

# ПОИСК НОВЫХ СВОЙСТВ ИЗВЕСТНЫХ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ С ПРИВЛЕЧЕНИЕМ МЕТОДОВ КВАНТОВОЙ ХИМИИ, МОЛЕКУЛЯРНОГО МОДЕЛИРОВАНИЯ И МНОГОМЕРНОЙ СТАТИСТИКИ

*Погребняк Л.В., Вотинцев Н.П., Погребняк А.В.*

Пятигорская государственная фармацевтическая академия, г.Пятигорск

[pspa2007@yandex.ru](mailto:pspa2007@yandex.ru)

*Методом многомерной статистики с использованием электронных и геометрических дескрипторов молекулярной структуры осуществлен прогноз новых видов биологической активности для шести известных лекарственных препаратов: дибунола, пропосола, клопидогреля, триметоприма, кислоты ацетилсалициловой и метилурацила. Результаты прогноза подтверждены фармакологическими экспериментами и могут служить основанием для изменения показаний к применению указанных препаратов.*

Эволюция методов поиска новых эффективных лекарственных препаратов в последние 20 лет претерпела существенные изменения, главным образом – за счет совершенствования теоретических приемов, описывающих свойства молекул на уровне атомов и электронов. В первую очередь это связано с двумя факторами: совершенствованием математического аппарата квантовой химии в части описания свойств молекул малого размера и возрастающей доступностью соответствующих вычислительных ресурсов. Так, удовлетворительное описание распределения электронной плотности молекулы, состоящей из 40-60 атомов сегодня можно получить в течение нескольких часов.

Поиск новых областей использования уже применяемых лекарственных препаратов имеет ряд преимуществ перед внедрением новых органических соединений, главные из которых:

- известна острая и хроническая токсичность, эмбриотоксичность, отрицательные побочные эффекты и пути их снижения;
- определены физико-химические свойства, определяющие вид и оптимальный состав лекарственных форм;
- изучены метаболические превращения и пути элиминации из организма;
- уточнена область амбулаторного и/или клинического применения, в т.ч. зависимость дозировок от пола, возраста и других факторов;
- у врачей и пациентов отсутствует психологический барьер, обычно сопровождающий внедрение нового препарата.

Все вышеперечисленное делает область изучения новых фармакологических свойств известных препаратов чрезвычайно привлекательным и экономически выгодным, особенно в условиях недостаточного финансирования программ по созданию новых лекарств.

Безусловно, речь идет не о подмене поиска новых биологически активных веществ ретроспективным исследованием «старых» препаратов, но о разумном балансе двух указанных путей, зависящем от текущей конъюнктуры.

Нижеприведенные экспериментальные примеры демонстрируют только начало данного пути. По мере получения новых данных мы рассчитываем продолжить серию подобных публикаций.

В современной медицинской химии существуют три способа поиска потенциальных лекарств: тотальный скрининг (новое вещество испытывается на большинство видов биологического действия); метод модификации (химическое изменение молекул с известной активностью) и виртуальный скрининг (для прогноза используется компьютерный алгоритм). Последний метод является самым новым и, соответственно, имеет наибольший потенциал развития.

Очевидное соответствие между тонким строением молекулы (в нашем случае - описывающими эту структуру геометрическими и электронными дескрипторами) и ее

биологическими свойствами позволяет использовать численные характеристики, получаемые в результате квантово-химических и иных расчетов, в качестве инструмента для прогнозирования новых свойств.

Ранее [1,2] нами был предложен метод прогноза биологической активности, основанный на сопоставлении совокупности электронных и геометрических дескрипторов с биологической активностью данной молекулы. Главная идея заключалась в том, что если нами рассчитана подобная совокупность для молекул «А» и «В» с заведомо известным биологическим действием, то логично предположить, что расчет этой же совокупности дескрипторов для новой молекулы «С» позволит определить степень близости (расстояние) последней до молекул «А» и «В» в n-мерном пространстве дескрипторов (где n – их общее количество). Оценивая эти расстояния можно полагать, что если «цифровой отпечаток» молекулы «С» ближе к молекуле «А», то и вероятность проявления молекулой «С» «А»-подобного действия существенно повышается.

Зачем нужны сложные расчеты и громоздкие алгоритмы прогнозирования, если можно провести высокоэффективный фармакологический скрининг? Дело в том, что существующий на сегодняшний день подход к созданию лекарств оборачивается значительными затратами: примерно одно соединение из  $10^5$  изученных веществ может рассматриваться в качестве кандидата в лекарства; весь процесс разработки от синтеза до реализации в аптеке длится 8-15 лет; стоимость нового лекарственного препарата составляет не менее  $10^9$  рублей. На конец 2006 года описано  $2,3 \cdot 10^7$  химических веществ, из них коммерчески доступны около  $6,5 \cdot 10^6$ . В качестве лекарственных препаратов в настоящее время зарегистрировано около 7 тысяч (данные ВОЗ 2003 г.). Включение химического соединения в официальный перечень не означает, что ВОЗ рекомендует данную субстанцию к применению в области медицины. В силу последнего обстоятельства на рынке лекарств активно используется не более 1700-2000 веществ.

