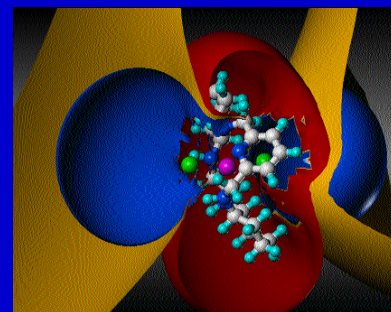




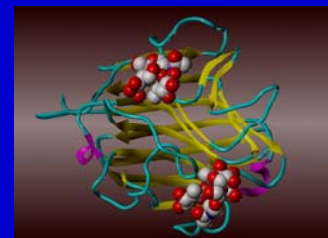
# 环境有机污染物的 结构-性质与活性





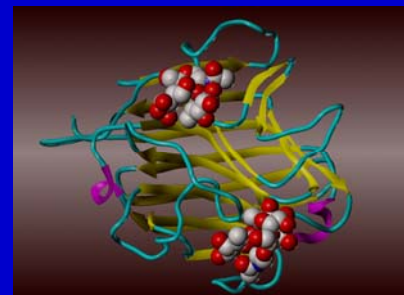
# Outlines

- 结构-性质/活性关系概论
- 结构-性质/活性关系模型的构建
- 有机污染物的分子结构表征
- 结构-活性关系研究的主要方法
- 结构-活性关系研究发展的主要方向



# 第一节 结构-性质/活性关系概论

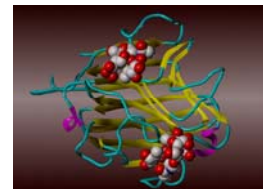
- 结构-性质/活性关系的含义
- 结构-性质/活性关系的工作原理
- 结构-性质/活性关系的功能
- 结构-性质/活性关系的发展历程
- 结构-性质/活性关系的发展特点





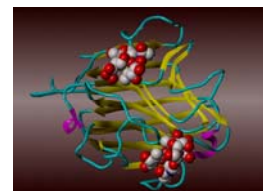
## （1）结构-性质/活性的含义

- 结构-性质/活性关系（**Structure—Activity/Properties Relationship**），是指对化学品的分子结构与性质或活性之间关系的研究，可以分为定性结构-性质/活性关系（**SAR**）和定量结构-性质/活性关系，即**QSAR**。
- 环境科学中的结构-活性关系研究是有机污染化学和生态毒理学中的一个前沿领域，是化学品环境风险评估的重要组成部分和值得信赖的重要手段之一。

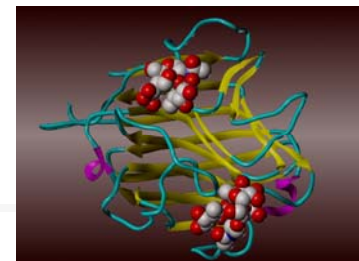


# 结构-性质-活性的含义

- **结构**：有机物的分子结构，通常用参数化的形式定量地表达。分子结构参数，分子结构描述符等。
- **性质**：包括污染物的理化性质例如污染物的水溶解度、疏水性、挥发性、熔点、沸点、极性等理化性质，以及表征污染物在不同介质之间的吸附、分配、迁移以及降解、水解、光解等环境行为的特性。因此可以概括为有机污染物环境行为或某种性质。
- **性质的双重角色**

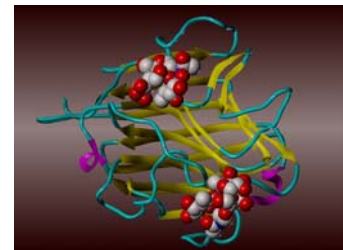


# 结构-性质-活性的含义



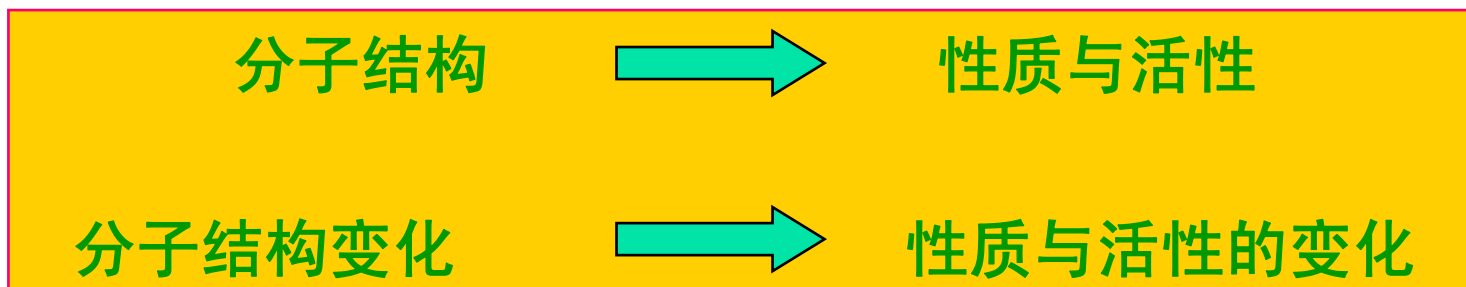
## ■生物活性

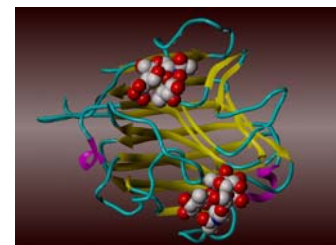
指污染物对不同生物物种和不同层次的测试终点的生态毒性和健康效应指标。例如污染物对微生物和水生生物的急性毒性、亚急性毒性、生长抑制毒性、酶抑制毒性等，对高等生物和人类的急性毒性、三致毒性（致癌性、致畸性和致突变性）、内分泌干扰活性（生殖和遗传毒性、免疫毒性、生长发育毒性、内分泌干扰、神经毒性等），对生物组织水平、细胞水平乃至分子水平的损伤等。



## (2) QSAR的工作原理

- 分子是构成物质的基础单位，化合物内部分子结构特征及分子间的组合方式等结构信息决定了化合物所表现的性质。
- 将化合物的生物活性归因于化学结构的影响是QSAR研究的理论基础。

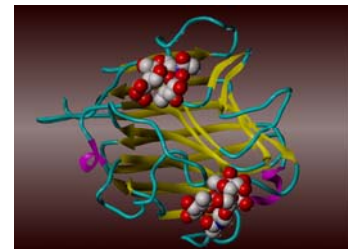




$$Y = f(x)$$
$$\Delta Y = f(\Delta x)$$

- 通过研究化合物的分子结构与其性质、活性之间的定量关系，应用数学方法，建立起结构-性质/活性关系模型，从而对已进入环境的污染物及尚未投入使用的新化合物的相关性质/活性进行预测、评价和筛选。这是结构活-性关系技术的工作原理。





### (3) 结构-活性关系的功能

- 利用模型的预测能力对其它未知化合物的相关性质/活性进行预测，在效应评价和暴露评价中弥补缺失的数据，对环境有机化学品进行筛选和评价；
- 根据模型的组成与形式，确定影响性质/活性的主要结构因素，结合已有的化学、生物学知识，探求有机污染物的环境过程机制和生态效应机理。
- 根据所阐明的结构-活性关系结果，为设计高效、低毒无害的新型环境友好化学品指明方向。



## (4) 结构-活性关系的发展历程

- 早在19世纪，就有人开始设法建立化合物的生物活性和结构之间的关系，这些研究者认为可以根据某些通用的规则，从化合物的化学结构探测活性。
- 20世纪初，人们普遍将化合物的生物效应与其物理性质如溶解度、表面张力和分配系数等联系起来，这些都属于结构-活性关系的范畴。
- 结构-活性关系的研究最初作为**定量药物设计**的一个研究分支领域，是为了适应合理设计生物活性分子的需要而发展起来的，它对于设计和筛选生物活性显著的药物，以及阐明药物的作用机理等均具有指导作用。



- 开创性的工作始于本世纪30年代Hammett等人的研究.
- Taft等人在50年代, Hansch等人在60年代, 为有机物定量结构—性质—活性相关研究做出了重要的贡献
- Hansch等借助于计算机技术建立的结构—活性关系表达式, 标志着QSAR时代的开始。
- 此后, QSAR在理论上日益得到发展, 同时在各个领域的应用范围日益广泛, 国内外在这方面的研究很活跃;



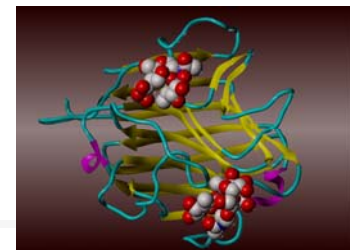
- 70年代以来，由于对进入环境中的大量有机化学品的生态风险评价的需要，QSAR开始在环境科学中得到应用，并得到了迅猛的发展；
- 1993年，专业杂志**SAR and QSAR in Environmental Research**在法国创刊.
- 并且世界各地已经开始有众多的环境科学家从事有机污染物的QSAR及QSPR的研究，每年有大量的科研论文发表。



## (5)QSAR的发展特点

---

- **综合性：**化学(环境化学、化学信息学、量子化学、图论、物理化学等),生物学(细胞生物学、分子生物学、毒理学、生态学等),数学(统计学、应用数学、数学模拟等),计算机科学(计算机图形学、分子模拟、人工智能等).
- **理论性:**体现在QSAR模型的建立不仅能够起到高效的预测作用,还能够根据模型的形式和组成,从参数的物理意义,提供理论上和机理上的信息,从而为探索有机污染物的环境行为规律和生态效应机理提供指导.

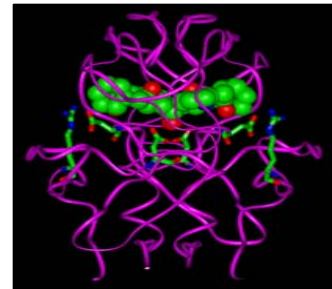


- **微观性:**90年代以后，随着量子化学、三维分子模拟和计算机图形学等在QSAR中的应用，随着效应指标体系的日益精确，暴露评价方法的日益完善，人们更注重定量模型的理论性和微观性，希望能够从本质上、从微观层次揭示和描述化学品的生物活性机制
- **智能化:** 由于化合物的结构的多样性以及生物效应作用过程和机理的复杂性，要在化合物的结构因素与生物活性之间建立满意的运算关系必须借助于多变量分析方法和计算机的自适应功能。因此多变量统计方法、人工智能、遗传算法等善于处理复杂问题等高级建模方法越来越多地应用于结构-活性关系研究。促进QSAR研究向智能化发展。



- **程序化与系统化：**结构-活性关系模型的建立，往往基于对大量化合物的众多参数的分析，从中筛选出对生物活性具有显著影响的变量，开发了大型的集成化的软件、专家系统。例如专家系统 **MULTICASE** 系统能够自动采集结构描述符、自动选择参数、自动建立QSAR模型并进行预测、自动提出对有机污染物的生态风险管理意见等。
- **实用性：**由于QSAR可以对化学品的暴露水平和生态效应作出评价，为化学品危险性评价提供了一种快速、简便、实用的途径和筛选工具。各国政府越来越重视QSAR的评价和预测功能，美国EPA在化学品评价中承认并接受QSAR的预测数据和结果。其次，通过QSAR可以发现并确定对化合物活性起关键作用的结构因素，因而对定向合成高效低毒的新化合物具有指导意义。

