

***Человек, не знающий  
математики, не способен  
ни к каким другим наукам  
Бэкон***

**Современные  
представления о структуре  
и действии лекарственных  
веществ**

# **Занятия №4-5**

**Количественные  
соотношения  
«структура – БА»**

# Поиск БАВ

- **Скрининг**
- В 1966 году США для поиска противоопухолевых препаратов **испытано 114 000** соединений, **отобрано 12** как потенциальные лекарства, но ни одно не стало выдающимся препаратом

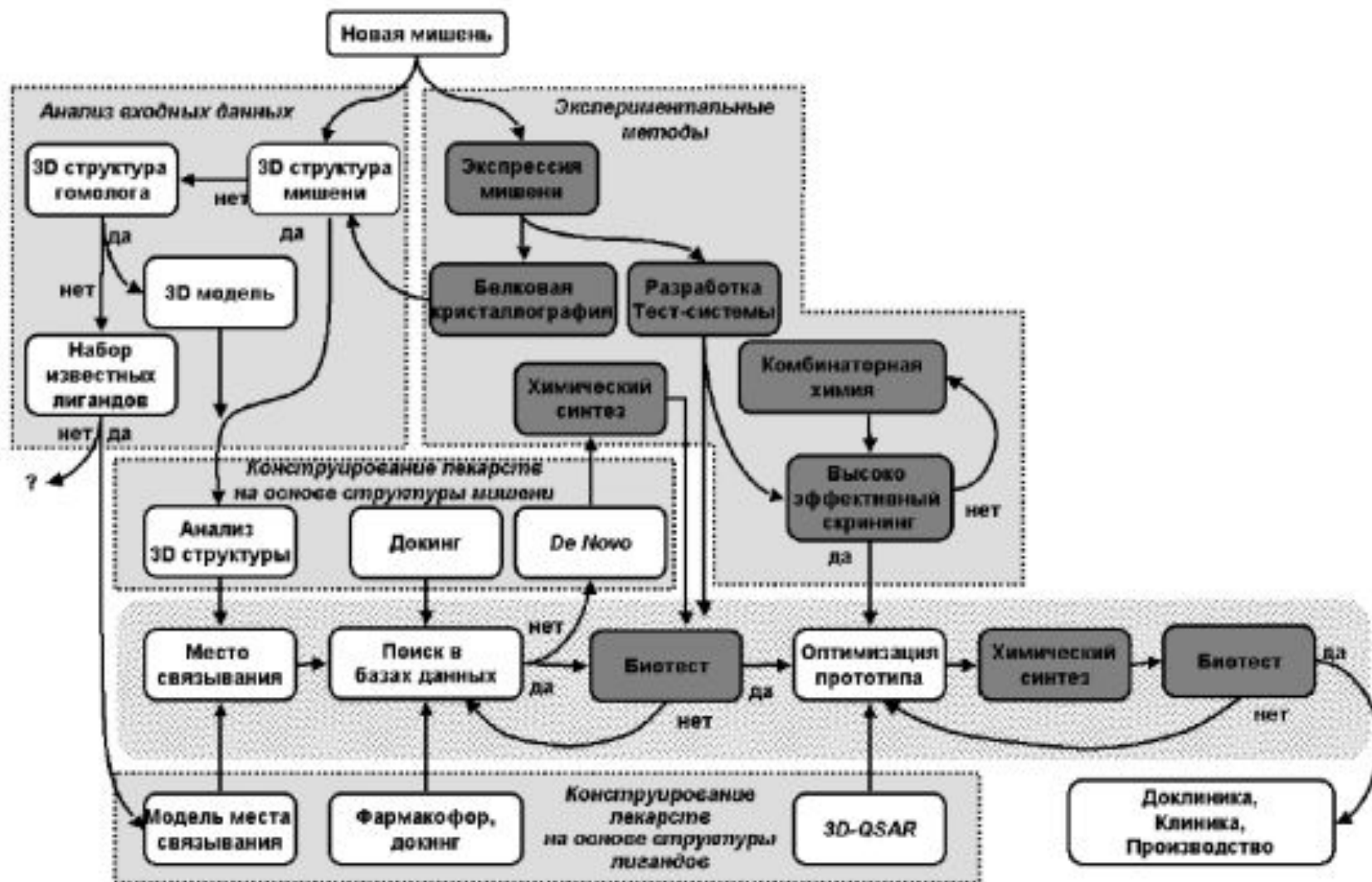
# Поиск БАВ

- Экономическая эффективность скрининга возрастает с ***увеличением числа тестируемых активностей***

# Поиск БАВ

- ***Математическое моделирование*** – анализ связи структуры веществ и биологической активностью

# Схема стратегии компьютерного конструирования лекарств





# Моделирование

- Применение нейронных вычислительных сетей при построении моделей «структура - СВОЙСТВО»

**Использование структурных  
параметров в  
количественных  
соотношениях «структура –  
биологическая активность»**

# Количественные соотношения «структура – БА»

- Наличие корреляций позволяет делать некоторые **заключения о природе действия ХС**

# Количественные соотношения «структура – БА»

- Была предсказана *антибутирилхолинэстеразная* активность **1-децил-3-(N-метилкарбамоил)пиперидина** гидробромида за 3 года до его синтеза

# Количественные соотношения «структура – БА»

- Позволяет найти **оптимальное лекарственное средство** «на кончике пера»
- 62 соединения фениламиндомочевины с антималярийным действием – уравнение – прогноз новых ХС – их синтез – испытание (оказались активнее известных)

# Показатели БА

- $LD_{50}$  – летальная доза для 50%
- $LK_{50}$  – летальная концентрация для 50% особей
- $NK_{50}$  – наркотическая концентрация (вызывает боковое положение 50% мышей)
- $C_{min}$  – пороговая концентрация, вызывающая изменения характеристики рефлекса

# Показатели БА

- ПДК – предельно допустимая концентрация
- Наиболее *низкие коэффициенты* зависимости устанавливаются для ЛД<sub>50</sub>
- Наиболее *лучшие корреляции* - для НК<sub>50</sub>

# Параметры структуры

- **Надмолекулярный** уровень ( $S_r$ ,  $S_{ж}$ ,  $dt_{кип}/dp_0$ )
- **Молекулярный** уровень [ $M$ ,  $d$ ,  $V_{(молек)}$ ,  $d_{крит}$ ,  $P_{крит}$ ,  $\chi$  (теплопроводность),  $C_p$  (теплоемкость),  $T_{пл}$ ,  $T_{кип}$ ,  $P_n$  (давление насыщенных паров),  $\sigma$  (поверхностное натяжение),  $\eta$  (вязкость),  $IgP$  или  $IgK_{ом}$ ,  $S_p$  (растворимость),  $n_D^{20}$ ,  $MR$ ,  $\mu$ ,  $\epsilon$ ,  $I$  (потенциал ионизации),  $D_a$  (абсолютная дисперсия), константы Гаммета,  $E_{1/2}$ ]



# Параметры структуры

- **Наилучшие корреляции** наблюдаются для свойств молекулярного уровня (но  $\eta$  и  $P_n$  не дают высоких корреляций)
- **Квантово-химические** параметры (энергии ВЗМО, НСМО,  $\Delta E$ , заряды на атомах, вклады отдельных атомов в МО и т.д.)

# Количественные соотношения «структура – БА»

- *Физическая интерпретация* или объяснения полученным зависимостям
- Бактериостатическая активность связана  $\approx 18$  показателями ХС
- Токсическое действие  $\approx 26$  показателей

# Количественные соотношения «структура – БА»

- Возможно **превращение ХС** под действием физических факторов (фото, термических, электрохимических), химических веществ
- Связь структура – БА очень **сложна**
- Нет **универсальных характеристик** ХС в определении БА

# Количественные соотношения «структура – БА»

- При поиске корреляций используется ***регрессивный анализ***

# Параметры структуры

- Х. Мейер (1899 г) и Е. Овертон (1901 г) предложили, что инертные вещества, вызывающие сон или наркотическое состояние и растворимые в липидах, способны вызывать эти состояния с эффективностью ~ **коэффициенту распределения** (липид-вода)  
$$K_{ом} = C(\text{олифковое масло})/C(\text{вода})$$

# Коэффициент распределения

- *Гемолитическое действие* некоторых неэлектролитов от коэффициента распределения (оливковое масло – вода)

$$\log C_{\text{гем}} = - 0.62 \log K + 2.45$$

# Параметры структуры

- Количественные *показатели гидрофобности ХС:  $I_d$  растворимости, работа адсорбции в системе вода – гептан (вода-воздух), энергия взаимодействия с неполярными органическими растворителями, поверхностная активность*

# Параметры структуры

- **Лазарев Н.В.** – отечественный фармаколог – 1944 г. «Неэлектролиты. Опыт биолого-физико-химической их систематики»
- Главное внимание *адгезии* веществ ( $\approx$  гидрофильность)



# Классификация неэлектролитов

- **А** – сильные неэлектролиты (сильно притягиваются молекулами воды)
- **В** - слабые неэлектролиты (слабо притягиваются молекулами воды)
- **С** – переходные электролиты

# Классификация неэлектролитов ( $K_{ом}$ )

- **A** – гр.1 ( $<10^{-2}$ ), гр.2 ( $10^{-2} \dots 10^{-1}$ )
- **B** – гр.3 ( $10^{-1} \dots 10^0$ ), гр.4 ( $10^0 \dots 10^1$ ), гр.5 ( $10^1 \dots 10^2$ )
- **C** – гр.6 ( $10^2 \dots 10^3$ ), гр.7 ( $10^3 \dots 10^4$ ), гр.8 ( $10^4 \dots 10^5$ ), гр.9 ( $>10^5$ )
- **Ряды:** 1 ряд – вещества с  $M < 100$   
2 ряд –  $M = 100 \dots 200$  и т.д.

# Классификация неэлектролитов ( $K_{om}$ )

- В гомологическом ряду переход от *низших* групп к *высшим*
- Специфическое *действие* ~ *номеру группы* (зависимость параболическая)

# Коэффициент распределения

- Количественные связи между БА и коэффициентом распределения (октанол-вода), который характеризует ***гидрофобность веществ*** (можно определять экспериментально, можно рассчитать на основании структуры ХС)

# Коэффициент распределения

Для расчета параметров гидрофобности соединений использовали метод Ханча<sup>\*</sup>, в основе расчетов которого лежит предположение о существовании зависимости между факторами биологической активности и физико-химическими параметрами веществ в гомологических рядах химических соединений. Гидрофобность соединения по методу Ханча описывается логарифмом коэффициента распределения соединения ( $\lg P$ ), и вычисляется по формуле:

$$\lg P = \sum a_j \pi_j, \quad (16)$$

где  $\pi_j$  – параметр гидрофобности фрагмента;  $a_j$  – число фрагментов данного типа в структуре.

