

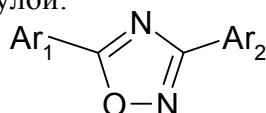
МАТЕМАТИЧЕСКИЕ АСПЕКТЫ ПОСТРОЕНИЯ QSAR МОДЕЛЕЙ.

Цыганкова И.Г.

Институт теоретической и экспериментальной биофизики РАН
142290 г. Пушкино Московской обл. tsygan@iteb.ru

Корреляционные соотношения структура – активность (Quantitative Structure-Activity Relationship, QSAR) дают возможность рассчитывать биологическую активность соединения по его молекулярной структуре без детального знания молекулярных механизмов активности. Для QSAR моделей существенным является предположение, что соединения с похожей молекулярной структурой ведут себя одинаково в одинаковых условиях. Степень похожести молекулярных структур определяется при помощи разнообразных молекулярных дескрипторов (см. например Todeschini R., Consonni V. *Molecular Descriptors for Chemoinformatics*; Wiley-VCH. 2009.).

При построении QSAR модели апоптоз-индуцирующей активности для 41 соединения, которые представляют собой производные 3,5 – диарил – 1,2,4 – оксадиазола с общей формулой:



были использованы фрагментные молекулярные дескрипторы, значения которых равны числу атомов или групп атомов определенного типа в структуре соединения. Корреляционное соотношение для расчета активности P было выбрано простой математической формы - в виде суммы вкладов:

$$P = \sum_t n_t P_t + \sum_{t,m,r} k_{t,m-r} P_{t,m-r}$$

Здесь n_t – число фрагментов, связанных с атомами определённого типа t ; $k_{t,m-r}$ – число пар фрагментов, которым соответствуют атомы типа t и m , разделенные r химическими связями. Фрагментные вклады P_t , $P_{t,m-r}$ - параметры корреляционного уравнения - должны быть определены по экспериментальным значениям активности. Поскольку число параметров корреляционного соотношения $n_t + k_{t,m-r}$ оказалось значительно больше числа экспериментальных точек, то для получения эффективных моделей было необходимо провести отбор переменных. На этом этапе было исследовано влияние критериев отбора переменных на качество получаемых моделей, проведено сравнение различных алгоритмов отбора. Разработанный ранее алгоритм с итерационным использованием пошаговой регрессии (Цыганкова И.Г., Женодарова С.М. // ЖОХ. 2008. Т. 78. № 9. С.1529) давал модели, не уступающие по точности моделям, полученным при помощи других алгоритмов, хотя работал быстрее. Использование в качестве критерия отбора коэффициента корреляции рассчитанных и экспериментально измеренных значений активности позволило получить модели, в которых было меньше «выбросов», т.е. единичных больших отклонений оценки активности от эксперимента. В целом, линейные QSAR модели с фрагментными молекулярными дескрипторами допускают наглядную интерпретацию, и в результате удалось предложить модификации молекулярной структуры для получения более активных индукторов апоптоза, что может оказаться полезным в поиске потенциальных противоопухолевых препаратов.