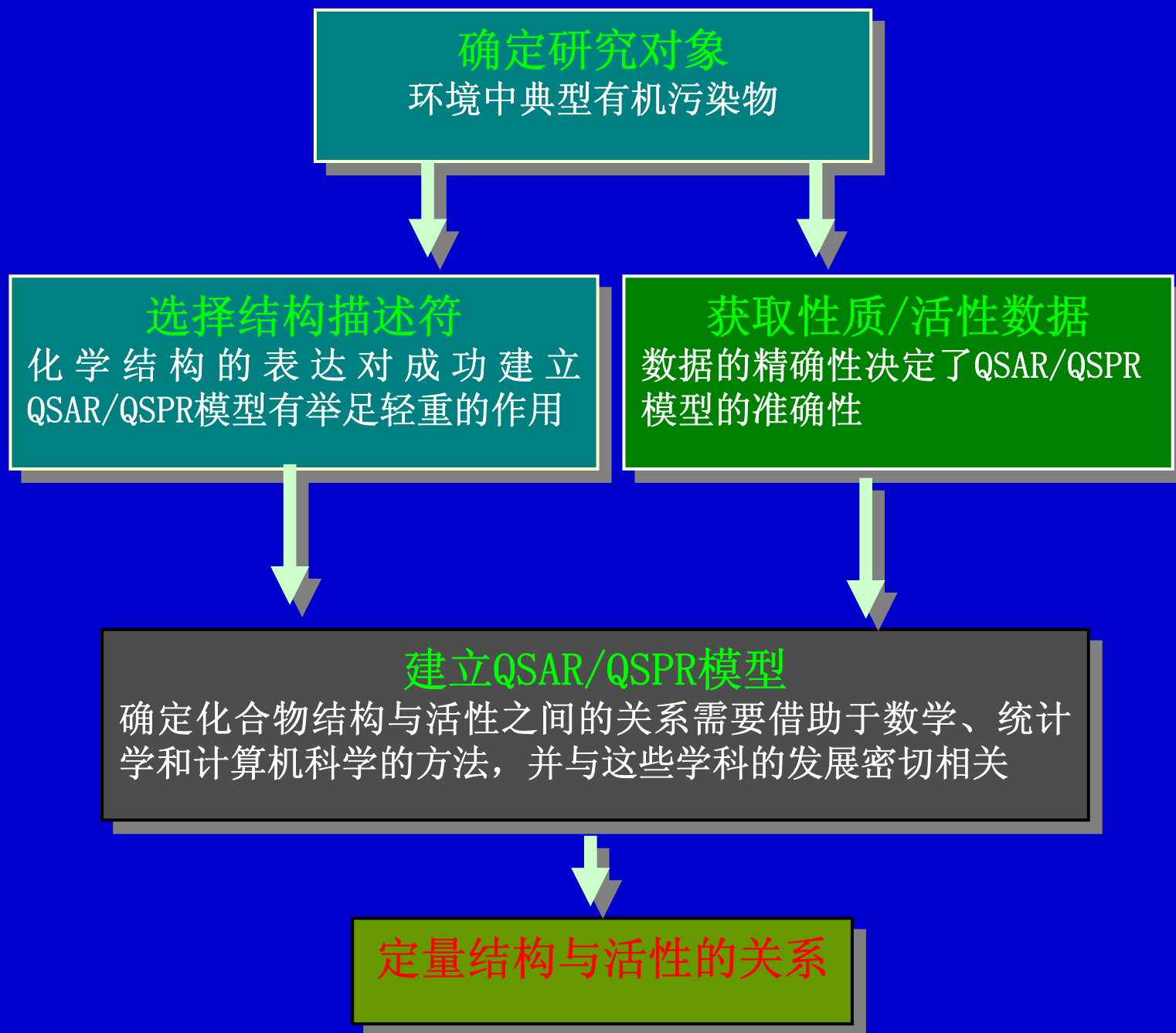
## 第二节 QSAR模型的构建



- 测试集系列化合物的选择
- 性质/活性数据的获得
- 分子结构描述
- QSAR模型构建
- QSAR模型的评价、验证
- QSAR模型的应用



# 结构活性关系研究总体思路





有机污染物

性质/活性

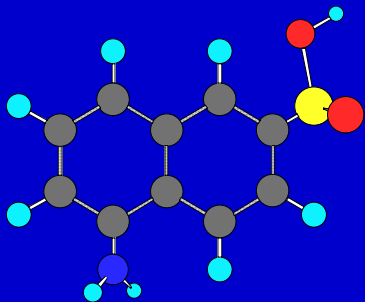
分子结构描述

统计分析方法

QSAR模型

模型验证与修正

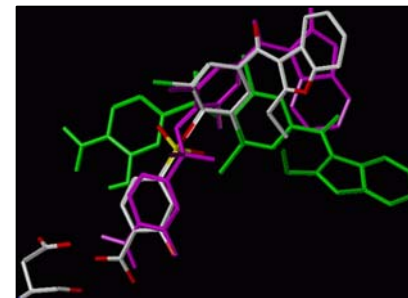
模型应用





## (1) 确定研究对象

---



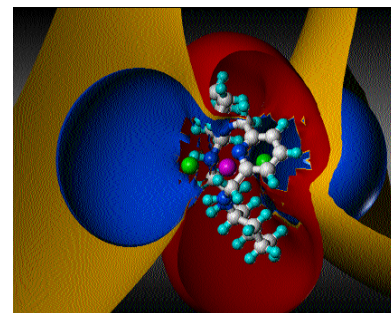
- 确定所要研究的对象，即环境中典型的有机污染物
- 根据一定的统计标准和结构标准选择一系列化合物，构成QSAR模型建立的训练集。
- 测试集化合物选择的条件是统计上的随机性和代表性，结构上的代表性、多样性和全面性，以及性质/活性数据的可获得性。



## (2) 性质/活性数据的获得

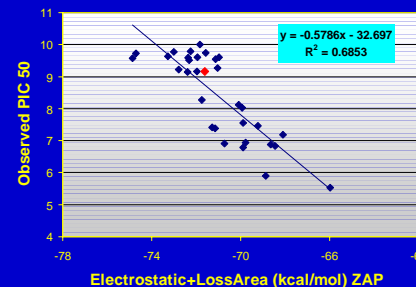
- 针对测试集中的系列化合物，收集所关注的理化性质/生物活性数据。
- 数据收集的途径主要有三种：科学文献的查阅、收集，活性/性质数据库中获取以及实验室中测试。
- 活性/性质数据必须要准确、可靠、标准化，另外毒性实验终点本身也应该具有显著的毒理学和生态学意义，并且容易进行机理上的解释。
- 一般数据具有两种：连续性响应数据（如急性毒性  $LC_{50}$ 、辛醇/水分配系数等）和非连续性响应数据（如致癌性数据阳性/阴性等）。

### (3) 分子结构描述



- 对于测试集化合物，首先应用分子模拟方法，构建正确的二维或三维分子结构，采用构象分析、分子力学等方法获得最优化的构象，进行分子结构计算，获取分子的结构信息。  
(以量子化学参数为例)
- 分子结构描述符的范畴包括理化性质参数例如辛醇/水分配系数，摩尔折射率等，一维分子组成参数例如碳原子数目、氢给体数目、分子量、平均原子极化率、平均原子范德化体积等，二维分子结构参数例如拓扑学参数、分子碎片常数等，以及三维分子结构参数例如三维分子力场参数、分子量子相似性参数、分子形状参数、分子电子结构参数等。

## (4) QSAR模型构建

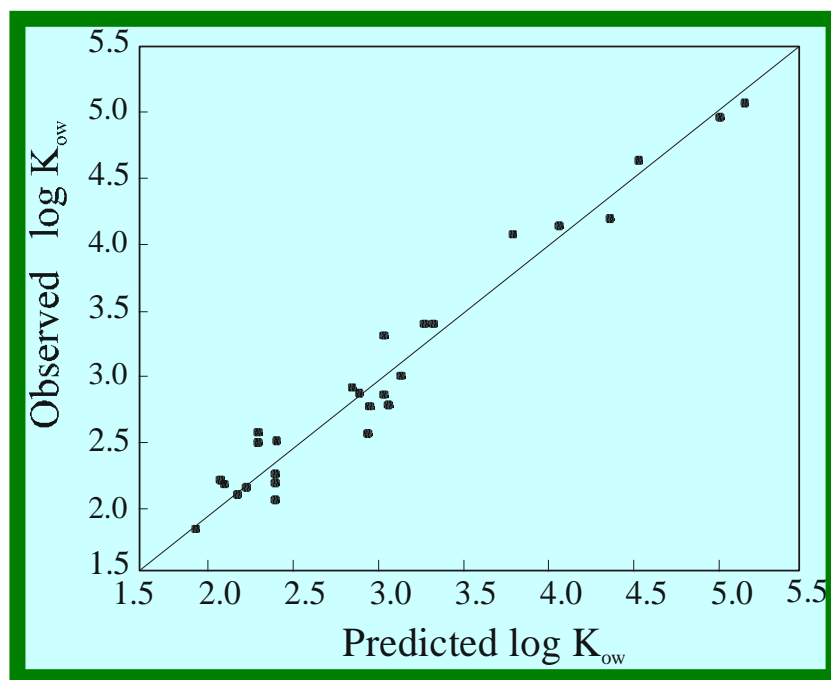
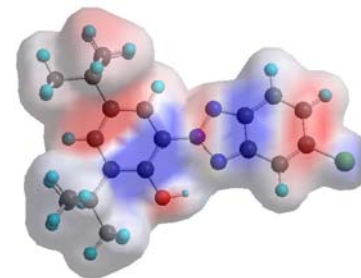


- 应用特征变量筛选方法筛选包含丰富结构信息的描述符；
- 应用统计方法或其他建模方法将训练集有机化合物的性质/活性数据与分子结构参数数据联系起来，建立QSAR模型；
- 常用的统计分析方法有回归分析方法，偏最小二乘分析、因子分析、主成分分析、聚类分析等，其中多元回归分析是目前应用最多的方法。
- 近来人工神经网络和遗传算法等高级建模方法也得到了越来越多的关注和应用。

## (5) QSAR模型的评价、验证



- 模型的拟合优度检验
- 拟合参数:
  - $r^2$  或  $r^2_{adj}$ : 回归系数的平方或自由度校正的  $r^2$ , 代表所解释的方差
  - P-显著性水平
  - F-检验值
  - 标准估计误差 (SE) 等。
- 比较: 真实值与拟合值(计算值)之间残差的大小
  - 如果大于2倍的标准偏差, 为离群值
  - 如果小于2倍的标准偏差, 可以接受



## 量化参数 (QC) 模型

$$-\log K_{ow} = 5.6256 - 7.4768 qH^+ + 0.8049 E_{lumo} + 0.0088 \alpha$$

$$R^2_{adj} = 0.97, SE = 0.097, F = 127.6, p = 0.0000, n = 13$$





## ■模型的稳健性: 交叉验证 (Cross-validation)

两种方式:

Leave-one-out;

