ОПИСАНИЕ АМБРОВЫХ ОДОРАНТОВ НА ОСНОВЕ ТРЕХМЕРНЫХ ДЕСКРИПТОРОВ МОЛЕКУЛЯРНЫХ ПОВЕРХНОСТЕЙ В ЗАДАЧЕ «СТРУКТУРА-СВОЙСТВО».

[2]Захаров А.М., [1]Кумсков М.И., [1]Свитанько И.В., [1]Пономарева Л.А., [2]Григорьева С.С., [2]Чичуа В.Т.

[1]Институт Органической Химии им.Н.Д.Зелинского РАН, г.Москва

[2]МГУ им. М.В.Ломоносова, мехмат, г.Москва

[2]Захаров А.М.
[1]Кумсков М.И
[1]Свитанько И.В
[1]Пономарева Л.А.
[2]Григорьевас С.С.
[2]Чичуа В.Т.
[1]ИОХ РАН
[2]Мехмат МГУ

Поставлена задача идентификации пространственной структуры «молекулярного ключа», соответствующего неизвестному «белковому замку», запускающему процесс проявления биологической активности на примере веществ, проявляющих амбровый запах. Описана методика поиска адекватных описаний структур молекулярных графов для задачи "структура-свойство" (QSAR - "Quantitative Structure-Activity Relationship"). Выбор описания молекул в виде вектора признаков заранее не задан. Правильное определение признаков под заданное биологическое свойство является ключевым моментом QSAR-моделирования. Адекватный выбор признаков позволяет содержательно проводить оценку и интерпретацию QSAR-моделей, в том числе формированием геометрической модели активного ценрта.

В работе предлагается использование нового подхода к проведению QSAR моделирования, который основан на аналогии между изображениями и молекулярными графами [1]. Обучающая выборка представляется молекулярными графами, имеющими пространственную укладку. Признаковые пространства, в которых проводится описание молекул обучающей выборки формируются автоматически с последовательным усложнением детализации описания и уровня представления молекул.

Особенностьюпредлагаемого подхода является представление молекулы не в виде конфигурации атомов, а в виде молекулярной поверхности, которая «раскрашивается» локальным физико-химическим свойством (ЛФХС) – зарядом, липофильностью, способностью отдавать или принимать электрон (донорно-акцепторные факторы). В результате «окраски» на молекулярной поверхности вычисляются экстремумы ЛФХС, которые называются «элементами описания» (ЭО) или особыми точками (ОТ). В результате объектом, подлежащим анализу и классификации, становится маркированный граф, вершины которого располагаются в ОТ на молекулярной поверхности.

Схема вычислений

1. Расчет пространственной структуры и заряда для 100 молекул, проявляющих амбровый запах и имеющих жесткую структуру.
2. Расчет молекулярных поверхностей молекул, распределения заряда и локальных экстремумов заряда на поверхности [4]
3. Расчет (идентификация) «точек описания» поверхности (или особых точек - ОТ) [5] (см.рис.1).
4. Классификации «ОТ по заряду» на молекулярной поверхности.
5. Формирование элементов описания (ЭО) в виде «пар, троек и четверок ОТ» с использованием «жесткой (четкой) классификации» расстояний между ОТ.
6. Расчет матрицы «молекула-ЭО» с различными вариантами класификации ОТ и дискретизации растояния между парами ОТ
8. Формирование и анализ прогностической устойчивости модели «структура-свойство» в заданном классификаторе.[6]

Предлагаемый подход можно отнести к 3D-QSAR методам, использующим информацию о пространственных конфигурациях молекул для построения зависимостей «структура-свойство».

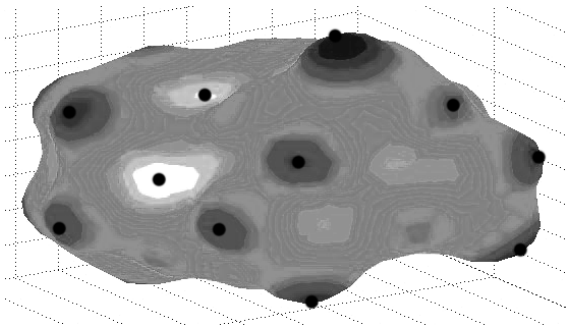

Рис.1. Молекулярная поверхность с выделенными особыми точками.

Наиболее разработанным и широко используемым 3D-QSAR методом описания молекул является Comparative Molecular Field Analysis (CoMFA) [2,3]. Его основные недостатки хорошо известны: (1) Необходимость нормализованного расположения молекулы - все молекулы обучающей выборки должны быть размещены в параллелопипеде и выровнены друг относительно друга. (2) Трудность обработки гибких молекул. CoMFA метод хорошо зарекомендовал себя для анализа жестких структур одного химического ряда. В предлагаемом подходе указанные недостатки преодолеваются, поскольку формы молекул описываются геометрическими фрагментами (ЭО) разных уровней детализации (двойки, тройки и четверки).

В рамках решения поставленной задачи уточняется методика и разрабатываются алгоритмы расчета и классификации «ОТ по заряду» на молекулярной поверхности, а также методика и алгоритмы расчета и классификации «геометрических ОТ» на молекулярной поверхности в системе MATLAB 7.0.

Работа поддержана РФФИ по гранту 06-01-00777.

Литература

1. *Sergunin S.Yu., Kvashnin K.M., Kumskov M.I.* Using Image Representation Families for Recognition Based on Symbol Marking of Singular Points // Pattern Recognition and Image Analysis, Vol. 13. №3. 2003. P.531–538.
2. *Cramer R.D.,III, Patterson D.E., Bunce J.D*. Comparative Molecular Field Analysis (CoMFA). 1.Effect of Shape on Binding of Steroids to Carrier Proteins. // J.Amer.Chem. Soc.1988. V.110. n.18. P.5959-5967.
3. *H. Kubinyi,* "QSAR in Drug Design. Theory, Methods and Applications" Escom, Leiden, 1993.
4. *Svitanko I.V., Kumskov M.I., Zyryanov I.L., SuslovI.A.* A Method for Describing the Molecular Electrostatic Potential for Determining the Structure-Activity Relationship // Mendeleev Communications, Volume 4(1994), Number 5, p.161-162.
5. *Захаров А.М., Свитанько И.В., Григорьева С.С., Чичуа В.Т*. Поиск особых точек на молекулярных поверхностях с использованием нечеткого кластер-анализа / Математические методы распознавания образов (ММРО-12-2005). С.333-335.
6. . *Zhuravlev Yu.I., Vinogradov A.P., Voronchikhin V.V., Ryazanov V.V., Senko O.V.* LOREG: Program System of Pattern Recognition and Data Analysis//Pattern Recognition and Image Understanding, Infix. 1999. P.244-250.