# Коэффициент распределения

- Расчет с помощью **ACD Labs**

## Задание

- Рассчитать  $\log P$  для ряда лекарственных препаратов

# Уравнение Ханча

$$A = -\psi_1 \cdot \ln^2 P + \psi_2 \cdot \ln P + \psi_3$$

- где  $A$  –  $1/c$ ,  
 $\psi_i$  – коэффициенты, зависящие от структурных особенностей
- Хорошие корреляции для *цитостатических свойств производных **нитрозомочевины**, канцерогенной активности полициклических аренов*

# Задание

- Какой интервал коэффициента распределения является оптимальным для биологической активности фенольных бактерицидов?



# Коэффициент распределения

- Хорошо коррелирует с **токсичностью** для веществ с высокой гидрофобностью
- Для гидрофильных соединений вклад в корреляционные соотношения вносят **степень ионизации, заряд иона,  $\mu$ , способность к образованию водородных связей** и т.п.

# Степень диссоциации

- Для электролитов:

$$BA \sim \alpha$$

- *Антибактериальная* активность *трифенилметановых красителей* (Альберт, 1941)
- Минимальная концентрация *производных акридина* (101 соедин.) для бактерий *Streptococcus pyogenes* (48 ч инкубации)

# Задание

- Найти корреляционную зависимость
- 9-аминоакридин  $1/C=16 \cdot 10^4$   $\alpha=99$
- 3-аминоакридин  $1/C=8 \cdot 10^4$   $\alpha=72$
- 2-аминоакридин  $1/C=1 \cdot 10^4$   $\alpha=2$
- 1-аминоакридин  $1/C=1 \cdot 10^4$   $\alpha=2$
- 3,9-диамино.....  $1/C=16 \cdot 10^4$   $\alpha=100$

**П (парахор)**

$$\mathbf{P} = (\mathbf{M}/\rho) \cdot \sqrt{\sqrt{\sigma}}$$

**M** – молярная масса

**ρ** - плотность

**σ** - поверхностное натяжение жидкости

# Корреляции БА – П (парахор)

Получены зависимости:

- Для замещенных в 5-м положении **пентиламинов** препятствовать *свертыванию крови*
- Активность *местных анестетиков*
- *Антимикробная активность пенициллинов*

# Корреляции БА – П (парахор)

Получены зависимости:

- *Фибринолитическая активность для **2- и 4-замещенных бензойных кислот** и **5-R-салициловых кислот***
- *Biochem Pharmacol. 1975, V. 24, №10, P. 1103-1109*

# Задание

- Составить корреляционные уравнения БА ХС с парахором (П)

# Свойства и действие ХС

- **Рише** (1893 г) установил, что некоторые вещества по силе их наркотического действия на мелких рыбок располагаются в *обратной зависимости* от их *растворимости* в воде



# Физические параметры

- БА ~ физические параметры  
( $T_{кип}$ ,  $T_{пл}$ ,  $\rho$ ,  $n_D^{20}$ ,  $\varepsilon$ ,  $T_{кип}/M$ ,  
 $T_{кип}/V(\text{молек})$ ,  $T_{кип}/MR$ ,  $T_{пл}/M$ )
- Зависимости ***не дают  
высоких корреляций***

# Задание

- Найдите корреляционные уравнения, связывающие физические, или химические, или физико-химические характеристики вещества с видом биологической активности

# Хроматографические параметры

- Хроматографические характеристики сравнительно легко определяются экспериментально и хорошо **коррелируют** со многими физико-химическими и структурными параметрами

# Хроматографические параметры

- Использование *хроматографических методов* для изучения зависимости «структура - биологическая активность»
- Корреляции типа «**удерживание** — **биологическая активность**»

# Задание

- Найдите корреляционные уравнения, связывающие хроматографические характеристики веществ с данными по их биологической активности

# Спектральные характеристики

- Параметры **флуоресценции** ( $\lambda_{\text{фл}}$  и квантовый выход - связывание с мишенью) ~ БА
- Многомерные вектора спектральных характеристик ХС и БА исследовать трудно, необходимо **упрощение** на подмножества

# Спектральные характеристики

- Возможны *ошибки*
- Отсутствие у *стероидов*  $\text{C}\text{H}_3$ - групп (C18), лишаящее их активности, практически не скажется на их УФ спектрах

# Спектральные данные

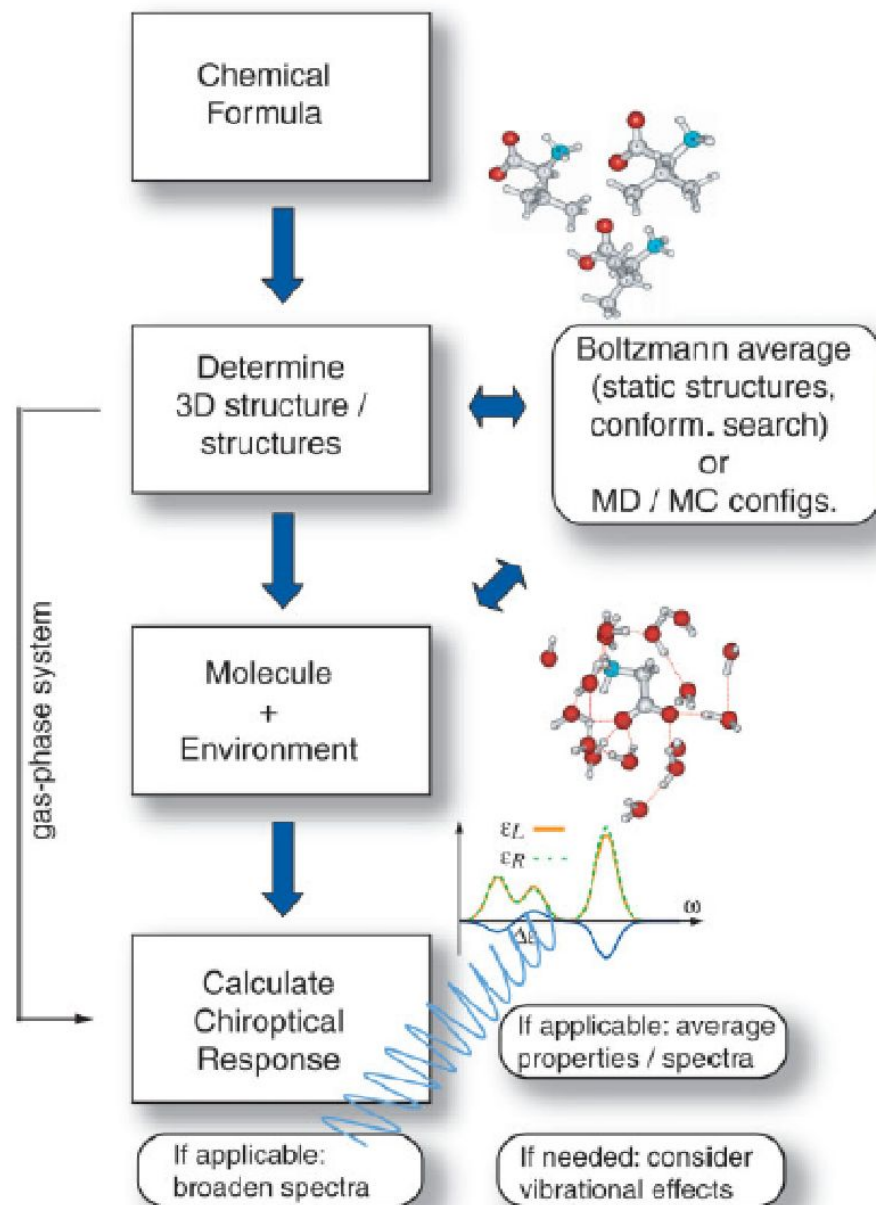
- В отдельных случаях выявляются корреляционные зависимости биологической активности со **специфическими пиками ИК спектроскопии**
- Для **замещенных бензолов** – колебания ***C-H* связей** вне плоскости ароматического кольца



# Задание

- Найдите корреляционные зависимости биологической активности выбранных вами химических соединений со специфическими пиками в их ИК спектрах

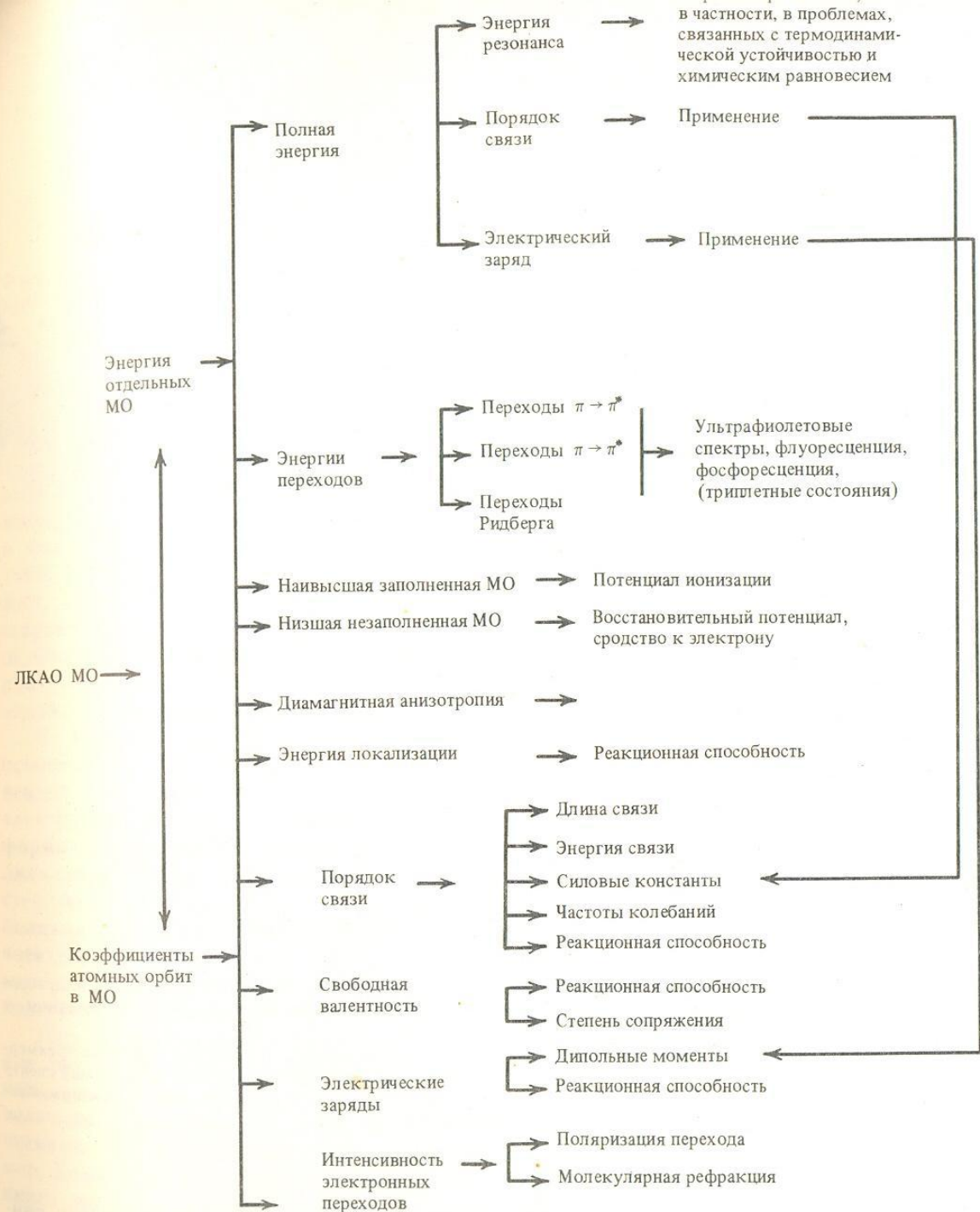
# Компьютерное моделирование (квантово-химические параметры)