Leave-more-out。

- 这样得到一个交叉验证的 $r^2$ , 即 $q^2$ 和一个标准预测误差 (SEP), 用来评价模型的稳健性。

$$q^2 = 1 - \left[ \frac{\sum (y - y_{pred})^2}{\sum (y - \bar{y})^2} \right]$$

$$SEP = \sqrt{\frac{\sum (y - y_{pred})^2}{N}}$$

# 交叉验证的程序

剔除 $m$ 个样本

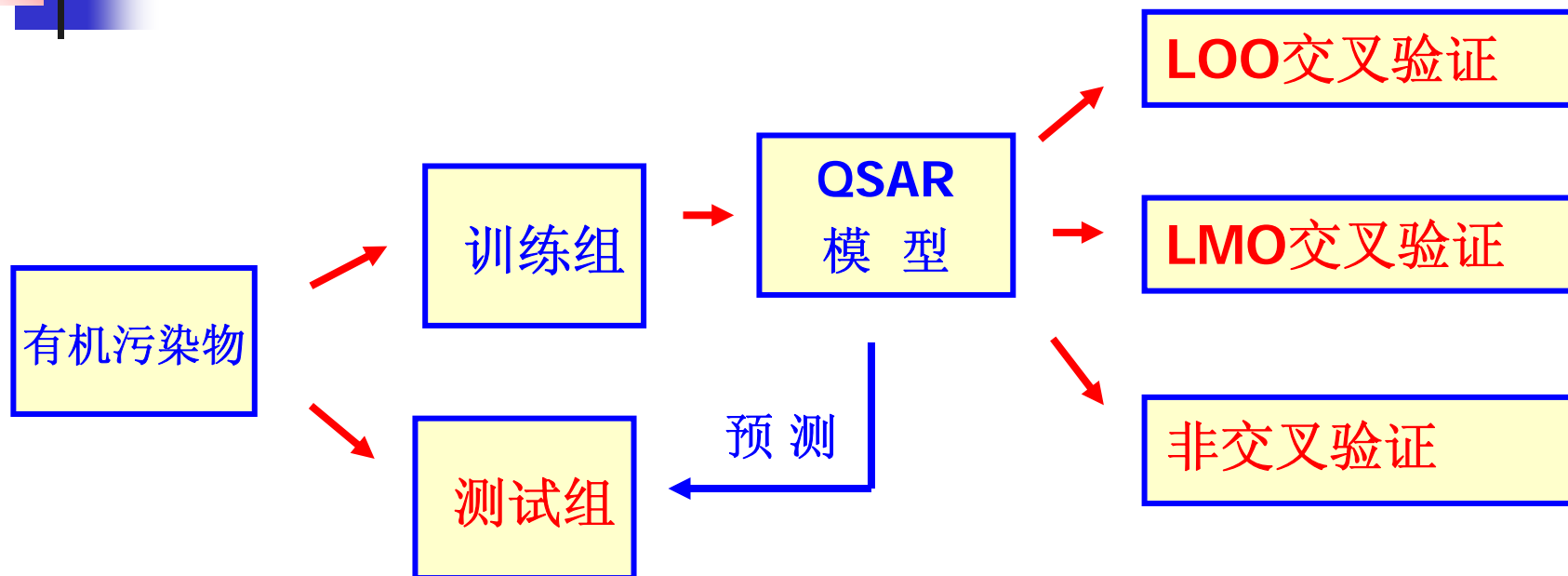
有放回地抽取 $n-m$ 个样本重新构造模型

计算模型估计值

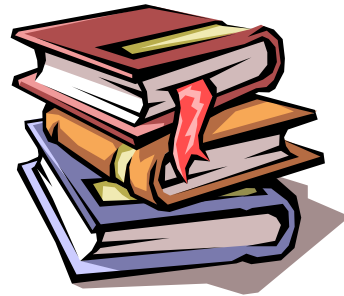
利用这些估计值的变异性作为原始估计值变异性的估计值

获取抽样变异性、置信区间、偏倚度  
等信息，并对分布未知的统计量进行推断

## 模型的预测能力检验—外部数据验证



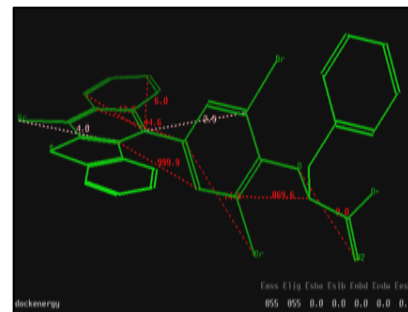
- 只有具有统计上的显著性、稳健的和具有高度预测能力的模型才能用来进行预测、评价和筛选，才有可能用来揭示有机污染物的活性机理和过程机制。



## (6) QSAR模型的应用

- 一是利用模型的预测能力对其它未知化合物的相关性质/活性进行预测，在效应评价和暴露评价中弥补缺失的数据，对环境有机化学品进行筛选和评价；
- 另一方面的应用是根据模型的组成与形式，结合已有的化学、生物学知识，探求有机污染物的环境过程机制和生态效应机理。
- 另外还可以根据所阐明的结构-活性关系结果，为设计高效、低毒无害的新型环境友好化学品指明方向。

## 第三节 分子结构表征

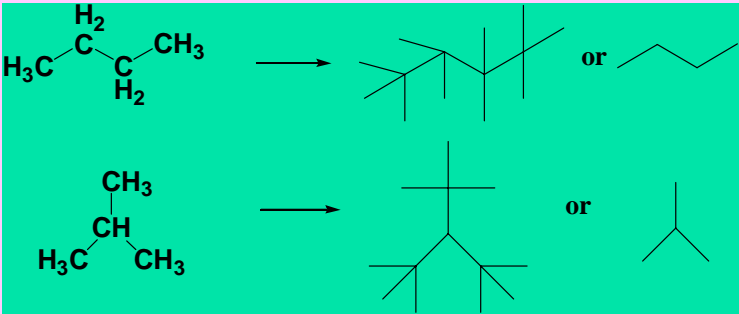


分子结构的表征,即化合物分子结构的参数化,把污染物的平面或空间立体结构表达为数值等计算机可以处理的信息.这是建立QSAR模型的关键步骤之一,理想的分子结构参数能够准确、完整地表达影响生物活性的结构信息,因此选取足够而合适的分子结构参数,对于QSAR模型的建立至关重要。

- 分子结构的表示
- 常用的分子结构参数

# 1. 分子结构的表示

## 分子结构式的五级表示方法

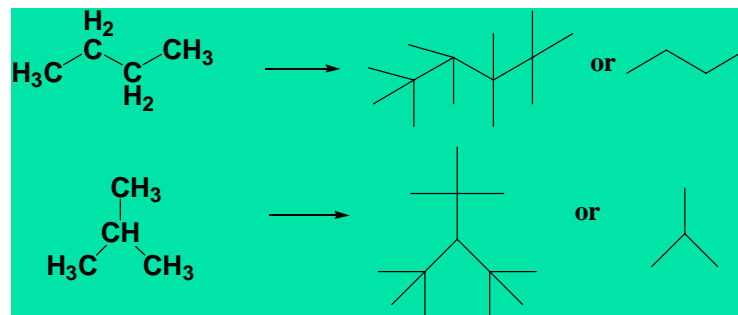
A	元素组成: C, H	
B	分子式: $C_nH_{2n+2}$	
C	具体分子式: $C_4H_{10}$	
D	分子拓扑结构	
E	量子结构	$\psi_1 = C_1\phi_1 + C_2\phi_2 + \dots$ $\psi_2 = C_1'\phi_2 + C_2'\phi_2 + \dots$



## A、B、C—级结构

- 从最简单的(A)级结构所得到的信息，仅限于构成分子的原子种类，但根据化学及物理方面的经验，可以推测出该分子具有烃的物理及化学性质，例如，该分子一般不溶于水，可燃烧等；
- 从(B)级结构，除了能得到从(A)级结构推得的相似理化性质外，还可以得出另外一些信息，例如，根据分子式可以知道分子中原子的种类和结合比，并可知道该分子具有开链烷烃的理化性质；
- 从(C)级结构就能获得更具体的信息，例如，可以推知该分子为丁烷，并可说出丁烷的某些物理性质，如溶解度、比重、化学反应性等，但由于异构现象，人们不能直接推出分子的结构式，即原子间的连接次序；

## D—级结构



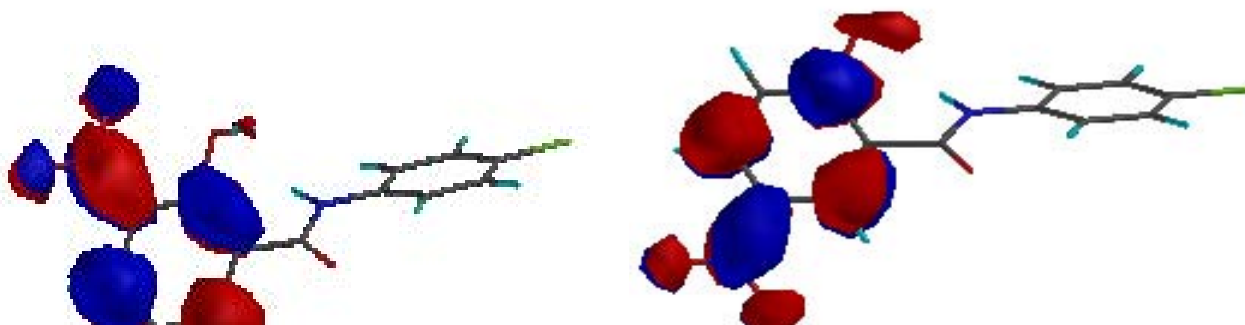
- (D)级结构实际上就是分子的平面拓扑结构，
- (D)级结构的表示是以分子中原子以共价键相连为基础的，其他作用力如氢键、范德华力等则没有表示，实际上只考虑了分子骨架。
- (D)级结构包含的信息为：(1)总原子数目；(2)原子的连接次序及键型，(3)各种原子的种类和数目。
- 根据(D)级结构，可以了解到分子中原子是如何排列和连接的；例如，图示分子中原子有两种排列方式，即异构体，根据物理和化学方面的经验，可以推测出大量有关该分子的性质。这种结构也是化学中经常使用的一种分子结构式。



## E—级结构

$$\psi_1 = C_1 \phi_1 + C_2 \phi_2 + \dots$$
$$\psi_2 = C_1' \phi_1 + C_2' \phi_2 + \dots$$

- (E)级结构是在量子力学水平上表示的，它原则上可以反映分子的全部信息，这些信息都是用数值来表示的，它包括分子中电子的几率、能量以及核的位置等。
- 由于在处理薛定谔方程时需考虑所有粒子间的作用，因此，量子力学计算较复杂，时间较长





## 2. 常见的分子结构参数

- 理化性质参数例如辛醇/水分配系数，摩尔折射率等
- 一维分子组成参数例如碳原子数目、氢给体数目、分子量、平均原子极化率、平均原子范德化体积等，
- 二维分子结构参数例如拓扑学参数、分子碎片常数等，
- 三维分子结构参数例如三维分子力场参数、分子量子相似性参数、分子形状参数、分子电子结构参数等。
- 在定量结构-活性关系研究中应用比较广泛的分子结构描述符主要有理化性质参数、拓扑学参数、量子化学参数，近来三维结构-活性关系方法在环境科学中受到了越来越多的关注



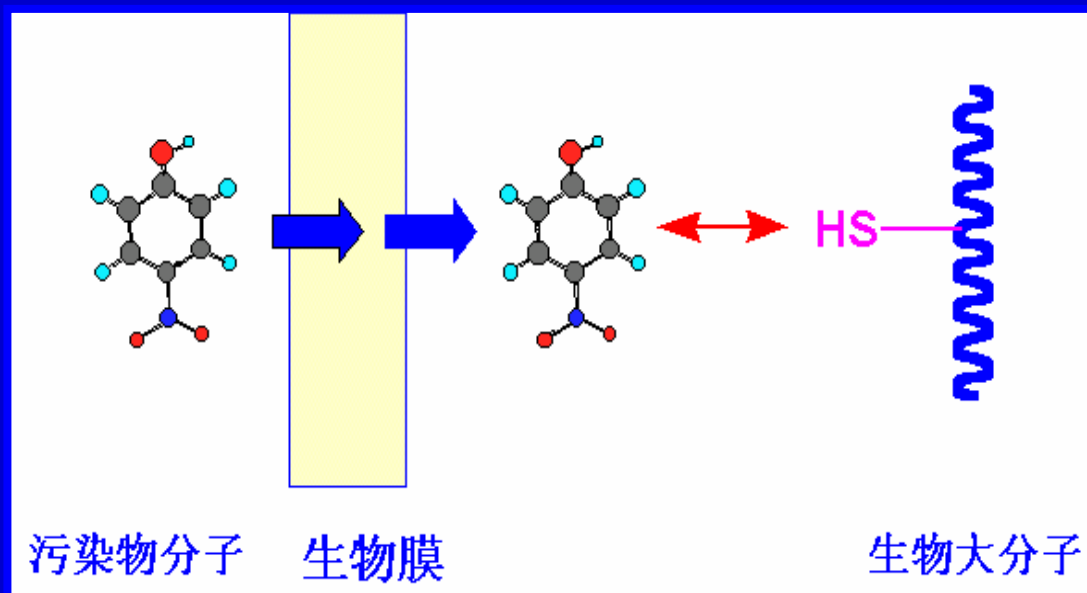
## (1).理化性质参数

---

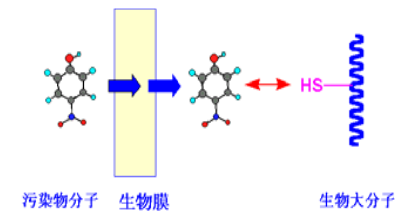
- 正辛醇/水分配系数 ( $K_{OW}$ )
- 溶解度 ( $S_w$ )
- 摩尔折射率 (MR)
- 酸离解常数 (pKa)