Последние и явились объектом нашего исследования.

Информация о химической структуре и фармакологической активности 1072 современных лекарственных веществ (далее – лекарств) заимствована нами из известного справочника М.Д. Машковского (XV изд. и более ранних), уточнение структур проводилось по оригинальным источникам.

При отборе соединений, включенных в обучающую базу данных, нами учитывалось, что время квантово-химического расчета растет в геометрической прогрессии с увеличением числа атомов в системе, а по достижении ими определенного количества самосогласованное решение попросту недостижимо (для обычно используемой двойной точности расчета, в программе МОРАС она задается ключевым словом PRECISE).

В этой связи в исходный набор были включены соединения, содержащие в молекуле до 100 неводородных атомов, или имеющие общее количество атомов не более 300. Самой «большой» структурой т.о. является ангиотензинамид ( $M_w=1031$ , 74 тяжелых атома); на нем с большим отрывом следует амфотерицин ( $M_w=924$ ). Самой «маленькой» молекулой является циклопропан ( $M_w=42$ ), и поэтому отсечение с данными границами не приводит ни к каким ограничениям. Вне расчета остались структуры, подобные гепарину, белки и т.п., при этом прогностическая емкость обучающей выборки остается на одинаковом уровне. Здесь уместно вспомнить, что однородность массива и близость к нему пробного объекта – важный критерий объяснимости и воспроизводимости результата любого прогнозирования.

Оптимизация геометрии молекул лекарственных веществ производилась при помощи итерационной процедуры поиска локального (в случае удачи - глобального) минимума на поверхности потенциальной энергии.

Для первичной оптимизации геометрии использовано силовое поле MMP2 [3]. В силу своей экономичности с точки зрения затрат машинного времени данный метод использовался в основном для первичного анализа конформационной заселенности. В качестве стартовой для методов МО была взята геометрия, полученная методом MMP2.

Метод молекулярной механики при оптимизации дает геометрию, более близкую к экспериментальной (в сравнении с полуэмпирическими и *ab initio* (в минимальном базисе STO-3G) моделями. В связи с этим, в случае невозможности проведения оптимизации геометрии, для расчета физико-химических дескрипторов методами *ab initio* нами была использована геометрия полученная методом ММР2.

Полученная в результате оптимизации методом AM1 [4] геометрия использовалась далее для расчета физико-химических дескрипторов полуэмпирическими методами (кроме PM3 [5], оптимизация геометрии в котором производилась независимо), а также для расчета дескрипторов молекулярной структуры (AM1).

При расчете и отборе дескрипторов, характеризующих наиболее значимые свойства молекул, присутствовали несколько важных особенностей: оптимизация геометрии молекул методами AM1 и 6-31G производилась как в газовой фазе, так и с учетом эффектов гидратации при физиологической температуре, для чего использовалась модель поляризуемого континуума (в иностранной литературе - PCM) [6].

Процедура поиска минимально-энергетической конформации проводилась методом молекулярной динамики в модификации Ланжевена (Langevin), позволяющим быстро, но достаточно точно учитывать влияние молекул растворителя [7].

Основными физико-химическими дескрипторами являются (в алфавитном порядке): адиабатический потенциал ионизации с уч. гидр. (AM1); адиабатическое сродство к электрону с уч. гидр. (AM1);  $\alpha$ -поляризуемость (6-31G\*\*);  $\beta$ -гиперполяризуемость (6-31G\*\*);  $\gamma$ -гиперполяризуемость (6-31G\*\*); дипольный момент; длина молекулы с уч. гидр. (AM1); молекулярная масса; объем ван-дер-ваальсовой модели молекулы (6-31G\*\*); площадь ван-дер-ваальсовой модели молекулы (6-31G\*\*); потенциал ионизации в основном состоянии (AM1); стерическая энергия (MM2); теплота образования аниона с уч. гидр. (AM1); теплота образования катиона с уч. гидр. (AM1); ширина молекулы с уч. гидр. (AM1); энергия ВЗМО (6-31G\*\*); энергия ВЗМО с уч. гидр. (AM1); энергия гидратации (6-31G\*\*, PCM); энергия гидратации (AM1, PCM); энергия НСМО (6-31G\*\*); энергия НСМО с уч. гидр. (AM1); энергия образования полости в воде (6-31G\*\*); энергия связей (MM2); энергия электростатических взаимодействий (MM2); энергия электростатического взаимодействия с гидратной оболочкой (6-31G\*\*).

Теплоты гидратации рассчитывали по уравнению (1):

$$H_{\text{hydr}} = H_{\text{aq}} - H_{\text{vac}}, (1)$$

где  $H_{\text{hydr}}$  – энергия гидратации молекулы;  $H_{\text{aq}}$  – теплота образования молекулы с учетом эффекта гидратации;  $H_{\text{vac}}$  – теплота образования без учета гидратации.