# Основы компьютерного прогнозирования связи "структура-свойство"

- Квантовохимические расчеты с помощью комплекса программ **HyperChem**
- Выбрать *стерические, электронные, поляризационные свойства* молекул

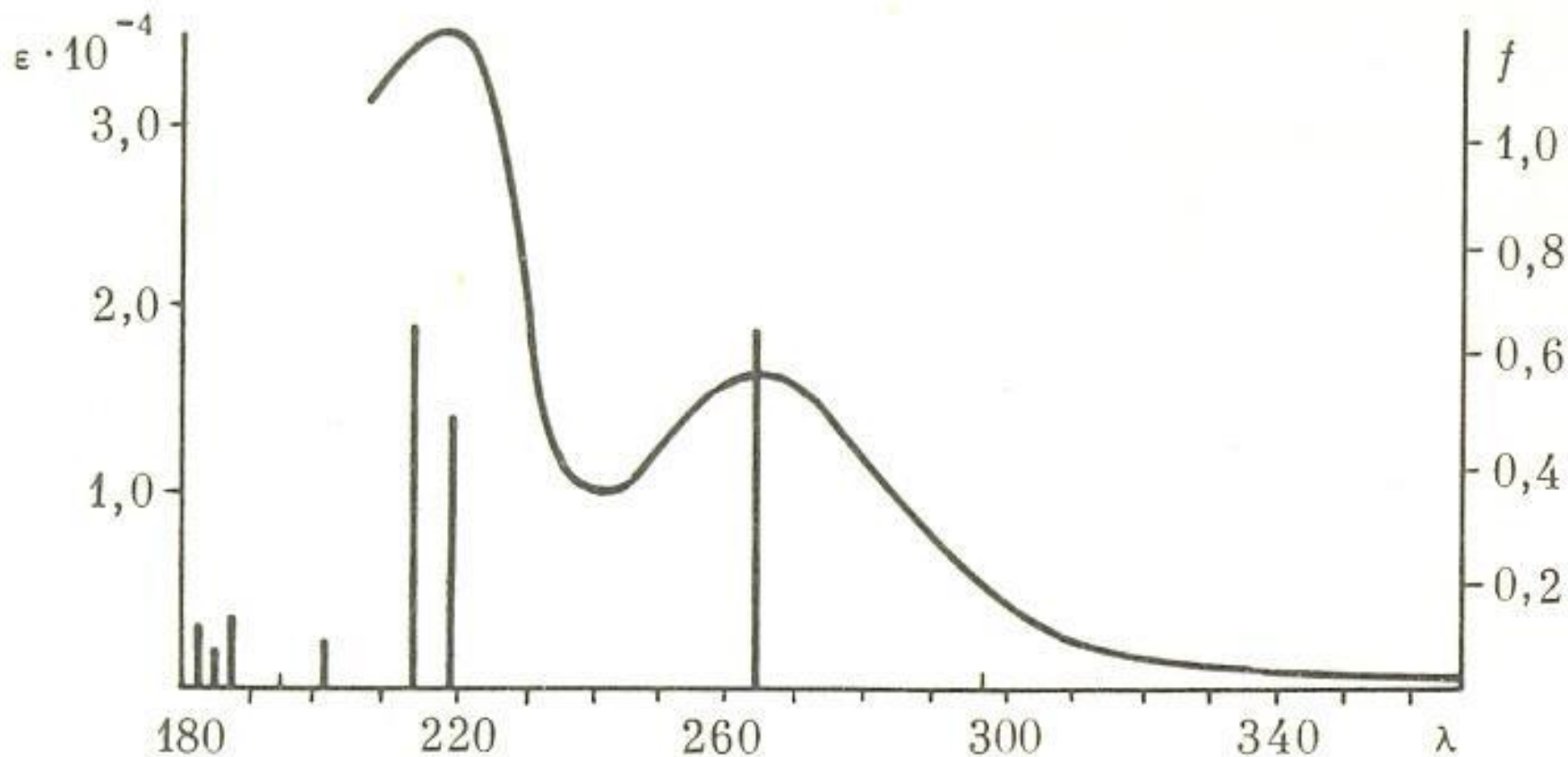
# Электронные характеристики



# Основы компьютерного прогнозирования связи "структура-свойство"

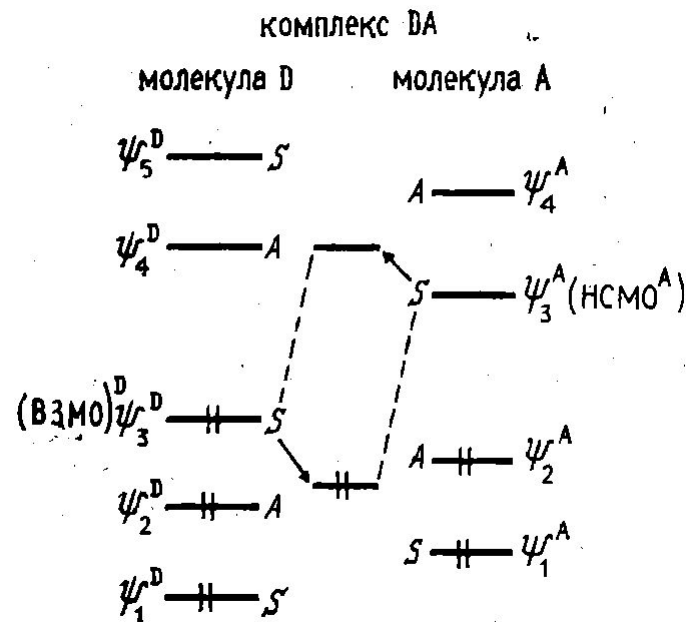
- Экспериментальные данные используются как ***критерий правильности*** теоретических расчетов

# Экспериментальный и теоретический спектры УФ



# Энергия орбиталей

- Есть корреляции **психотропной активности** препаратов со значениями энергий их **ВЗМО**



# Ряд психотропных препаратов

- Дофамин
- Мескалин
- 2,5-диметокси-4-метил-амфетамин
- N,N-диметилтриптамин
- Серотонин
- Хлорпромазин
- Диэтиламид лизергиновой кислоты (ЛСД)
- **Предсказать ряд активности по энергиям ВЗМО**



# Потенциалы ионизации

- В оценках *нейротропной активности* следует учитывать не только первый, но и последующий ***потенциалы ионизации*** молекул препаратов

# Потенциалы ионизации

- Корреляция биологической активности со значениями потенциалов ионизации соответствует ***механизму переноса заряда***, при котором электронодонорная π-система молекулы психотропного препарата вступает во взаимодействие с электроноакцепторным центром рецептора

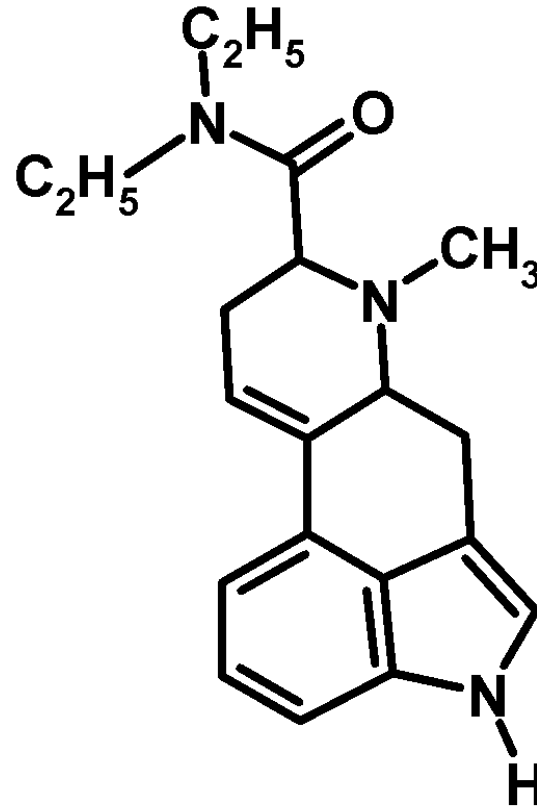
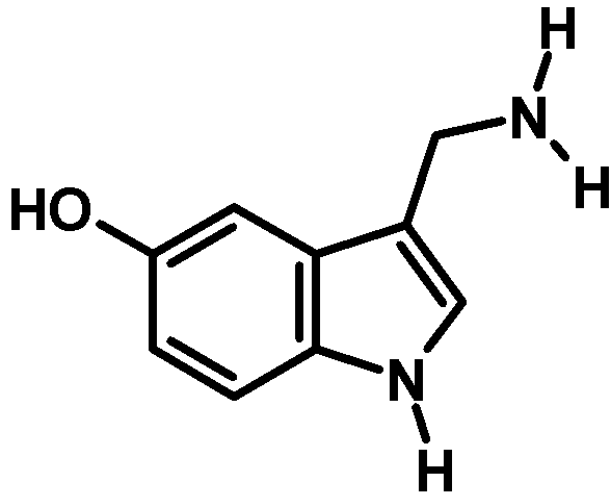
# Энергия ВЗМО