$K_{OW}$ 是有机化合物在水-辛醇平衡体系中两相浓度之比，通常取对数以后，用来作为有机污染物疏水性的量度。在定量结构-活性关系研究中， $K_{OW}$ 常模拟污染物或药物透过生物膜进入生物体内的能力，与生物富集系数、土壤吸附系数、有机碳吸附系数、生物活性等具有显著的相关性

## ◇ 有机污染物的毒性作用一般过程



$$\text{Log}[1/C] = A [\text{透过}] + B [\text{作用}] + C$$



## ■ 临界体积理论与麻醉性QSAR方程

➤ 非反应性有机污染物穿过细胞膜进入了细胞类脂双层结构组织，当达到**临界体积**，由于化学物质的膨胀阻塞了离子通道，破坏了细胞正常代谢作用，产生了**可逆的麻醉毒性作用**，毒性效应的大小和化合物的亲脂性大小成正比，即：

$$\log (1/C) = a \log K_{ow} + C$$

- 可以用 **$\log K_{ow}$**  来定量预测麻醉性毒性



## 取代苯酚对高等植物的抑制毒性

- 研究对象：28种取代苯酚
- 毒性终点：黄瓜种子根生长的半抑制浓度（ $RC_{50}$ ）
- 毒性作用机制：极性麻醉与生物反应性
  - 极性麻醉作用：

$$LC_{50} = 0.94 \log K_{ow} + 2.04$$

$$n = 22 \quad r^2 = 0.89 \quad SE = 0.10 \quad P = 0.00001$$

- 生物反应性：不能够建立与疏水性的相关关系

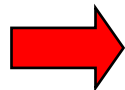
## ❖ 反应性化合物的毒性机制

### ■ 反应性化合物的亲电性质:

硝基取代苯酚 (3)

对苯二酚

氨基苯酚和萘醌 (3)



氧化磷酸化解偶联  
醌类作用机理



亲电性质



### ■ 反应性与 $E_{LUMO}$ 的关系

$$RC_{50} = -0.91 E_{lumo} + 2.73$$

$$n = 6 \quad r^2 = 0.81 \quad SE = 0.14 \quad P < 0.001$$

- Michael 加成反应
- 氧化压力



## (2) 电子效应常数

- 化合物分子电子性质的参数可以用很多参数来表示：Hammett电性效应参数、场常数 $F$ 和共轭参数 $R$ 、 $pK_a$ 值、光谱参数、电荷转移常数、偶极矩、氢键参数和量子化学电子结构参数。电子效应反映的是分子局部原子和基团的性质。其中最著名的是Hammett电性效应参数

$$\sigma_x = \log K_x - \log K_H$$

$K_x$  是苯甲酸25°C在水中的离解常数，二者对数之差代表了取代基的电性效应





### (3).立体效应参数

- 化合物的立体结构性质可以用很多参数来描述，Taft常数 $E_s$ 是经典的取代基立体参数，用来描述化合物分子取代基的空间立体效应的作用：

$$E_s = \log \left( \frac{K_x}{K_H} \right)_H$$

- 分子摩尔体积、范德化半径、原子间距以及部分分子连接性指数、Charton立体参数、最小立体差异等参数都曾经被用来作为立体参数，表征分子间作用的立体效应。



## (4).溶剂化自由能

---

- 化合物的大多数性质和反应是在溶液中发生的，溶剂对化合物性质和反应有重要影响；对于化合物的生物活性，溶剂效应的影响也是非常重要的，如：对酶的催化活性、蛋白质的折叠和展开方式、生物大分子的构象等有着较大的影响，化合物在体内的传输、与受体的作用和代谢也离不开溶剂效应的作用。溶剂化自由能是化合物分子与溶剂介质相互作用的宏观反映，可以用来估测与溶剂效应相关的化合物性质。

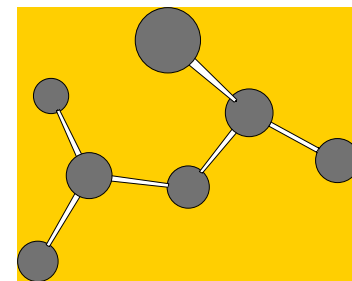


## (5)分子拓扑参数

---

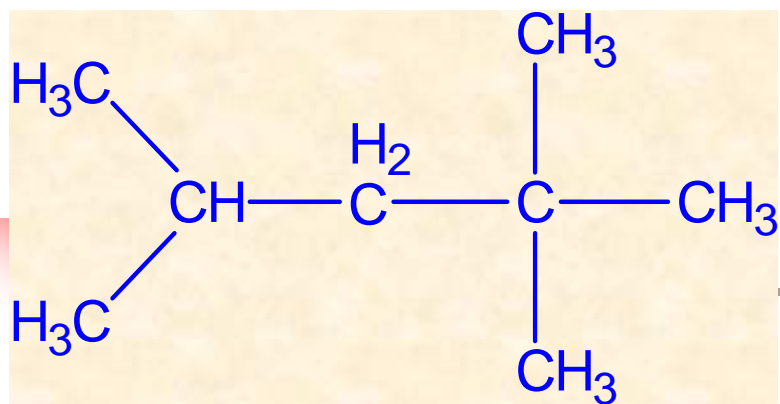
- 分子拓扑指数是分子拓扑结构信息的反映
- 应用拓扑指数提炼分子结构信息，一般有三个步骤：分子结构的图形化、图形结构的矩阵化和数值化。
- 理想的拓扑指数应满足如下要求：能够反映分子骨架中原子的种类、数目以及化学键的数目，不饱和键的位置和数目，环的大小和数目，碳桥的位置和数目，原子的排列顺序等。
- 分子拓扑指数表示方法很多，文献报导的拓扑指数已超过百种。目前比较有影响的一些拓扑指数有：Wiener指数、Hosoya指数、Balaban指数、分子连接性指数等，其中应用最广泛的是分子连接性指数。

# 分子连接性指数

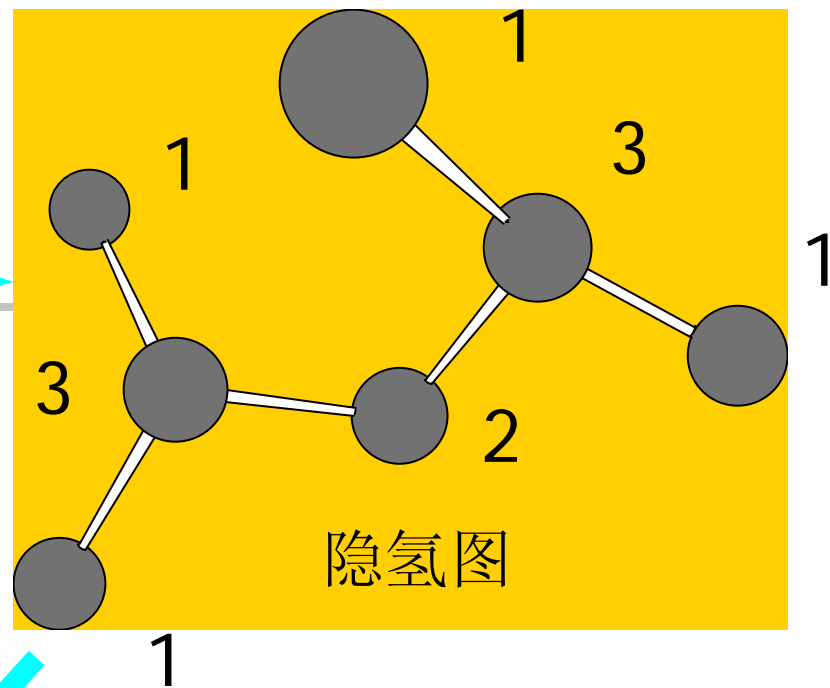


- 分子连接性指数由Randic于1975年提出，之后Kier和Hall又发展形成了分子的价连接性指数。它们都是根据分子结构的隐氢图得到的拓扑学指数，以所有不同形式相连的子图求和而得：（1）根据化合物结构式写出隐氢图；（2）标出原子的点价；（3）代入如下公式进行计算

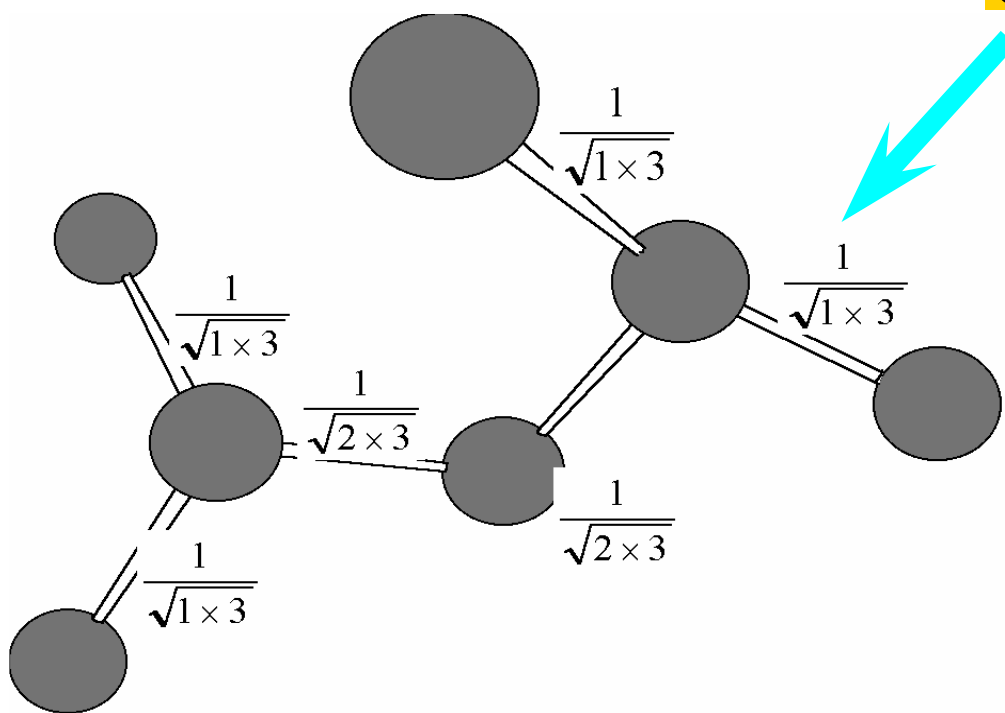
$$^n\chi = \sum \left( \delta_i \delta_j \cdots \delta_n \right)^{-1/2}$$



2,4-二甲基戊烷



隐氢图



全部边值和  
 ${}^1X=3.126$

## 分子连接性指数与土壤吸附系数的相关性

- ❖ Sablijic研究了一阶分子连接性指数与土壤吸附 ( $K_{om}$ ) 的相关性，包括建立预测模型以及检验预测模型对极性化合物土壤吸附系数的预测结果。
- ❖ 通过对72个化合物的试验，包括多环芳烃、杂环和取代多环芳烃、烷基或氯代苯、氯代苯、氯代烷烃和烯烃以及氯酚等，建立了以下的相关方程：

$$\log K_{om} = 0.53^I X + 0.54$$

$$n=72, \quad r=0.976$$



## (6) 色谱保留指数

---

- 色谱保留指数不仅是化合物色谱定性、定量分析的重要参数，而且还能够作为表示结构的参数应用于定量结构-性质/活性的研究，因为它是一种表示分子相互作用的综合性热力学参数。
- 色谱保留指数广泛地用于预测正辛醇-水分配系数, 水溶解度等理化性质
- 在QSAR研究中常用的色谱保留指数主要为气相色谱中的Kovats指数及高效液相色谱中的容量因子等。



## ■ 气相色谱Kovats指数

$$I_i(T) = 100 \times \frac{\log t'_{Ri} - \log t'_{RZ}}{\log t'_{R(Z+1)} - \log t'_{RZ}} + 100Z$$

