур.

Напоминание: отпечаток пальцев – строка битов.

1	0	1	1	1	0
---	---	---	---	---	---

Для двух структур A и B:

S_{AB} – коэффициент Танимото,

a – количество "единиц" у A,

b – количество "единиц" у B,

c – количество "единиц",
общих для A и B.

$$S_{AB} = \frac{c}{a + b - c}$$

$$0 < S_{AB} < 1$$

Пример расчета коэффициента Танимото

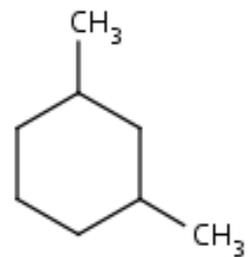
A	1	0	1	1	1	0	1	1	0	0	1	1	$a=8$
			↕	↕			↕				↕	↕	$c=5$
B	0	0	1	1	0	0	1	0	1	0	1	1	$b=6$

$$S_{AB} = \frac{5}{8+6-5} = 0.56$$

Структуры А и В подобны на 56 %

Structure [i](#) 

View Help



Powered by [ChemAxon Marvin](#)

Structure Search Options [i](#)

Substructure Search

Similarity Search [v](#) %

Exact (parent only)

Увеличить
или уменьшить
список
результатов
поиска