- Этим взаимодействием может быть *комплексообразование*, *электрофильное замещение* по ароматическому кольцу препарата, образование  *$\pi$ -связи* рецептора с ароматическим кольцом и т.д.
- Роль *граничных орбиталей* (ГО) в определении психотропного действия представляется весьма *убедительной*

# Распределение электростатического потенциала

- Одним из примеров корреляции *биологической активности с распределением электростатического потенциала* в молекуле может служить изучение взаимодействия **ЛСД** и **серотонина** (и его аналогов – **изомерных гидрокситриптофанов**) с соответствующим рецептором

# ЛСД и серотонин



# Распределение электростатического потенциала

- Были *рассчитаны контуры молекулярных электростатических потенциалов*, показывающие векторы предпочтительной *ориентации молекулы* по отношению к полю, генерируемому положительным зарядом

# Распределение электростатического потенциала

- Оказалось, что направления этих векторов существенно *различаются* для различных *гидрокситриптаминов*: вектор молекулы *6-гидрокситриптамина* почти *перпендикулярен* вектору *5-гидрокситриптамина*

# Распределение электростатического потенциала

- Такое различие в условиях взаимодействия с рецептором может быть *причиной неодинаковой биологической активности серотонина и его аналогов*



# Распределение электростатического потенциала

- Карта электростатического потенциала **серотонина** в значительной мере воспроизводится в молекуле **ЛСД**, что и объясняет их чувствительность к одному и тому же рецептору (оба соединения способны оказывать сильные эффекты на деятельность центральной нервной системы)

# Уравнение Гаммета



$$\lg(k/k_0) = \sigma\rho, \text{ где}$$

$k_0$  - константа скорости или равновесия при  $X=\text{H}$ ,

$k$  - константа скорости при замещении водорода группой  $X$

$\rho$  - константа, характеризующая данную реакцию в данных условиях

$\sigma$  - константа, характеризующая заместитель  $X$  (мера суммарных электронных эффектов)

# Уравнение Гаммета

- Принято  $\rho=1$  для диссоциации *мета*- и *пара*-замещенных бензойных кислот в воде при 25°C
- Были рассчитаны  $\sigma_m$  и  $\sigma_p$  для различных заместителей
- Можно найти  $\rho$  для других реакций

Зная  $\rho$  и  $\sigma$ , можно предсказать скорость реакции до ее проведения (если механизм одинаков)

# Уравнение Гаммета

- К *орто*-заместителям этот подход не оправдывается
- Браун предложил  $\sigma^+$  для реакций, при которых электронодонорная группа взаимодействует с образующимся в переходном состоянии положительным зарядом

# Уравнение Гаммета

- Предложены  $\sigma$ - (Тафт) для реакций, при которых образуется отрицательный заряд в переходном состоянии

# Уравнение Гаммета

- Существуют ***таблицы*** с этими константами
- Аналогичные расчеты можно проводить для других циклических систем - ***нафталина, гетероциклов, для этиленовых, ацетиленовых*** углеводородов

# Уравнение Гаммета

- Уравнение Гаммета представляет собой **линейное соотношение свободных энергий**:

$$\Delta G = -R \cdot T \cdot \ln K$$

$$\lg K - \lg K_0 = \sigma \rho$$

$$-\Delta G / 2.3RT + \Delta G_0 / 2.3RT = \sigma \rho$$

$$-\Delta G = \sigma \rho \cdot 2.3 \cdot RT - \Delta G_0$$

# Уравнение Гаммета

Для БА ХС

$$\lg(A/A_0) = \sigma\rho, \text{ где}$$

$A_0$  – БА для соединения с  $X=H$ ,

$A$  - БА для ХС, с замещенным водородом группой  $X$

$\rho$  - константа, характеризующая данную БА

$\sigma$ - константа, характеризующая заместитель  $X$  (мера суммарных электронных эффектов)



# Аддитивная модель Фри-Вилсона

- БА -  $\Sigma$  активности ***типového соединения*** в гомологическом ряду и активностей ***заместителя***

# Задание

- Составить уравнения оценки вклада заместителя в общую активность ХС

$$\log(1/c) = X + \log(1/c_0)$$

# Количественные соотношения «структура – БА»

- Наличие корреляций позволяет делать некоторые **заключения** **о природе действия ХС**

# Нелинейные эффекты

- Уравнения описывают *линейную аддитивную комбинацию* свойств заместителей в молекуле, между тем ***нелинейные взаимодействия*** вносят существенный вклад в биологическую активность

# Нелинейные эффекты

- Обусловлены:
- - кинетикой транспорта и распределения соединения в биологической системе
- - лимитом пространства для взаимодействия гидрофобных групп с сайтом связывания
- - аллостерическими эффектами (связывание ХС с участками фермента вне активного центра влечет за собой конформационные изменения и его активности)

# Нелинейные эффекты

- Обусловлены:
- - возрастающим метаболизмом липофильных аналогов
- - ингибированием конечного продукта липофильными продуктами ферментных реакций
- - ограниченной растворимостью высших аналогов
- - принципом минимальной оккупации рецептора

# Уравнение Ганча

- Для математического описания основных нелинейных эффектов, связанных с липофильностью, Ганч сформулировал ***параболическую модель***

# Трехпараметровые уравнения Ганча-Фуджиты

- Использование характеристик электронной структуры (**константы Гаммета**), стерических взаимодействий (**стерические константы Тафта**) и транспортных свойств (**коэффициент распределения в системе октанол - вода**)



# Уравнение Ганча

$$\log(A) = a\sigma + b\pi + c\pi^2 + dE_s + eMR + f$$

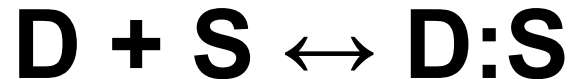
- где  $\pi$  - липофильность (гидрофобность),  $E_s$  -стерические факторы Тафта,  $MR$  - молярная рефракция и дисперсионные силы,  $\sigma$  - константы Гаммета
- Величины  $\sigma$ ,  $\pi$ ,  $E_s$ ,  $MR$  относятся к отдельным заместителям и определяются по таблицам

# Модель Ганча

- Была сделана **компьютерная модель** транспорта в биологической системе с гипотетическими константами и она не показала существенного отклонения от сторон параболы, что говорит о хорошем соответствии **модели Ганча** конкретным взаимодействиям

# Взаимодействие ХС с рецептором

- 1) *Обратимое образование комплекса* вещества D с рецептором S



- 2) *Необратимое превращение комплекса* в продукты



# Взаимодействие ХС с рецептором

- Скорость образования продукта

$$dQ/dt = k_2[D:S]$$

- Константа равновесия

$$K_1 = [D:S]/([D][S]) = \exp(-\Delta G/RT)$$

$\Delta G$  – свободная энергия Гиббса образования комплекса

# Взаимодействие ХС с рецептором

- Скорость образования продукта

$$dQ/dt = k_2[D][S]\exp(-\Delta G/RT)$$

- После интегрирования

$$Q = k_2[D][S]\exp(-\Delta G/RT)t$$

- Так как

$$[D] = AD_0$$

**A** – вероятность достижения рецептора  
(постоянная для данного ХС)

# Взаимодействие ХС с рецептором

$$\ln(1/D_0) = \ln A - \Delta G/RT + \ln|S| + \ln(t/Q) + \ln k_2$$

Где  $D_0$  – ЛД<sub>50</sub>, ЕД<sub>50</sub> и т.д.

- Принимая, что механизм действия одинаков -  $\ln|S|$ ,  $\ln(t/Q)$ ,  $\ln k_2$  – const

$$\ln(1/D_0) = \ln A - \Delta G/RT + C$$

# Взаимодействие ХС с рецептором

- Из подхода Гаммета-Тафта

$$\Delta G \sim \sum \sigma_i \text{ и } \sum E_i$$

$$\ln(1/D_0) = \ln A + a_1 \sum \sigma_i + a_2 \sum E_i + a_0$$

- **A** – связана с транспортом вещества до рецептора
- **Модель**: вещество путем диффузии проходит поочередно через водные и липидные фазы

# Взаимодействие ХС с рецептором

- Принимают, что объем всех фаз и поверхность раздела между ними одинаковы
- Считаем, что константы скоростей диффузии из воды в липид и обратно не зависят от номера фазы



# Концентрация ХС в $i$ -фазе

$$\frac{dC_1}{dt} = -k_{12}C_1 + k_{21}C_2,$$

$$\frac{dC_{2i}}{dt} = -2k_{21}C_{2i} + k_{12}(C_{2i-1} + C_{2i+1}),$$

$$\frac{dC_{2i+1}}{dt} = -2k_{12}C_{2i+1} + k_{21}(C_{2i} + C_{2i+2}),$$

$$\frac{dC_{n-1}}{dt} = \begin{cases} -(k_{21} + k_m)C_{n-1} + k_{12}C_{n-2} & (n - \text{нечетное}) \\ -(k_{12} + k_m)C_{n-1} + k_{21}C_{n-2} & (n - \text{четное}), \end{cases}$$

$$\frac{dC_n}{dt} = k_m C_{n-1},$$

# Коэффициент распределения

- Вводят новую переменную  $\ln P$ , где  $P$  – коэффициент распределения между водой и липидной фазой (от 0.001 до 1000)

$$P = k_{12}/k_{21}$$

- Решая систему уравнений вычисляли  $\ln C_n$  – полученные точки хорошо обрабатываются параболой

# Коэффициент распределения

$$\ln C_n = a \ln^2 P + b \ln P + c$$

- Так как,  $C_n = A \cdot D_0$

$$\ln A = a \ln^2 P + b \ln P + c.$$

$$\ln(1/D_0) = \ln A + a_1 \sum \sigma_i + a_2 \sum E_i + a_0$$

# Уравнение Ганча

- Методология Ганча хорошо работает в узких рядах гомологичных соединений
- *Как* отнести новое ХС к определенному ряду?