$I_i(T)$  在一定温度下，在指定柱上物质i的色谱Kovats保留指数；

$t'_{RZ}$  是指碳数目为Z的标准同系物的调整保留时间；

$t'_{R(Z+1)}$  是指碳数目为Z+1的另一个标准物的类似参数；

$t'_{Ri}$  是物质I的调整保留时间。





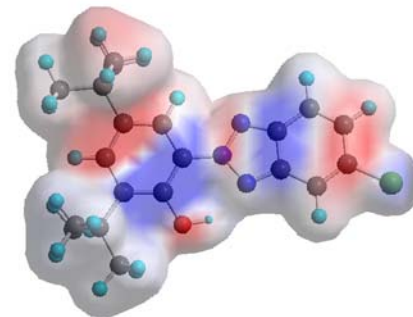
## 高效液相色谱的容量因子

化合物的HPLC保留性质可以用相容量比或容量因子来表示，容量因子  $k'$  定义为：

$$k' = \frac{t_R - t_M}{t_M} = \frac{V_R - V_M}{V_M}$$

$t_R$  和  $V_R$  —分别为溶质的色谱保留时间和保留体积。

$t_M$  和  $V_M$  —分别为所谓的无保留溶质的保留时间和保留体积



## (7)量子化学参数

### ■ 微观粒子运动 的Schrödinger方程

Schrödinger方程

$$\hat{H} \psi = E \psi$$

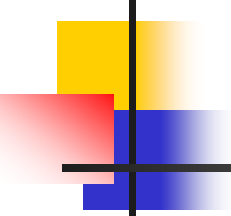


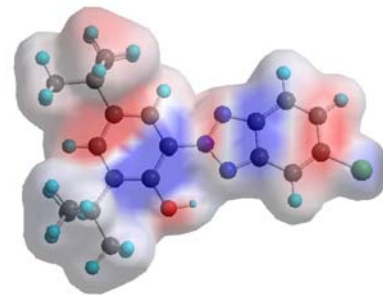
$$\hat{H} = -\frac{1}{2} \sum_p \nabla_p^2 + \frac{1}{2} \sum_{p \neq q} \frac{1}{r_{Ap}} - \sum_A \sum_P \frac{Z_A}{r_{aP}} - \frac{1}{2} \sum_A \frac{1}{M_A} \nabla_A^2 + \frac{1}{2} \sum_{A \neq B} \frac{Z_A Z_B}{R_{AB}}$$



Hartree-Fock-Roothaan 方程

$$\sum_p (F_{\mu\nu} - \epsilon_i S_{\mu\nu}) C_{\nu i} = 0$$


$$\sum_p (F_{\mu\nu} - \varepsilon_i S_{\mu\nu}) C_{\nu i} = 0$$

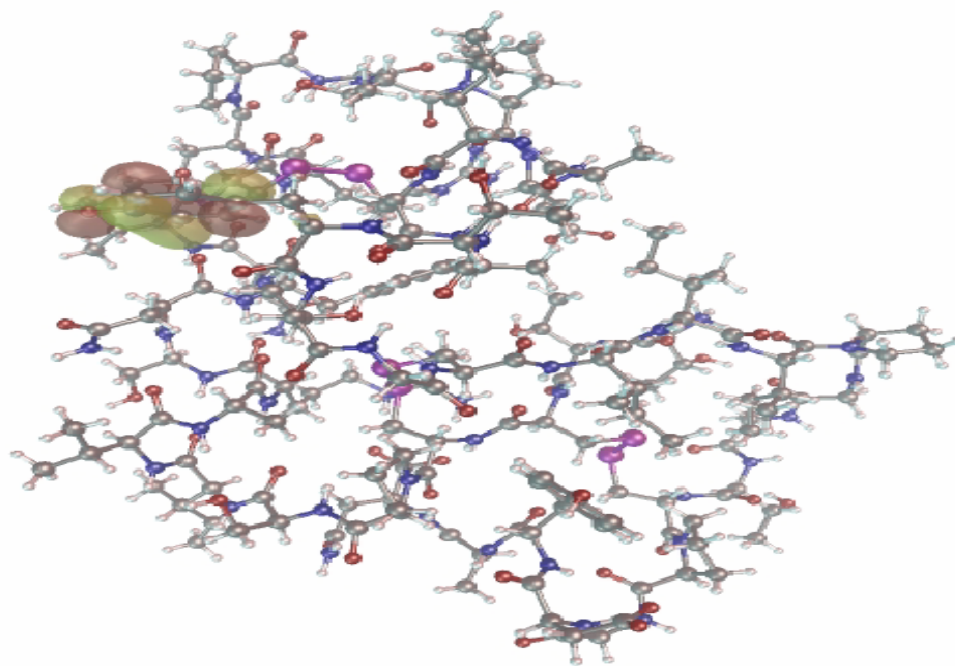


- 根据对Hartree-Fock-Roothaan方程的求解过程中是否引入其他的近似和简化，产生了不同的分子轨道计算方法
  - 从头计算法 (*ab initio*)
  - Hückel分子轨道法 (HMO)
  - 推广的Hückel分子轨道法 (EHMO)
  - 半经验分子轨道法 (AM1, PM3, MNDO)
  - 其它方法：密度泛函理论 (DFT) 等



# 半经验分子轨道计算 (MOPAC)

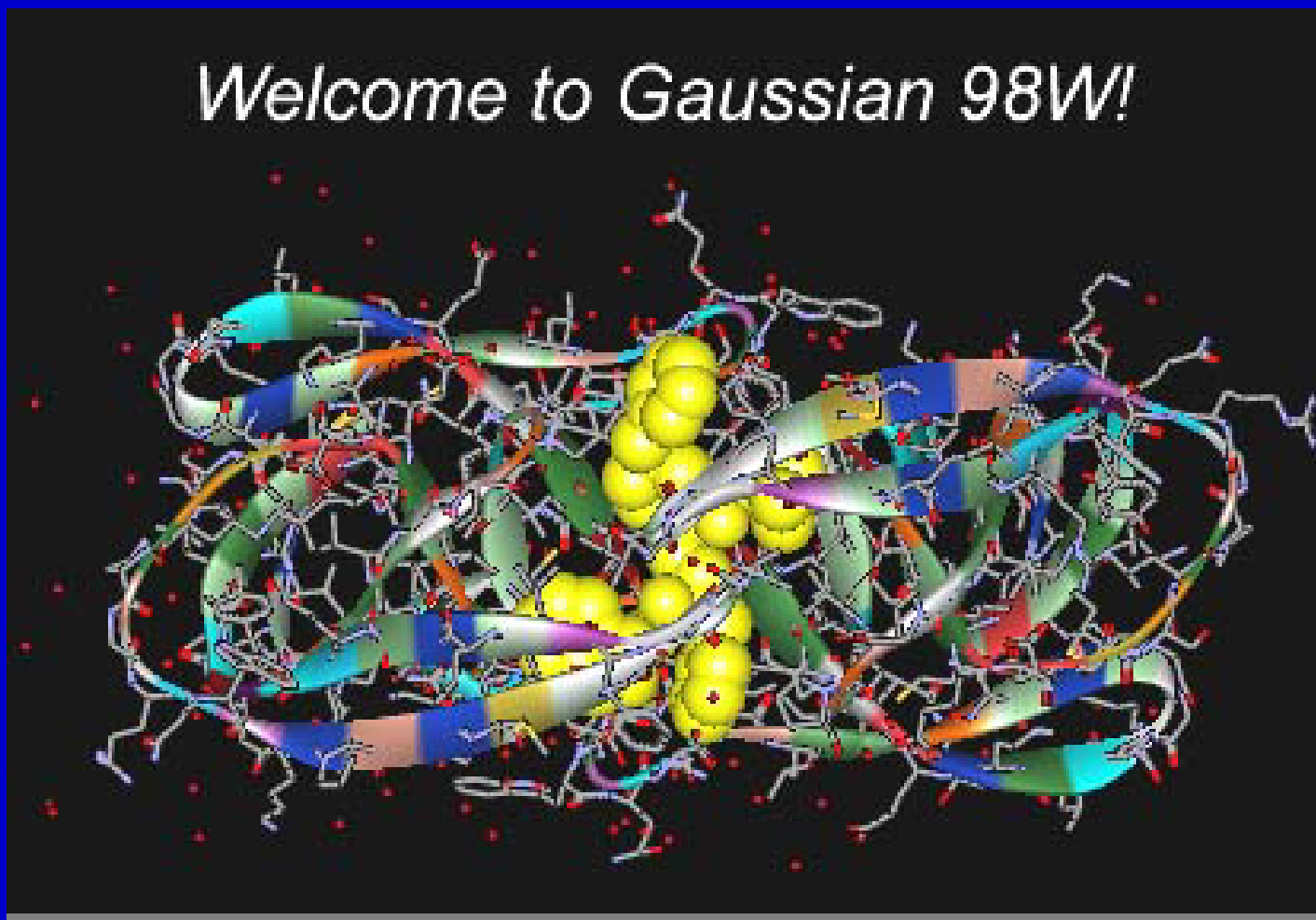
**MOPAC 2000 V1.3 User's Manual**  
James J. P. Stewart



**FUJITSU**

# 从头计算法量子化学计算-Gaussian 98

*Welcome to Gaussian 98W!*





## ■化学桌面办公软件-Chemoffice

**CambridgeSoft**  
www.cambridgesoft.com

**CS ChemOffice**  
Drawing, Modeling  
and Information

**ChemOffice**  
Drawing, Modeling and Information

**Chem3D**  
Ultra version 7.0.0

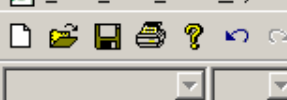
?1986-2001 CambridgeSoft.Com  
All Rights Reserved

Licensed to: W&Y  
nju  
7654321

**CS ChemDraw**  
Structure Drawing

**CS Chem3D**  
Molecular Modeling

**CS ChemFinder**  
Chemical Information



Tools



Atom Properties...  
Bond Properties...  
Bracket Properties...

Check Structure  
Clean Up Structure Shift+Ctrl+K  
Expand Label  
Contract Label

Add Multi-Center Attachment  
Add Variable Attachment

Map Reaction Atoms  
Clear Reaction Map

<sup>1</sup>H-NMR Shifts  
<sup>13</sup>C-NMR Shifts  
Make Spectrum-Structure Assignment

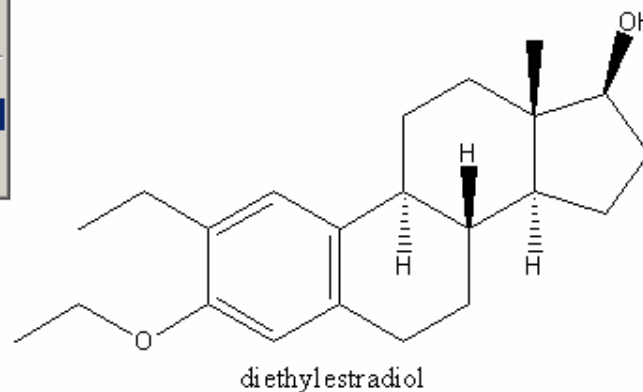
Define Nickname...

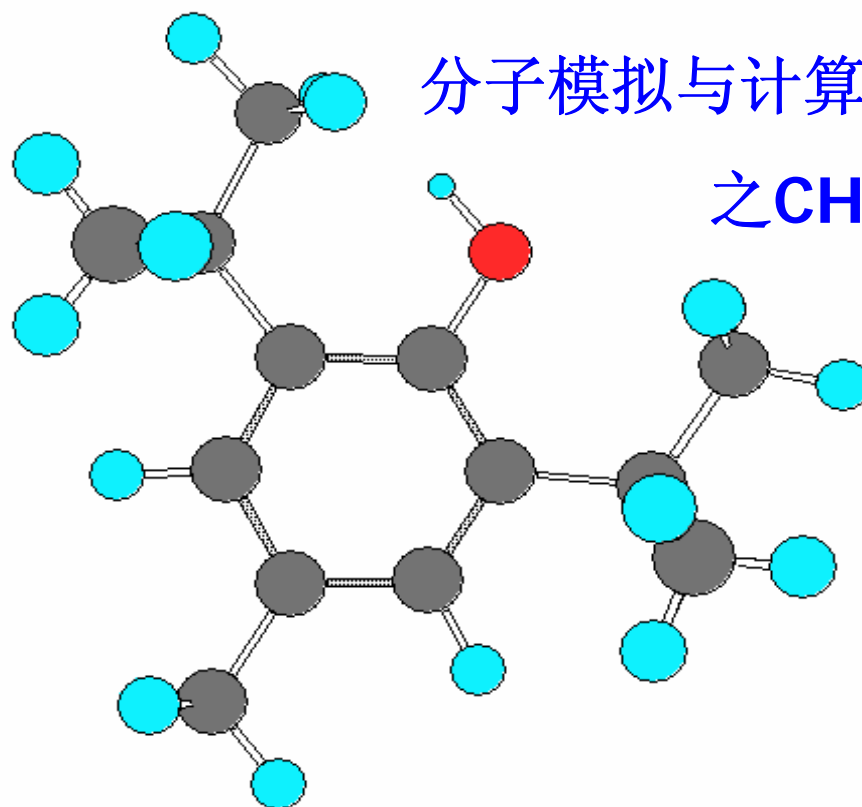
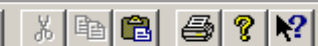
**Convert Name to Structure** Shift+Ctrl+N

Convert Structure to Name Alt+Shift+Ctrl+N

About AutoNom

# 分子构建之CHEMDRAW





# 分子模拟与计算-Chemoffice 之CHEM3D

195 messages

Adding lone pairs to O[7]

System: 2 3 4 5 1 6

Note: Some parameters are not finalized (Quality = 3).

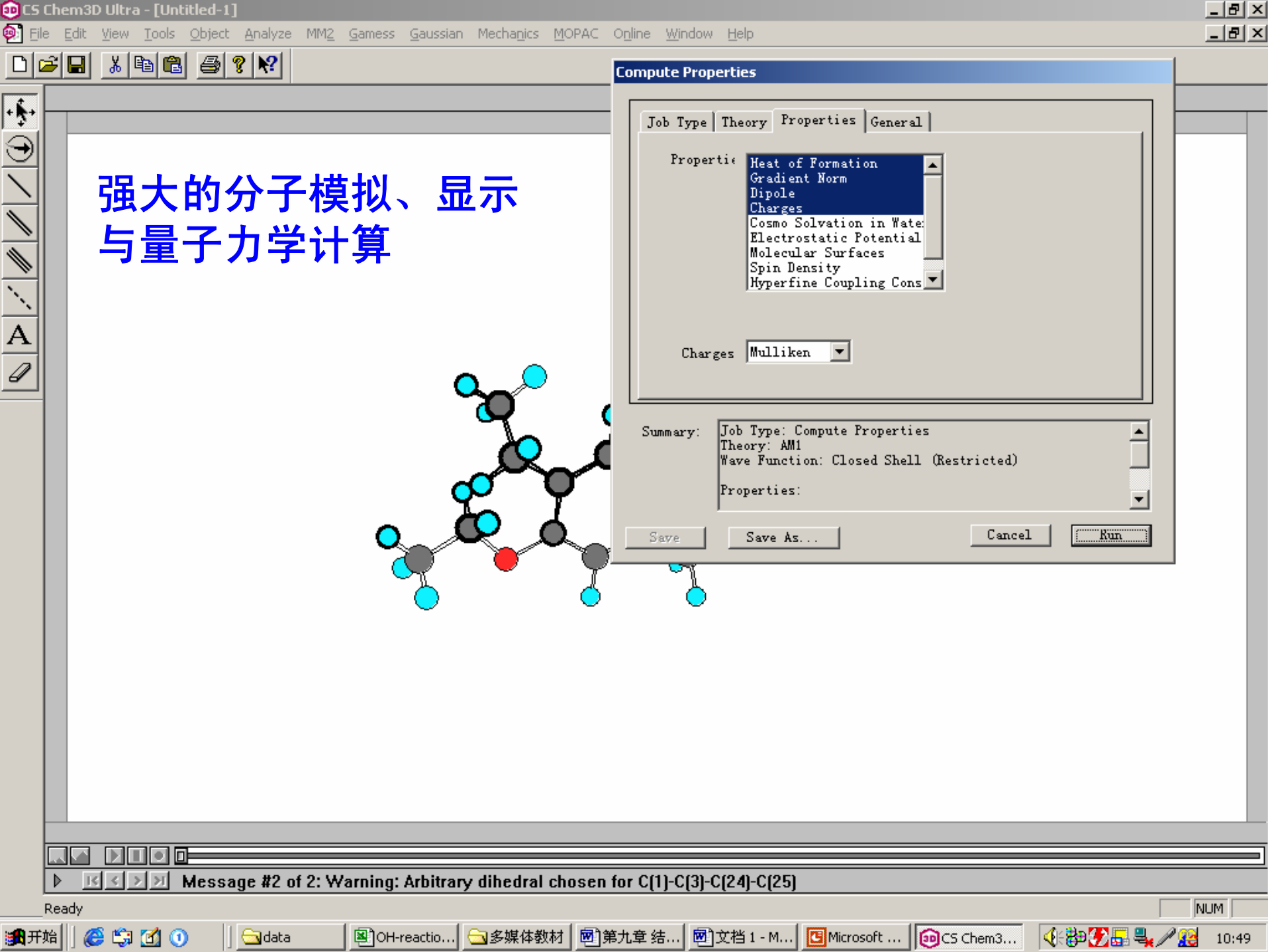
O[7]