**Методология  
QSAR  
(КССА)**

Элементарный состав  
химического соединения и  
расположение его атомов и  
связей между ними в  
пространстве полностью  
определяют физические,  
химические свойства  
соединения, а также его  
биологическую активность

# QSAR

Quantitative Structure-Activity  
Relationships

# QSAR

- Направление, возникшее, на стыке ***органической химии, хемометрики, математического моделирования и компьютерной химии***



# QSAR

- Целью методов QSAR является поиск корреляций между любой из форм биологической активности и структурными, физико-химическими и квантово-химическими параметрами исследуемого химического соединения

**КССА**

**Количественные  
соотношения структура -  
активность**

# QSAR

- В основе метода QSAR лежит утверждение о том, что активность и параметры соединения связаны известной математической функцией:

$$BA = F(S)$$

где **BA** - биологическая активность,  
**S** - физико-химические и/или  
квантово-химические характеристики  
исследуемого соединения

# QSAR

- Функция  $F$  обычно представляет собой математическое выражение, полученное с помощью статистических методов (примером может служить широко используемая многофакторная линейная регрессия)

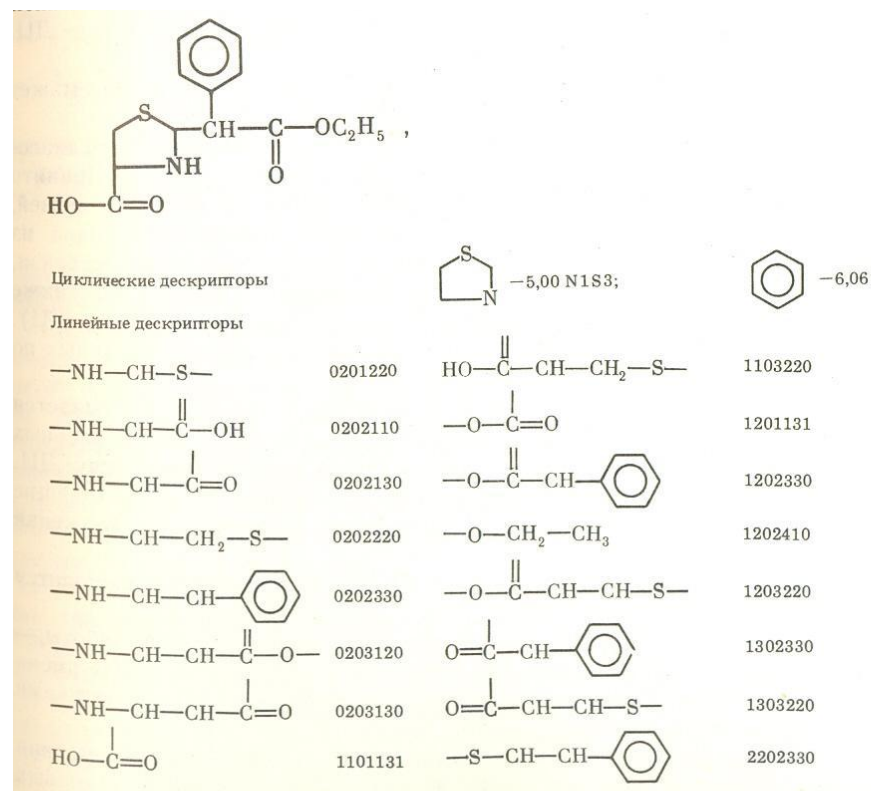
# Линейно-регрессионное QSAR-уравнение

$$\log(A) = \sum x_i \cdot z_i + \mu$$

- **A** - это какой-либо биологический ответ
- **x<sub>i</sub>** - дескриптор, который описывает присутствие или отсутствие *i*-ого заместителя
- **z<sub>i</sub>** - вклад *i*-ого заместителя в  $\log(A)$
- **μ** - константа

# Химическое соединение

- ХС разбивается на **фрагменты** многократно, с суперпозицией отдельных фрагментов



# Достоинства метода QSAR

- При использовании QSAR-метода исследователю **не нужно знать вид зависимости**  $BA = F(S)$
- Число **структурных параметров**  $S$  может быть достаточно большим
- Для QSAR-метода **не требуется молекула-эталон** (выбор ее зачастую субъективен)

# Достоинства метода QSAR

- Достаточно знать ранговую оценку свойства  $BA$  (это важно, когда  $BA$  определяется с достаточно большой *погрешностью*)
- Классификацию молекул по свойствам на основе структурных характеристик можно проводить, ***варьируя число классов***



# Достоинства метода QSAR

- Количественная оценка соответствия между структурой и свойствами позволяет лучше **понять их взаимосвязь**, которая иногда может быть не очень четко выражена в случае большого количества разрозненных данных

# Достоинства метода QSAR

- QSAR позволяет **синтезировать химические соединения** с заданными биологическими свойствами
- Возможно **предсказание свойств** еще не синтезированных химических соединений

# Достоинства метода QSAR

- Результаты могут быть использованы для лучшего ***понимания взаимодействий*** имеющихся между функциональными группами молекул и функциональными группами молекул мишеней

# Недостатки метода QSAR

- Возникновение ***недостовверных корреляций*** из-за чрезмерного доверия к биологическим данным, полученным экспериментальным путем

# Недостатки метода QSAR

- Иногда, эксперимент, от которого зависят результаты QSAR-анализа, может быть *плохо продуман*, из-за чего полученные в результате эксперимента данные могут быть ошибочными

# Недостатки метода QSAR

- Некоторые физико-химические и квантово-химические параметры соединений могут оказаться **взаимокоррелируемыми**, что может привести к неверным выводам о биологической активности исследуемого соединения

# Недостатки метода QSAR

- При применении QSAR наиболее эффективными структурными параметрами  $S$  являются лишь те, которые имеют ***монотонный характер влияния*** на исследуемое свойство ***ВА***

# Дескрипторы химической структуры

- Число или *набор чисел*, которые характеризуют структуру химического (органического) соединения, причем характеризуют так, что они описывают наиболее существенные черты этой структуры



# Дескрипторы химической структуры

- Любое число, которое можно рассчитать исходя из структурной формулы (***молекулярный вес, число определенных атомов, связей или групп, молекулярный объем, частичные заряды*** на атомах и т.д.)

# Дескрипторы

- При создании моделей структура - биологическая активность большое значение в ***отборе дескрипторов*** имеют современные представления о процессах в организме, происходящих с участием химических веществ

# Дескрипторы элементного уровня

- Молекулярная масса
- Число атомов какого-либо сорта
- Атомные массы заместителей
- Молярная рефракция и др.

содержат слишком ***мало информации*** о молекулярной  
структуре

# Топологические дескрипторы

- Рассчитываются на основе описания структурной формулы соединения с помощью **молекулярного графа  $G$** , представляющего собой двумерное отображение молекулы (вершины соответствуют атомам, а ребра - химическим связям молекулы)

# Топологические дескрипторы

- Фрагменты структуры, индексы атомов и связей, *каппа*-индексы, описывающие форму молекулы, индексы молекулярных связей
- ***Индекс Плата***  $F(G)$  равен сумме степеней каждого ребра в графе  $G$

# Топологические дескрипторы

- ***Число Винера*** ( $\Sigma$  кратчайших путей)
- Индекс Балабана
- Теоретико-информационные индексы графов обеспечивают количественную меру симметрии соответствующей структуры

# Дескрипторы электронной структуры

- **Заряды на атомах** (для сильно полярных реагентов) или электронные плотности
- Значения **граничных орбиталей** (ВЗМО и НСМО)
- Значения **вкладов отдельных атомов** в граничные орбитали
- **Дипольные моменты**
- **Поляризуемость**

# Липофильность

характеризует транспортные свойства соединений в биологических объектах

- Распределение ( $P$ ) соединений в модельной системе **октанол-вода**
- $IgP$  для нейтральной формы соединения
- $pK_a$  для ионизированных веществ



# Геометрические параметры молекул

- Длины связей
- Валентные углы
- Торсионные углы
- Размеры молекулы и т.д.

# Методика QSAR

Весь массив соединений с известной характеристикой биологической активности делят на:

- ***обучающий ряд*** (обучающая выборка)
- ***тестируемый ряд*** (контрольная выборка)

# Обучающий ряд

- ***Статистический анализ***

позволяет выявить зависимость и степень зависимости биологической активности от значения того или иного дескриптора или группы дескрипторов

# Обучающий ряд

*Математическая  
зависимость* данной  
биологической активности от  
значений *наиболее значимых  
дескрипторов*

# Тестируемый ряд

*Проверка* полученной  
математической зависимости  
данной биологической  
активности от значений  
выбранных дескрипторов

# Предсказанные и опытные значения $LC_{50}$

# QSAR

***Справочников*** по QSAR-  
уравнениям пока нет

Обзоры в ***Chemical Rev.***

# Примеры уравнений QSAR



# Примеры уравнений

- Производные 1,2,3-тиадиазола – фунгицидная активность (AF)  
*Verticilium dahliae* на хлопке

$$\frac{1000}{100 - AF} = -130,63 - 8,77AMW + 352,76Mv - 93,96Mp + 12,16nCl + 1,33PCWTe - 0,06RBN + 269,86RNCG, n = 26, R^2 = 0,991, Q^2 = 0,917$$

# Примеры уравнений

- средний молекулярный вес ***AMW***
- значение атомного Вандерваальсового радиуса для атомов углерода ***Mv***
- значение атомной поляризуемости для атомов углерода ***Mr***
- количество атомов хлора ***nCl***
- топологический электронный дескриптор частично тяжелого заряда ***PCWTe***
- количество вращающихся связей ***RBN***
- относительный отрицательный заряд ***RNCG***

# Примеры уравнений

- Производные 1,2,3-триазола – фунгицидная активность (AF)  
*Verticilium dahliae* на хлопке

$$\frac{1000}{100 - AF} = -22,79 + 4,05RBN - 62,12RBF - 0,69nO - 1,25ARR + 0,78D - 19,73E_{HSMO},$$

$n = 18, R^2 = 0,963, Q^2 = 0,910$

- Производные 1,2,3-триазола – фунгицидная активность *Gibberella zeae* на пшенице (G)

$$\frac{1000}{100 - G} = -168,42 - 2,11AMW - 3,98Mv + 53,64Mp + 160,25Me + 0,33nH + 0,05D - 0,02Ss, n = 18, R^2 = 0,960, Q^2 = 0,883$$

# Мутагенная/канцерогенная активность

- Определение множества дескрипторов для каждой молекулы, их корреляций с мутагенной/канцерогенной активностью, отбор наиболее значимых дескрипторов, вносящих наибольший вклад в биологическую активность, требуют использования ***специальных компьютерных программ***

# Предсказание генотоксичности

- Пакеты программ **CASE, MULTICASE, МЕТА**
- *Rosenkranz H.S., Klopman G. // Toxicol. Ind. Health. 1988. V.4. P. 533-540*
- *Klopman G., Rosenkranz H.S. // Mutat. Res. 1994. V. 305. P.33-46*
- *Klopman G., Ddimayuga M., Talafos J. // J. Chem. Inf. Comput. Sci. 1994. V.34. P. 1320-1325*

# CASE

- ***Обучаемая программа***, начинает свою работу со считывания всех структур в обучающей выборке и их фрагментирования на цепочки по 2-10 связанных неводородных атомов, затем программа для этих фрагментов оценивает вероятность их соотношения с активностью, используя биномиальный и другие статистические методы

# MULTICASE

- *Отбирает* наиболее значимые из этих фрагментов как **биофоры**, считая их ответственными за активность, наблюдаемую у соединений, содержащих эти фрагменты
- Биофор понимается как *участок локализации основных реакций*, приведших к появлению активности

# MULTICASE

- Затем программа *идентифицирует* внутри уменьшенной обучающей выборки молекул, содержащих биофоры, ***фрагменты*** и физико-химические свойства, играющие роль в модуляции активности биофоров



# CASE и MULTICASE –анализ

- Полностью автоматизирован и удобен для пользователя
- Для оценки **новой молекулы** программа сначала проверяет наличие в ней биофоров, если их нет, то молекула предсказывается как неактивная; если молекула содержит биофоры, тогда исследуется наличие модуляторов для этих биофоров – результатом является **качественное** и **количественное предсказание** активности

# **CASE и MULTICASE –анализ**

- Получаемые уравнения **QSAR** действительны только внутри домена, определяемого соответствующими биофорами

# МЕТА – программа

- **Канцерогенность и мутагенность** часто ассоциируется с **электрофильным потенциалом молекулы**, однако, электрофильность молекулы может выявиться или сразу, или после одного или нескольких этапов метаболической трансформации

# МЕТА – программа

- Для решения проблемы метаболической трансформации ксенобиотиков была создана ***экспертная программа МЕТА***

# МЕТА – программа

- Использует предварительно установленные правила в *идентификации структурных мишеней* для некоторых ферментов и преобразует эти мишени, ***имитируя метаболическую трансформацию***

# Метаболиты

- По их потенциальной биологической активности:  
*терапевтически активные,*  
*неактивный экскретируемый,*  
*неактивный неэкскретируемый,*  
*потенциально токсичный* т.д.

# Метаболиты

- Различные метаболиты могут ***локализоваться в различных тканях*** в зависимости от способа введения, коэффициента распределения, способности связываться различными белками и от специфических для соединения механизмов действия

# МЕТА – программа

- Проблема влияния *растворимости* на метаболическую трансформацию решается с помощью программы **LOGP**, входящей в пакет **МЕТА**, и рассчитывающей коэффициенты распределения *октанол-вода*



# МЕТА – программа

- В зависимости от расчетной величины *logP*, молекула подлелжит или не подлелжит *трансформации* соответствующим *ферментом*, также оценивается ее способность проходить через мембрану или двигаться от органа к органу

# Пути биотрансформации

- Большинство данных касается путей биотрансформации различных метаболитов **для животных**, а не для человека
- Первый этап оценки стабильности метаболитов, основывается на **общих для всех видов реакциях**, а специальные задачи могут быть решены с учетом **различий метаболизма** между видами

# Пути биотрансформации

- В отдельных случаях могут рассматриваться генетически обусловленные ***индивидуальные вариации метаболизма***, такие как *полиморфизм*, связанный с *N-ацетилированием* и *P-450-опосредованным окислением*

# МЕТА – программа

- После каждого этапа трансформации программа с помощью словаря спонтанных реакций **оценивает стабильность** полученного метаболита и продолжает **трансформировать структуру** до образования конечного продукта, нестабильные метаболиты автоматически разрушаются

# МЕТА – программа

- Метаболиты, оцениваемые как потенциально мутагенные/канцерогенные, ***вносятся в базу данных CASE или MULTICASE***

# Пример использования QSAR-анализа

- Проведено сравнительное изучение **мутагенной активности** производных **бифенила, флуоренона, фенантренхинона, пирена** и его **гетероциклических аналогов** в тесте Эймса на штамме *Salmonella typhimurium* TA 1538 (всего **60 соединений**)

# Пример использования

- Большинство из исследованных соединений показало ***мутагенную активность*** и некоторые из них индуцируют до ***7000 ревертантов/нмоль***

# Пример использования

- Любимова И.К., Абилев С.К., Мигачев Г.И. // *Генетика*, 1995. N 2. С. 268-272
- Баскин И.И., Любимова И.К., Абилев С.К., Палюлин В. А., Зефиоров Н.С. // *Доклады АН*. 1993. Т. 332. N 5. С. 587-589
- Баскин И.И., Любимова И.К., Абилев С.К., Палюлин В. А., Зефиоров Н.С. // *Доклады АН*. 1994. Т. 339. N 1. С. 106-108
- Любимова И.К. *Зависимость мутагенной активности полициклических ароматических соединений от их структуры. Автореферат канд. диссертации. М. 1994. 25 с.*



# Анализ данных

- **Пара-положение** нитро- и amino-групп определяет **высокую** мутагенную активность, наличие же заместителей (нитро-, ацетил-, карбонил- ) в других положениях **ослабляет** или **усиливает** эту активность

# Уравнения QSAR

- Для получения уравнений QSAR на выборке исследованных соединений были выбраны следующие дескрипторы:
- $d_1$  – энергия нижней незанятой молекулярной орбитали (LUMO);
- $d_2$  – величина максимального заряда на атоме азота;
- $d_3$  – величина максимального заряда на атоме кислорода;
- $d_4$  –  $\log P$  (липофильность);
- $d_5$  – наличие нитрогруппы в *пара*-положении;
- $d_6$  – наличие аминогруппы в *пара*-положении;
- $d_7$  – наличие заместителей в *мета*- и *орто*-положениях

# Квантово-химические параметры

- Рассчитывали по *методу Хюккеля*, анализ *корреляций между дескрипторами и мутагенной активностью* проводили с помощью *программы “EMMA”* (Effective Modeling of Molecular Activity), которая позволяет осуществить множественный линейный регрессионный анализ

# Уравнения QSAR

- Соединения были разделены на **2 группы**:  
первая включала **гетероциклические аналоги пирена**, производные **фенантрена** и **флуоренона**;  
вторая – производные **бифенила**

# Уравнения QSAR

- Была получена серия регрессионных уравнений со следующими параметрами:
- $N_{his^+}$  - число  $his^+$  ревертантов при дозе 7.5 нмоль/чашку;
- $R$  – коэффициент корреляции между экспериментальными и прогнозируемыми значениями логарифма числа  $his^+$  ревертантов;
- $S_v$  – среднеквадратичная ошибка прогноза логарифма  $his^+$  ревертантов в контрольной выборке;
- $F$  - значение критерия Фишера

# Уравнения QSAR

- Для гетероциклических аналогов *пирена* и производных *фенантрена* и *флуоренона* (всего **39 соединений**)

$$\ln(N_{\text{his}}^+) = 0.88 d_1 + 2.08 d_5 + 1.07$$
$$R = 0.72; S_v = 1.31; F = 14.72$$

$$\ln(N_{\text{his}}^+) = -1.05 d_1 + 1.84 d_5 - 0.34 d_7 + 1.04$$
$$R = 0.75; S_v = 1.24; F = 11.25$$

$$\ln(N_{\text{his}}^+) = -2.64 d_1 - 1.10 d_2 - 8.93 d_3 -$$
$$1.34 d_5 + 0.89 d_6 - 3.93$$
$$R = 0.80; S_v = 1.21; F = 6.68$$

# Уравнения QSAR

- Для замещенных *бифенилов* (21 соединение)

$$\ln(N_{\text{his}}^+) = 2.1 d_1 - 0.57 d_4 - 0.13$$
$$R = 0.82; S_v = 1.12; F = 16.37$$

- Для всех соединений (объединенной группы соединений)

$$\ln(N_{\text{his}}^+) = 1.77 d_1 - 0.22 d_2 - 0.9 d_5 - 0.52$$
$$R = 0.75; S_v = 1.94; F = 18.80$$

# Уравнения QSAR

- Все эти соединения ***удовлетворительно прогнозировали мутагенную активность*** контрольных выборок соединений (3 производных пирена и фенантрена и 2 производных бифенила)



# Уравнения QSAR

- ***Во всех уравнениях*** присутствует ***дескриптор  $d1$  – энергия LUMO***, которая характеризует способность соединения реагировать с нуклеофилами, т.е. его ***электрофильность***

# Анализ уравнений QSAR

- Характеризующих взаимосвязь *структура-активность* в группе разнородных соединений, включающей гетероциклические аналоги пирена, производные фенантрена и флуоренона показывает, что величина ***R*** ***зависит от числа*** включенных в уравнение ***дескрипторов***

# Анализ уравнений QSAR

- $R = 0.80$  достигается только при включении в QSAR-модель таких дескрипторов как ***d2*** и ***d3***, характеризующих величину максимального ***заряда на атомах азота и кислорода***, соответственно

# Анализ уравнений QSAR

- Наличие в этом уравнении дескрипторов ***d5*** и ***d6***, характеризующих наличие нитро- и amino- группы в *пара*-положении, соответственно, ***отражает эмпирические выводы***, сделанные на основе анализа экспериментальных данных

# Анализ уравнений QSAR

- Дальнейшее *повышение коэффициента корреляции* между экспериментальным и прогнозируемым значением мутагенной активности (логарифм числа his+ ревертантов) достигается только при *увеличении числа* квантово-химических и подструктурных *дескрипторов*

# Анализ уравнений QSAR

- Увеличение числа дескрипторов в уравнениях QSAR, **сужает круг** прогнозируемых структур
- Такое уравнение будет прогнозировать активность в реальной ситуации **не лучше**, чем уравнения с дескрипторами, характеризующими наиболее общие свойства обучающей выборки