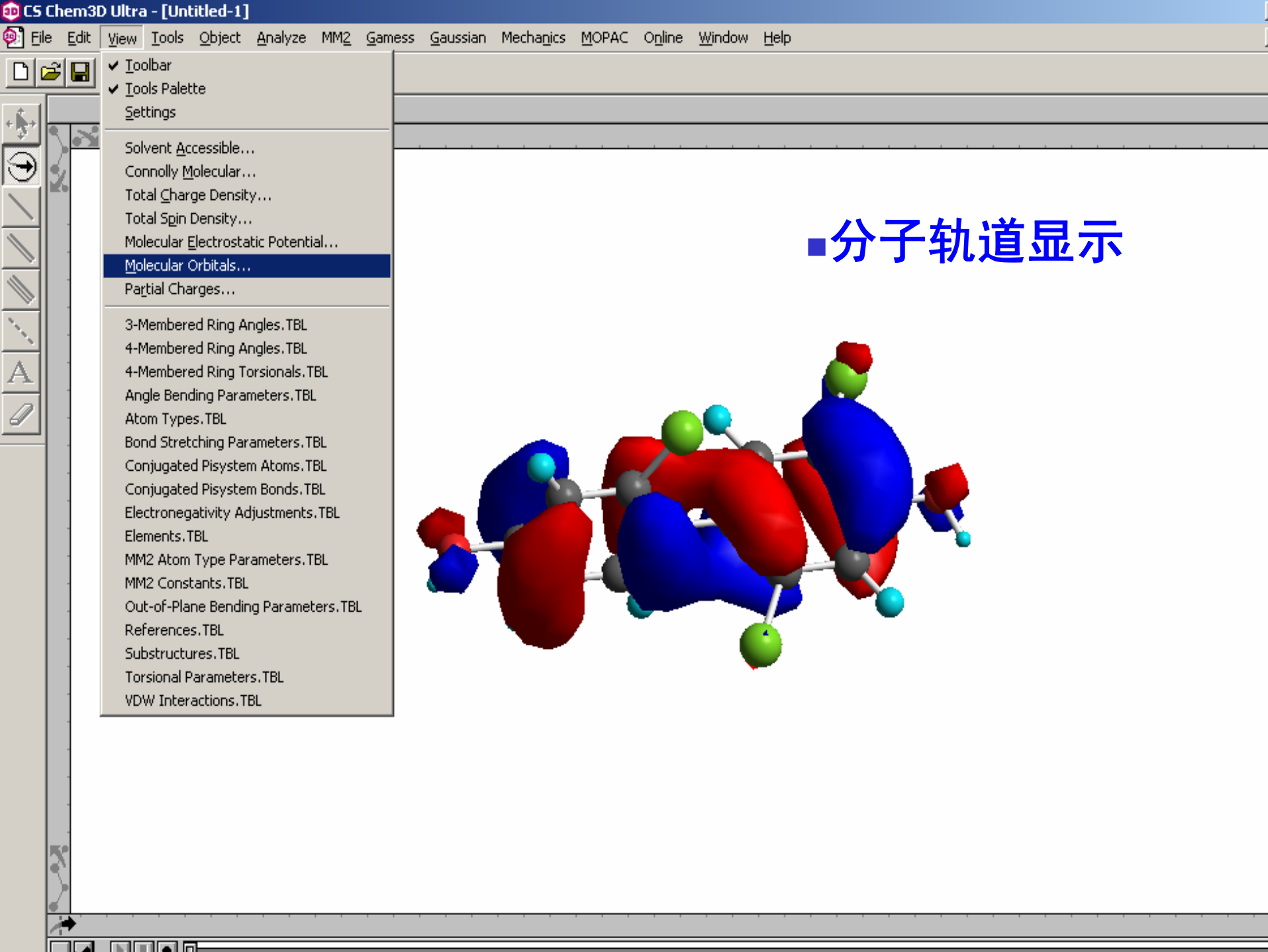
-

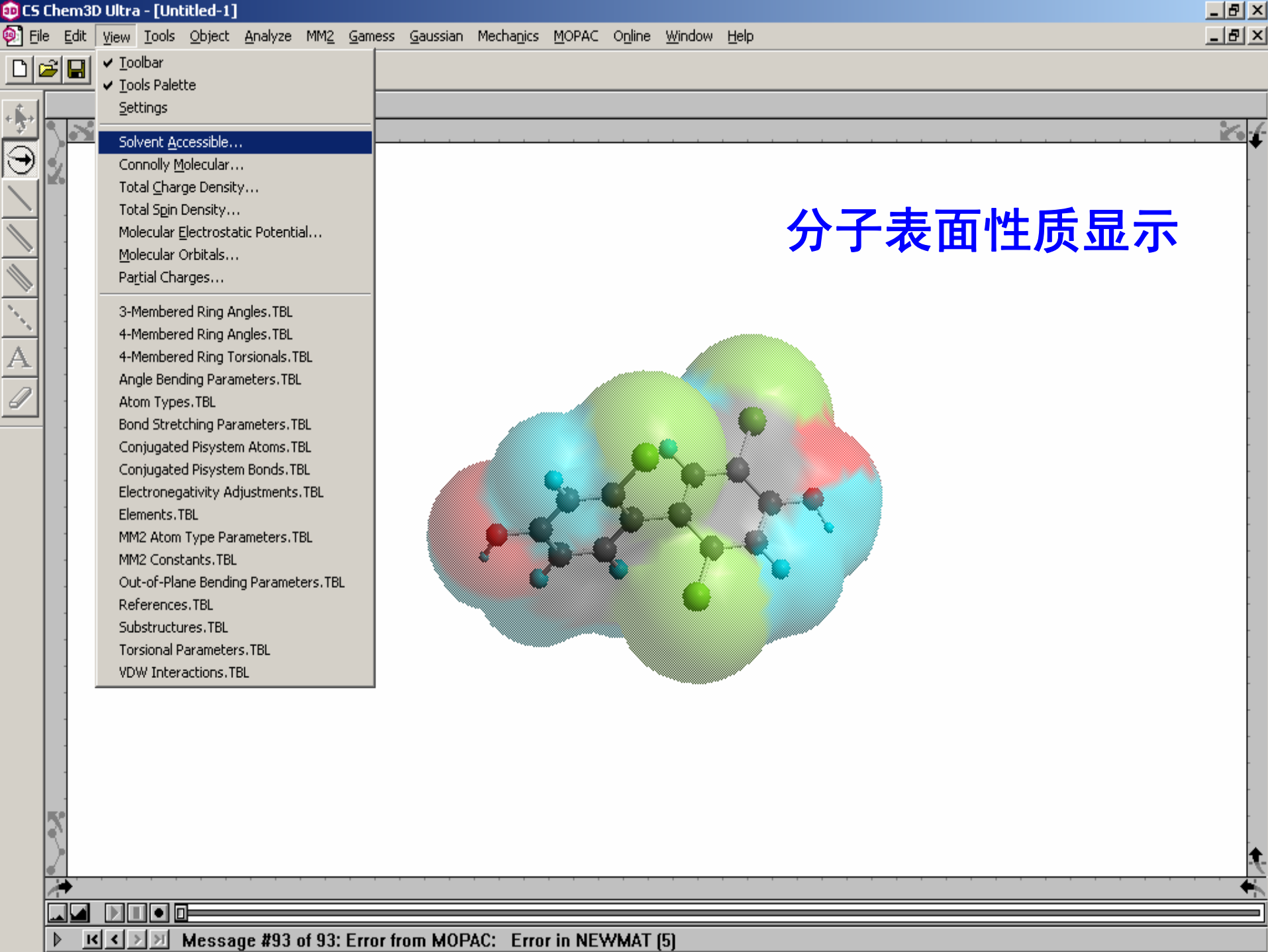
-

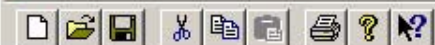
NUM



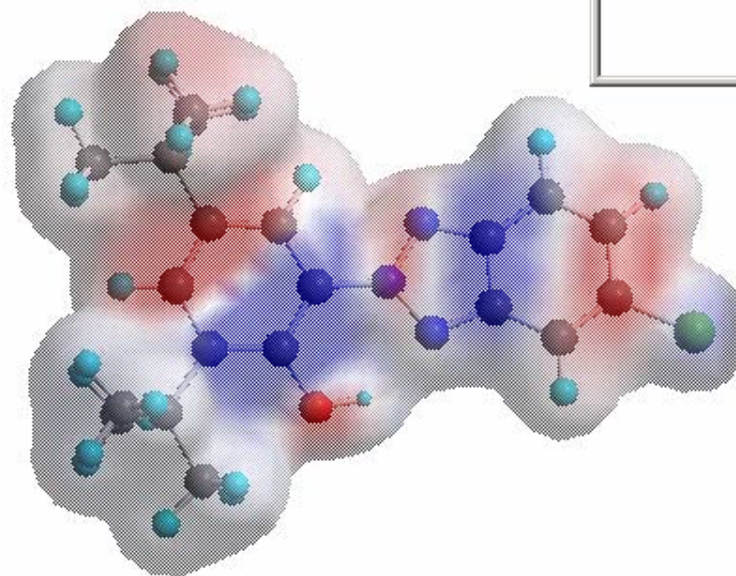








# 分子电荷性质显示



Total Charge Density Surface

Calculation: AM1

Surface Type: Translucent

Map Property: Molecular Orbital (HOMO)

Solid Color: None (Solid Color)  
Molecular Orbital (HOMO)  
Electrostatic Potential

Isocharge: Partial Charges

Grid: 30x30x30

Set Grid...

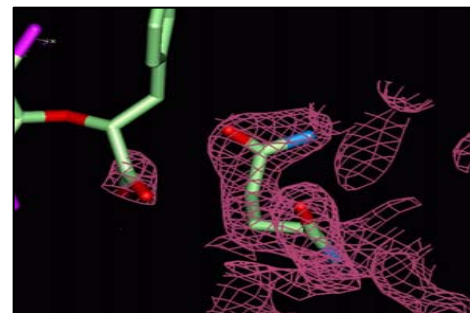
Hide Surface



Message #174 of 174: H[28] 88.16069

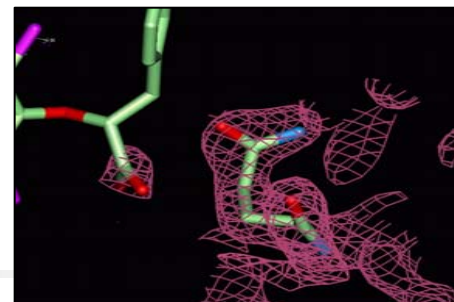
Ready

# 常用的量子化学参数



- 电荷参数:  $Q_A$ 、 $Q_{min}$ 、 $Q_{max}$ 、静电势电荷
- 分子轨道能量:  $E_{LUMO}$ 、 $E_{HOMO}$ 、绝对硬度、电负性等
- 轨道电子密度: 原子上的  $\sigma$  电子和  $\pi$  电子密度  $Q_A$ ,  $\sigma Q_A$ ,  $\pi$
- 超离域度: 亲电和亲核超离域度
- 原子-原子极化率: 原子自极化率
- 分子极化率:  $\alpha$
- 偶极矩和极化指数: 分子偶极矩  $\mu$
- 能量参数: 分子总能量、生成热、离子化势、结合能等
- 立体结构参数: 分子表面积、分子形状参数、分子体积

# 量子化学参数的优点

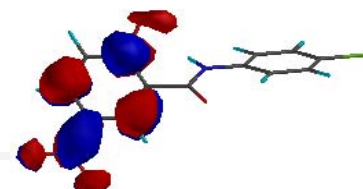


- 通过对化合物进行量子化学计算，可以获得有关分子、分子碎片以及取代基等的反应性、形状和键合性质，能够获得有关分子的电子结构、电子分布、立体化学及静态和动态的能量分布等信息
- 量子化学参数具有明确的物理意义，有利于探讨影响污染物的环境行为和过程机制的结构因素，探讨污染物的生态毒性机理；
- 同时，量化参数完全实现理论计算，不需要实验测定；
- 另外，量子化学参数不限于同系物的研究，因此量子化学参数在有机污染物的结构活性关系研究中具有广泛的应用前景。



# 应用举例—LUMO能量预测

## 含氮芳烃对高等植物的反应性毒性

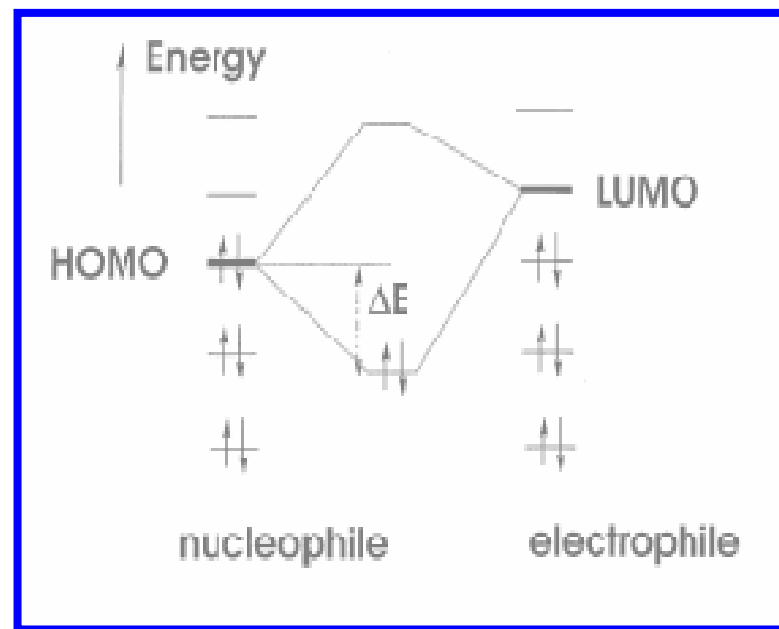


### ➤ 前线分子轨道理论:

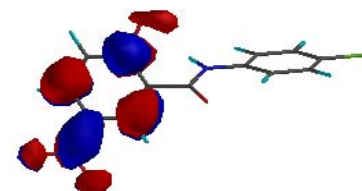
大多数化学反应在一个反应物的HOMO和另一个反应物的LUMO能够产生最大重叠位置及方向上发生。亲核反应物主要以其HOMO参与反应，亲电反应物主要以其LUMO参与反应，反应性的大小取决于 $\text{HOMO}_{\text{NU}}$ 与 $\text{LUMO}_{\text{ECE}}$ 之差

➤ 有机污染物的生物反应性取决于污染物的前线轨道能量

➤ 有机污染物的亲电反应性取决于本身的 $E_{\text{LUMO}}$



# (1) 应用举例—LUMO能量预测 含氮芳烃对高等植物的反应性毒性

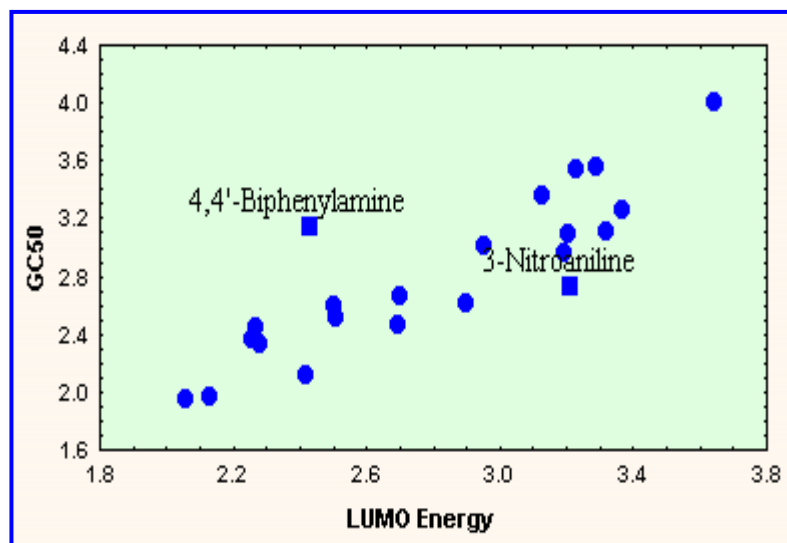


- 研究对象：21种含氮芳烃
- 毒性终点：黄瓜种子发芽抑制

$$GC_{50} = -0.936E_{\text{lumo}} + 2.414$$

$n=21$   $r^2=0.88$   $SE=0.2$   $P < 0.00001$

植物毒性与LUMO能量的关系

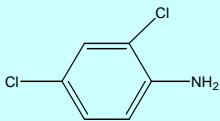
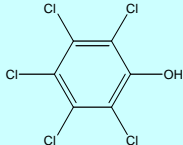
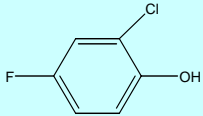
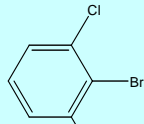
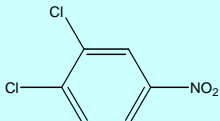
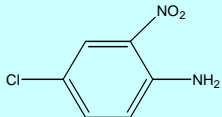
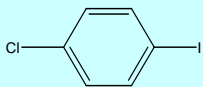
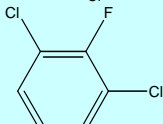
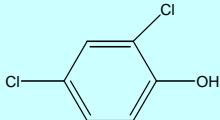
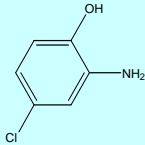
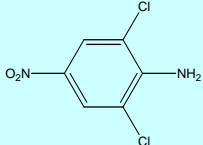
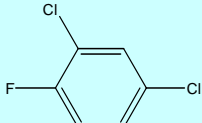

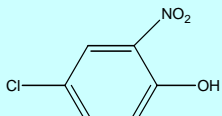
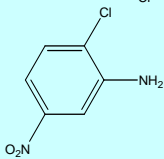
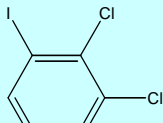
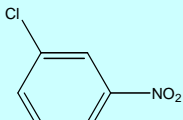
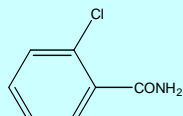
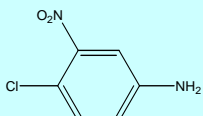
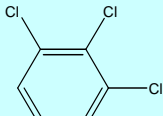
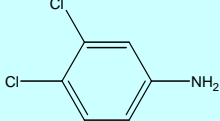

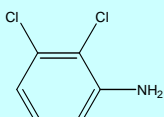
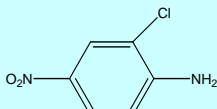
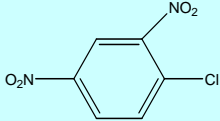
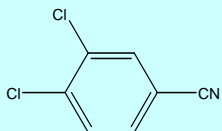
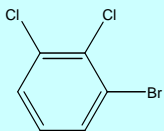
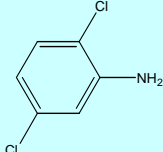


XD Wang et al., 2001. 913-916.

Environ. Toxicol. Chem.