# Анализ уравнений QSAR

- Прогностическая ценность уравнений QSAR зависит от ряда общих проблем, касающихся ***принципов подготовки материала*** для анализа методами SAR и QSAR

# Анализ уравнений QSAR

- Подходы к изучению взаимосвязи структура-активность основаны **на существующих знаниях** о физико-химических **свойствах** исследуемых соединений и о **механизмах** их биологической активности
- Все это зависит от современного уровня знаний, точек зрения, предположений, которые могут быть **верными** и **неверными**



# Проблемы использования метода QSAR

- Размер и неоднородность базы данных
- От размера и однородности обучающей выборки зависит ***прогностическая способность*** уравнений QSAR

# Проблемы использования метода QSAR

- ***Большая и однородная*** база данных, полученная на группе родственных соединений, позволяет получить уравнение с ***высокими предсказательными возможностями*** и с ***ограниченным числом дескрипторов***, описывающих главные свойства химических соединений, с которыми связано наличие или отсутствие биологической активности

# Проблемы использования метода QSAR

- **Неоднородная база данных** и, соответственно, такая же обучающая выборка, **не позволяет** получить уравнения с высокими предсказательными возможностями
- Увеличения значения  $R$  для таких выборок можно достичь только за счет **увеличения ее размера** или **числа дескрипторов**

# Проблемы использования метода QSAR

- *Проблема выделения дескрипторов*
- Можно использовать **всевозможные** квантовые, физико-химические свойства и структурные **характеристики**, а затем с помощью **статистических процедур** отбирать те из них, которые лучше всех **коррелируют с активностью**

# Проблемы использования метода QSAR

- Такой подход нуждается в проведении ***большой предварительной работы***, (расчета квантово-химических параметров для каждого соединения, включенного в базу данных, с помощью специальных программ)

# Проблемы использования метода QSAR

- Поэтому *при анализе* структура-активность химических мутагенов **особое внимание** уделяется **механизмам мутагенного действия** химических соединений (алкилирование, интеркаляция и т.д.)

# Проблемы использования метода QSAR

- Известно, что **мутагенная активность** отражает способность молекул к **реакции** с азотистыми основаниями ДНК
- В таком случае **общим свойством** соединений с мутагенной активностью является наличие в их молекуле **электрофильных свойств** или эти сайты образуются в результате **метаболизма**

# Проблемы использования метода QSAR

- Для выделения таких “структурных алертов” необходимы **глубокие знания химии и метаболизма соединений**, чего нет на практике
- **Электрофильная теория** не может быть признана **единственно верной**



# Проблемы использования метода QSAR

- **Анализ баз данных** по мутагенной активности **нитро-** и **аминоароматических** соединений
- Чтобы **отличить мутагенные соединения** от немутагенных при наличии у всех структурного алерта – нитро- или амино- группы, необходимо обязательно учитывать наличие **дескрипторов-модуляторов** мутагенной активности

# Проблемы использования метода QSAR

- Если *энергию LUMO*, отражающую *электрофильность* молекулы и *logP* - показатель ее *биологической доступности*, принять за основные дескрипторы, то *другие дескрипторы* являются *модуляторами* мутагенной активности

# Уравнение QSAR

- Уравнение для *бифенильных соединений*:

$$\ln(N_{\text{his}}^+) = - 111 d1 - 2.66 d2 - 36.7 d3 - 10.4 d4 + 0.967$$

$$R = 0.95; Sv = 0.69; F = 35$$

# Уравнение QSAR

- где  $d_1$  – минимальный квадрат коэффициента вклада атомной орбитали углерода в низшую свободную молекулярную орбиталь
- $d_2$  – минимальный квадрат коэффициента вклада атомной орбитали азота в низшую свободную молекулярную орбиталь
- $d_3$  – максимальный индекс свободной валентности на атоме углерода
- $d_4$  – среднее значение индекса свободной валентности на атомах кислорода

# Уравнение QSAR

- Уравнение отличается от прежнего уравнения более **высоким уровнем** коэффициента корреляции  **$R=0.95$**
- Но, оно не отражает возможного **механизма мутагенного действия** соединения и не имеет очевидного биологического смысла, что затрудняет его восприятие с точки зрения **биолога-экспериментатора**

# Уравнение QSAR

- Возникает дилемма
- Выбирать дескрипторы *исходя из механизмов* мутагенного действия и результатов анализа экспериментальных данных
- *Перебирать всевозможные* квантовохимические параметры

# Уравнение QSAR

- Реальная предсказательная *ценность уравнений* второго типа пока остается неизвестной

# QSAR – анализ

- При существующем **разнообразии программ QSAR и баз данных** по мутагенной активности химических соединений в различных тест-системах, достаточно **корректное уравнение QSAR** можно получить только при **совместной работе** биологов, химиков, математиков



# QSAR – анализ

- **Определение цели** – получить корреляции между активностью и структурными особенностями молекулы (**качественный анализ – SAR**) или получить коэффициент корреляции между экспериментальными и прогнозируемыми значениями мутагенной активности (**количественный анализ – QSAR**)

# QSAR – анализ

- В последнем случае количественные **экспериментальные данные** анализируются и **формализуются**
- Решается **вопрос об однородности** базы данных
- Если принимается обоснованное решение о разделении базы данных на подгруппы, то дальнейший анализ проводится **для каждой подгруппы** по отдельности

# QSAR – анализ

- **Выбор дескрипторов** зависит от **общих знаний механизма** мутагенного действия, от **возможности программ**, которые предполагается использовать для расчета квантовохимических параметров и анализа фрагментов структур, а также для регрессионного анализа связи между рассматриваемыми дескрипторами и мутагенной активностью

# QSAR – анализ

- В конечном итоге ***выбираются*** те ***уравнения QSAR***, которые достаточно удовлетворительно ***прогнозируют*** активность соединений, отнесенных к ***контрольной выборке***

# QSAR – анализ

- Речи сегодня о замене *теста Эймса* или другого теста методологией **QSAR** нет
- Это **возможно** в будущем
- Сегодня **QSAR** является хорошим инструментом для **внеэкспериментального прогноза**

# Задание

- Используя известное QSAR уравнение осуществите теоретическое конструирование новых ХС с данным видом БА

# Задание

- Найдите корреляционные уравнения, связывающие  $LD_{50}$  химических веществ с параметрами их структуры

# Задание

- Обработайте массив экспериментальных данных зависимости БА ~ параметры ХС
- Создайте систему данных, аналогичную системе элементов Д.И. Менделеева



**PASS**

Prediction of Activity  
Spectra for Substances

# Биологическая активность

- Обнаружение у веществ определенных свойств – по-прежнему **случайное событие**, но **вероятность проявления** такого события **резко возрастает**

# Пример

- *Аминазин (нейролептик)* – использовался для непсихотропных целей, прежде чем были открыты его свойства нейролептика

# Примеры

- ***ВИАГРА*** - *силденафил* - как антиагрегантное средство начал применяться в 1988 г. и как препарат, стимулирующий эрекцию - в 1994 г.
- ***Аспирин*** был предложен в качестве анальгетика в 1899 г., а его антиагрегантное действие было открыто лишь в 1971 г.; и т.д.

# Пример

- *Талидомид*, обладающий анксиолитическим и снотворным эффектами, был введен в медицинскую практику в 50-х годах
- В начале 60-х годов из-за наличия тератогенности он стал причиной врожденных дефектов у более чем 8000 новорожденных, что привело к запрету на его применение

# Пример

- Сорок лет спустя, *талидомид* переживает «второе рождение» - он активно испытывается в клинике как потенциальное противоопухолевое и антиметастатическое средство, как препарат для симптоматической терапии СПИДа

# Компьютерная система прогноза

- Собрать всю известную информацию о биологически активных соединениях, создать на этой основе обучающую выборку, провести анализ связей "структура-активность" для веществ из обучающей выборки и построить соответствующие зависимости
- "Подставив" в эти зависимости данные о структуре нового вещества, можно получить в результате оценку его биологической активности

# КССА

- Применимы к ***соединениям одного и того же химического класса***
- Оперируют с ***одним видом биологической активности***
- Можно ли разработать подобные методы для веществ, гетерогенных как по химической структуре, так и по проявляемому ими биологическому действию?



# PASS

- **Основа** – характеристика активных частей молекулы
- Расчет **частоты субструктурной активности** (ЧСА) – отношение БАС, содержащих данную структуру, к числу соединений, имеющих эту структуру

# PASS

- Вычисляют из данных ЧСА *среднюю арифметическую ЧСА* для всей молекулы
- *ЧСА* – предсказывает БАВ

# PASS

- Выделение *фрагментов фармакофоров*
- *Определение частоты встречаемости* этих фрагментов для определения видов БА

# Дескрипторы

- **Топологические** – не меняющиеся при деформациях молекулы (брутто-формула)
- **Топографические** – пространственные соотношения между компонентами молекулы (расстояния, углы)
- **Квантовомеханические** или **квантовохимические** – электронная структура ( $E_{MO}$ ,  $q$ , порядок связи, индекс свободной валентности)

# Дескрипторы

- **Химические** – изменение состава, числа связей ( $\Delta G$ ,  $pK$ )
- **Физические** ( $V_{\text{молекулы}}$ ,  $T_{\text{пл}}$ ,  $T_{\text{кип}}$ ,  $n_D^{20}$ )
- **Физико-химические** – взаимодействие с другими веществами без химической реакции (растворимость, хроматографические индексы)

# PASS

- ***Мера информативности дескриптора*** – отношение частоты встречаемости в подмассиве с данной активностью к частоте встречаемости во всем массиве

$$n_i/N_i : n_j/N$$

# PASS

- Версия компьютерной системы предсказания спектра биологической активности PASS реализована в **1998 году** Институтом биомедицинской химии (Москва)

*Рос. хим. ж. (Ж. Рос. хим. об-ва им. Д.И. Менделеева), 2006, т. L, № 2*

**Прогноз спектра биологической активности органических соединений**

Д. А. Филимонов, В. В. Поройков

# PASS

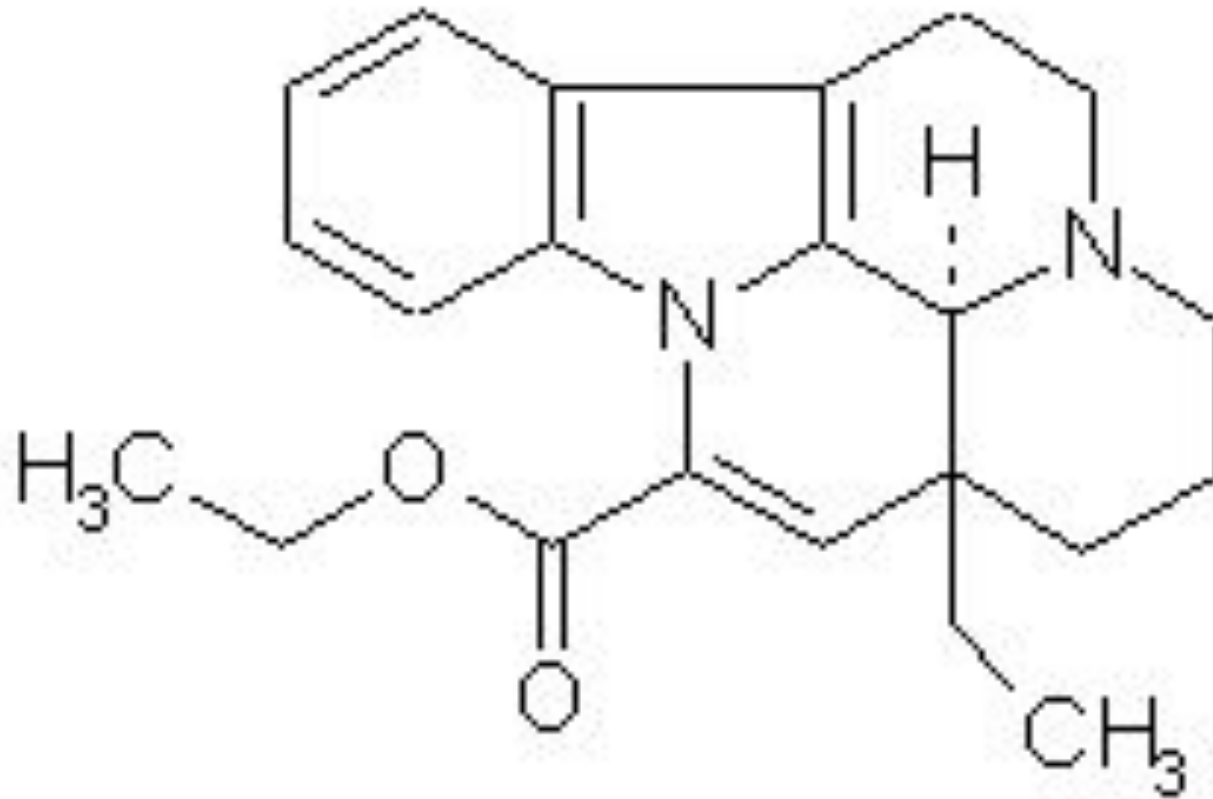
- Она включала в себя обучающую выборку, содержащую более **30000** БАВ с известной биологической активностью, и охватывала более **400** (900) фармакологических эффектов, механизмов действия, а также мутагенность, канцерогенность, тератогенность и эмбриотоксичность



**PASS online**

**[www.pharmaexpert.ru/  
PASSOnline/](http://www.pharmaexpert.ru/PASSOnline/)**

# Cavinton



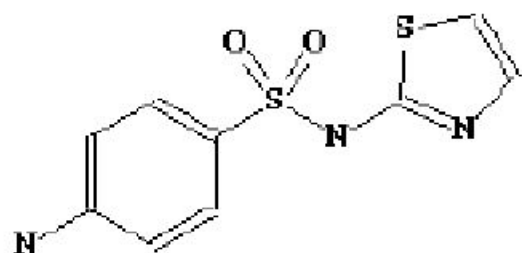
# PASS online

- Биологическая активность описывается в PASS качественным образом ("да"/"нет")
- Выдаваемые результаты прогноза помимо **названий активности** включают в себя **оценки вероятностей наличия** ( $P_a$ ) и **отсутствия** каждой активности ( $P_i$ ), имеющие значения от 0 до 1
- Эти вероятности рассчитываются независимо, их **сумма не равна единице**

# Cavinton

Pa	Pi	Activity
0.929	0.004	Peripheral vasodilator
0.900	0.000	Multiple sclerosis treatment

No Selected Activity



## Activity Spectrum

Chart   General   **Effects**   Mechanisms   Toxicity

## Known Activities:

Antibacterial  
 Dihydropteroate synthase inhibitor  
 Iodide peroxidase inhibitor

139 of 2005 Possible Activities at Pa &gt; Pi

0.889 0.005 Antiobesity  
 0.835 0.005 Para amino benzoic acid antagonist  
 0.736 0.006 Dihydropteroate synthase inhibitor  
 0.721 0.006 Antidiabetic  
 0.556 0.006 Antiprotozoal (Coccidal)  
 0.551 0.019 Prostaglandin E1 antagonist  
 0.509 0.026 Prostaglandin H2 antagonist  
 0.485 0.045 Potassium channel antagonist  
 0.453 0.013 Cyclooxygenase inhibitor  
 0.468 0.028 Antiprotozoal  
 0.443 0.012 Antibacterial  
 0.412 0.021 Diuretic inhibitor  
 0.408 0.024 Gingipain R inhibitor  
 0.421 0.053 Antinfective  
 0.371 0.006 Hypoglycemic  
 0.328 0.015 Antineoplastic (breast cancer)  
 0.362 0.054 Antimycobacterial  
 0.351 0.047 Antituberculosis  
 0.325 0.023 Saluretic  
 0.345 0.052 Myelodysplastic syndrome treatment

> <CAS>  
 72-14-0

> <GENERIC\_NAME>  
 SULFATHIAZOLE

32 Substructure Descriptors: 0 new.  
 There are 3 known activities.  
 Drug-Likeness: 0.156

139 of 2005 Possible Activities  
 35 of 224 Possible Pharmacological Effects  
 95 of 1756 Possible Molecular Mechanisms  
 9 of 25 Possible Side Effects and Toxicity

# Задание

- Рассчитать спектр предсказания биологической активности для *Талидомида*

# PASS

- Если целью исследования является **поиск базовых структур лекарств**, целесообразно отбирать из массива доступных веществ не те структуры, для которых величина  **$P_a$**  близка к единице (они могут оказаться близкими аналогами известных препаратов), а соединения с  **$P_a < 0.7$**

# PASS

- ***Риск*** получения отрицательного результата в эксперименте тем больше, чем меньше величина  ***$P_a$*** , однако и ***новизна*** такой структуры (при подтверждении прогноза в эксперименте) будет более высокой



# PASS

- Если поставлена цель ***поиска*** ***близкого аналога*** известного препарата, то из массива имеющихся образцов следует отобрать вещества с наибольшими значениями ***Pa***

# PASS

- При отборе **перспективных** для исследований соединений можно руководствоваться **комбинированным критерием**:
  - - наличие в прогнозируемом спектре требуемых эффектов/механизмов
  - - отсутствие нежелательных эффектов/механизмов

# PASS

- Может быть выполнен уже *на этапе планирования синтеза*
- Будут синтезированы лишь некоторые из теоретически возможных производных, в наибольшей степени удовлетворяющие критериям задачи

# Компьютерная система интерпретации спектров биологической активности

- Можно составить ***большое количество комбинаций*** из требуемых и нежелательных эффектов

# Компьютерная система интерпретации спектров биологической активности

- Для их анализа сотрудник Лаборатории структурно-функционального конструирования лекарств НИИ Биомедхимии РАМН А.А. Лагунин разработал специальную компьютерную систему интерпретации спектров биологической активности веществ **IBIAC**, основанную на знаниях об известных взаимосвязях между фармакологическими эффектами и механизмом действия биологически активных веществ (более 2000 терминов, описывающих биологическую активность)

# Компьютерная система интерпретации спектров биологической активности

- С использованием системы ***IBIAC*** генерация перечня эффектов, соответствующих определенному механизму действия и, наоборот, списка вероятных механизмов, ответственных за проявление определенного эффекта, осуществляется автоматически

# Ограничения PASS

- Имеет смысл прогноз для ***синтетических и биополимеров, для неорганических веществ и т. п.***

# Ограничения PASS

- Наличие **не менее 5 веществ** с известной активностью для формирования обучающей выборки
- В случае принципиально **НОВЫХ мишеней** действия лекарственных препаратов, для которых имеются данные только об 1-2 лигандах, предсказание биологической активности таким методом не может быть реализовано



# Ограничения PASS

- В случае существенной по отношению к соединениям обучающей выборки **новизны химической структуры** прогнозируемого вещества (более 3-х дескрипторов ни разу не встретились в обучающей выборке) результаты прогноза могут иметь **значительную погрешность**

# Ограничения PASS

- *Целесообразно протестировать* вещество на требуемые виды активности независимо от результатов прогноза, так как результатом может оказаться принципиально новая базовая структура

# Ограничения PASS

- В некоторых случаях вещество прогнозируется одновременно как **агонист и антагонист** (стимулятор и блокатор, активатор и ингибитор) по отношению к одним и тем же рецепторам (ферментам и т.п.)
- Это означает, что система не может **дифференцировать внутреннюю активность вещества**, а лишь указывает на его способность к связыванию с данным рецептором (ферментом)

# Система PASS

- ***Не может предсказать***, станет ли конкретное вещество лекарственным препаратом, поскольку это будет зависеть также от многих других факторов (сравнительной оценки безопасности и клинической эффективности; наличия необходимых для разработки и внедрения инвестиций, и т.д.)

# Система PASS

- Прогноз, однако, может помочь определить, **какие тесты** наиболее адекватны для изучения биологической активности конкретного химического вещества, и **какие вещества** из имеющихся в распоряжении исследователя наиболее вероятно проявят требуемые эффекты

# Задание

- Провести анализ прогноза видов БА для известного лекарственного препарата  
(*Сульфадимезин, Норсульфазол и др.*)

# Методы компьютерного моделирования

- Позволяют ученым строить компьютерные модели «мишени» и определять структуру потенциального «ключа»
- Химики синтезируют это соединение

# Лекарство от болезни Альцгеймера

- С помощью компьютерного моделирования химики МГУ (Зефиоров Н.С.) выяснили, что **димебон** – лекарство от аллергии – может и должен обладать нейропротекторными свойствами



# Лекарство от болезни Альцгеймера

- *Димебон* недавно прошел клинические испытания в России
- Российский препарат даже пропустили на американский лекарственный рынок!

**Зачет?**