## 2) 举例:应用QSAR模型预测卤代芳烃的 $\lg K_{ow}$

序号	化合物	序号	化合物	序号	化合物	序号	化合物
1		8		15		22	
2		9		16		23	
3		10		17		24	
4		11		18		25	
5		12		19		26	
6		13		20		27	
7		14		21		28	

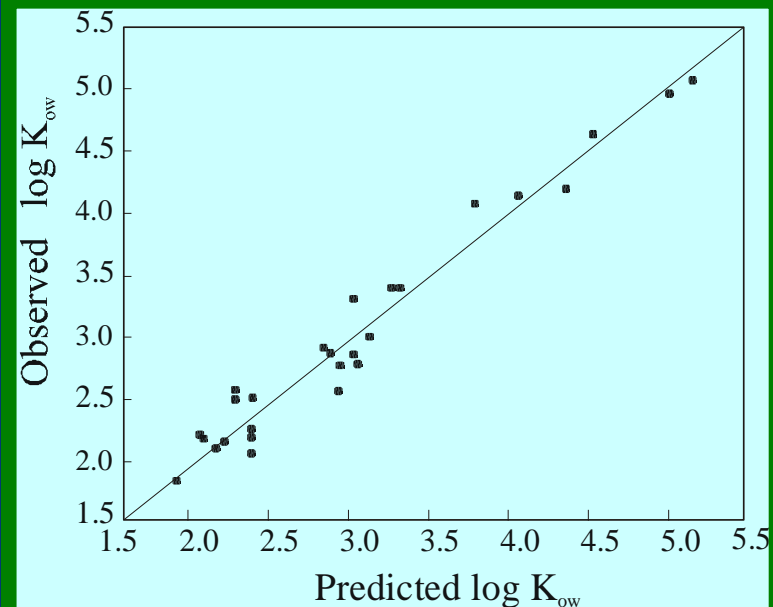
## 基于三种分子结构参数的QSAR模型

方法	回归方程	$R^2_{adj}$	$SE$	$F$
MCI	$\log K_{ow} = -0.102 + 0.998^2 \chi^v - 0.964 \Delta^0 \chi^v + 1.132^4 \chi_p^v$	0.955	0.196	185.3
TLSER	$\log K_{ow} = 10.752 - 3.015 \pi^* + 6.18 \times 10^{-2} V_{mc} - 1.465 q - 36.276 \varepsilon_h$	0.870	0.330	44.67
静电势法	$\log K_{ow} = 0.264 + 1.514 \times 10^{-2} MW + 1.677 V - 0.613 TE$	0.918	0.264	97.92

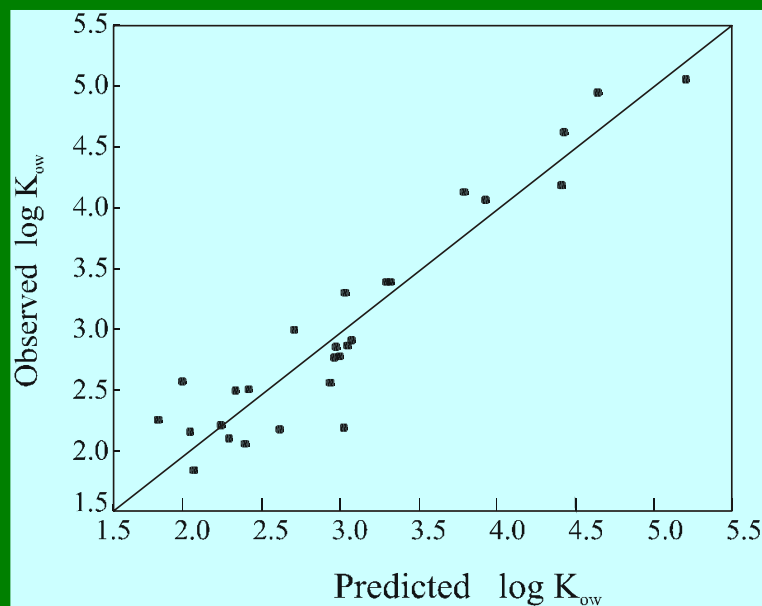
- MCI模型虽能较准确地预测分配行为，但其参数缺乏明确的理化意义
- TLSER和静电势模型中的描述符都具有明确的物理意义，能反映该组化合物的分配机制，但静电势模型的预测能力相对更高

预测值与观  
测值间关系

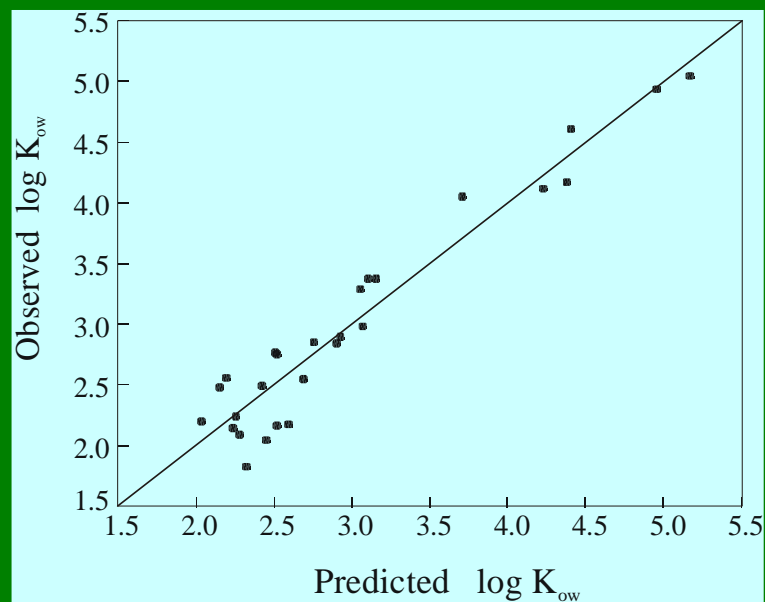
MCI 模型



TLSE模型



静电势模型




$$\log K_{ow} = 0.264 + 1.514 \times 10^{-2} MW + 1.677V - 0.613TE$$

---

***MW***: 分子量，与分子体积呈正相关。体积越大，化合物在水中形成与之匹配的空穴需更高能量，因而在正辛醇相分配更多，使Kow增大

***V***: 负静电荷电势，与分子间相互作用有关。该值越小，化合物与溶剂水分子之间的静电作用越弱，化合物分子更易分配进入正辛醇相，使Kow值升高

***TE***: 分子总能量，该值越负，表明化合物越稳定，其与水分子间作用的可能性越小，更多份额能进入正辛醇相，使Kow值增大

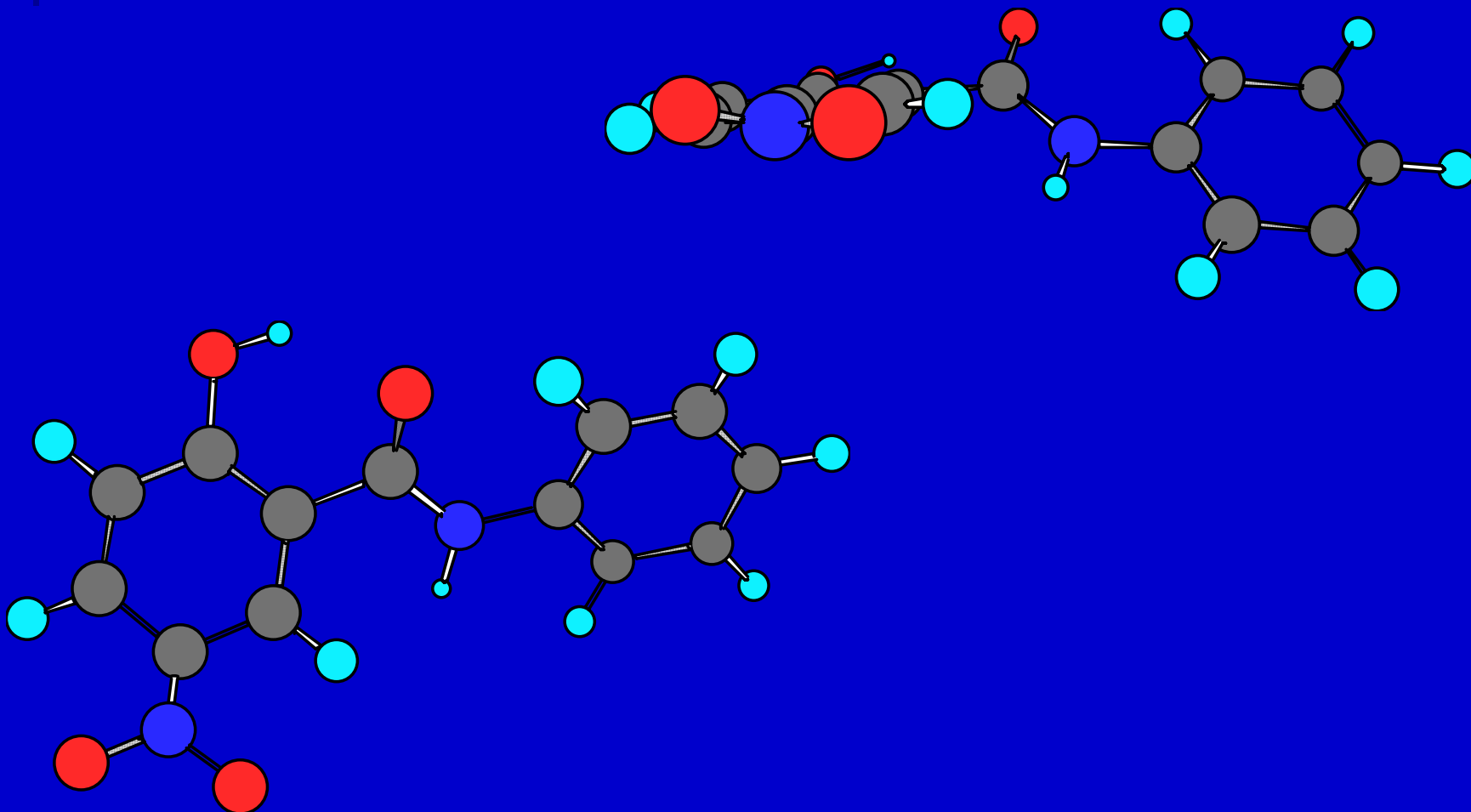
# (3) 举例: 量子化学参数预测苯甲酰苯胺类化合物 对大型蚤的毒性lgEC<sub>50</sub>

## 14个苯甲酰苯胺类化合物

序号	结构式	序号	结构式	序号	结构式
1		6		11	
2		7		12	
3		8		13	
4		9		14	
5		10			



## 5'-NO<sub>2</sub>-R<sub>1</sub>-R<sub>2</sub>分子的空间结构




$$-\log LC_{50}=9.4718-6.6019Q_{N-H}-2.9534Q_{N-O}-5.8379P_H$$

$$R^2_{adj}=0.8269 \quad SE=0.1132 \quad F=21.6938 \quad p=0.0001$$

## 量子化学参数对毒性机制的探讨:

1.  $Q_{N-H}$  与  $Q_{N-O}$  分别代表酰胺基中N原子和硝基中O原子与外来H原子形成氢键的倾向。当  $Q_N$  (  $Q_O$  ) 越负, 或  $Q_H$  (  $Q_N$  ) 越低, 二者的代数和  $Q_{N-H}$  与  $Q_{N-O}$  就越负, 分子与生物靶分子中的H原子形成氢键的作用越强, 因而化合物毒性增强
2.  $P_H$  描述化合物分子的电负性。当  $P_H$  越小, 化合物越难与溶剂水分子间形成氢键, 化合物的毒性因而增强

三个参数揭示了该类苯甲酰苯胺化合物与生物靶分子形成氢键而致毒的机理



# 小 结

---

- 结构-活性关系研究是有机污染化学和生态毒理学中的一个前沿领域
- 结构活性关系技术作为现代化学、生命科学、数学和计算机科学的交叉学科生长点，依托先进的信息处理技术，为化学品的环境风险评价开创了一条简便、实用的途径
- 在结构-活性关系研究中，分子结构的表征是关键步骤之一。理想的分子结构参数能够准确、完整地表达分子的结构信息，使QSAR模型兼具高度的预测能力和明确的物理意义，有助于解释有机污染物的生物活性作用机制，有助于揭示污染物在环境中的环境行为规律。