Для расчета адиабатического потенциала ионизации ( $I_p$ ) использовались данные о теплотах образования соответствующего катиона и молекулы в основном состоянии (во всех случаях учитывались эффекты гидратации). Использовалось уравнение (2):

$$I_p = H_{f+} - H_{f0}, (2)$$

где  $I_p$  – потенциал ионизации;  $H_{f+}$  – теплота образования соответствующего катиона (с учетом эффекта гидратации);  $H_{f0}$  – теплота образования молекулы в основном состоянии (с учетом эффекта гидратации).

Аналогично производился расчет (3) адиабатического сродства к электрону ( $A_e$ ):

$$A_e = H_{f-} - H_{f0}, (3)$$

где  $A_e$  – сродство к электрону (с учетом эффекта гидратации);  $H_{f-}$  – теплота образования соответствующего аниона;  $H_{f0}$  – теплота образования молекулы препарата (с учетом эффекта гидратации).

В схему прогнозирования были включены также дескрипторы следующих групп (в круглых скобках указано общее количество): аддитивные дескрипторы (11), топологические дескрипторы (32), индексы обхода (10), геометрические дескрипторы (6), дескрипторы 3D-MoRSE (10), дескрипторы WHIM (24), дескрипторы GETAWAY (3). Во избежание лишних математических трудностей из массива данных нами были исключены дескрипторы,

имеющие одинаковое значение для 5 и более молекул. Общее число топологических дескрипторов на одну молекулу – 96.

В отличие от большинства молекулярные дескрипторы (далее - МД), получаемые квантово-химическим путем, четче детерминированы и имеют более ясный физический смысл. Последнее позволяет описывать и трактовать результаты прогнозирования в терминах понятных химику и фармакологу.

При разработке метода прогнозирования фармакологического действия общий ход наших рассуждений выглядел так: каждую молекулу (после соответствующих расчетов МД) можно представить точкой в n-мерном пространстве, где n – количество МД. Зная такие «координаты» для молекул базового набора (обучающей выборки) и рассчитав аналогичный набор дескрипторов для нового соединения можно, на основании оценки расстояния между данным соединением и молекулами тестового массива, сделать заключение о возможной активности испытуемого вещества.

Существует два рациональных подхода к достижению указанной цели: первый представляет собой известную процедуру поиска расстояния между точками или массивами в n-мерном пространстве (какая-либо разновидность кластерного анализа, например метод k-средних или сравнительно новая модификация метода в редакции Worth'a и Cronin'a – названная ими «встроенный кластер»); второй, более гибкий и точный, является комбинированным и предлагается нами в оригинальной модификации.

На наш взгляд весьма перспективным методом с точки зрения изучения количественной связи «структура-активность» является метод потенциальных функций (применение см., например, [8]). Для проведения процедуры прогнозирования нами выполнена модификация метода потенциальных функций. Во избежание разночтений отметим, что метод потенциальных функций по некоторым источникам является лишь иным названием метода восстановления функции распределения по Парзену-Надараю.

В качестве меры фармакологического действия нами использовался логарифм обратной среднетерапевтической дозы для взрослого человека (для каждого лекарственного препарата базового набора эта величина заведомо известна).

Численные значения фармакологических активностей аппроксимировались в пространстве дескрипторов с помощью потенциальных функций вида (4-6):

$$P_i = \sum_{j=1}^m k_j f_{ij} \quad (4)$$

$$f_{ij} = f(r_{ij}) \quad (5)$$

$$r_{ij} = \sqrt{\sum_{c=1}^v W_c^2 (X_{ic} - X_{jc})^2} \quad (6)$$

где  $P_i$  - значение потенциальной функции в точке i-го препарата;  $k_j$  - нормировочный множитель для j-й функции распределения; m - количество функций распределения;  $r_{ij}$  – матрица дистанций;  $f_{ij}$  – матрица значений функции распределения;  $W_c$  – вектор весов дескрипторов;  $X_{ic}$  и  $X_{jc}$  – матрица значений дескрипторов; f – функция распределения; i – номер препарата в обучающей выборке; j – номер препарата в активной подгруппе обучающей выборки; c – номер дескриптора; v – количество дескрипторов.

Следует отметить, что функции распределения центрированы на точках соответствующих самым активным (для данного вида биологической активности) препаратам обучающей выборки.

В качестве функции распределения нами была использована функция Гаусса, определенная на всем пространстве МД и принимающая значения от 0 до 1. Функция имеет параметр сглаживания (7), который может варьироваться для оптимизации модели:

$$f(r) = e^{\left(-\frac{r^2}{\sigma^2}\right)} \quad (7)$$

где  $r$  – радиальная координата;  $\sigma$  – параметр сглаживания (характеризует относительную степень разброса структурных требований для данного вида активности).

Дополнительную гибкость модели придает возможность использования различных видов функций распределения с варьированием параметров в них. Пример функции двух аргументов, используемой в качестве потенциальной функции для прогнозирования фармакологической активности, приведен на рисунке 1.

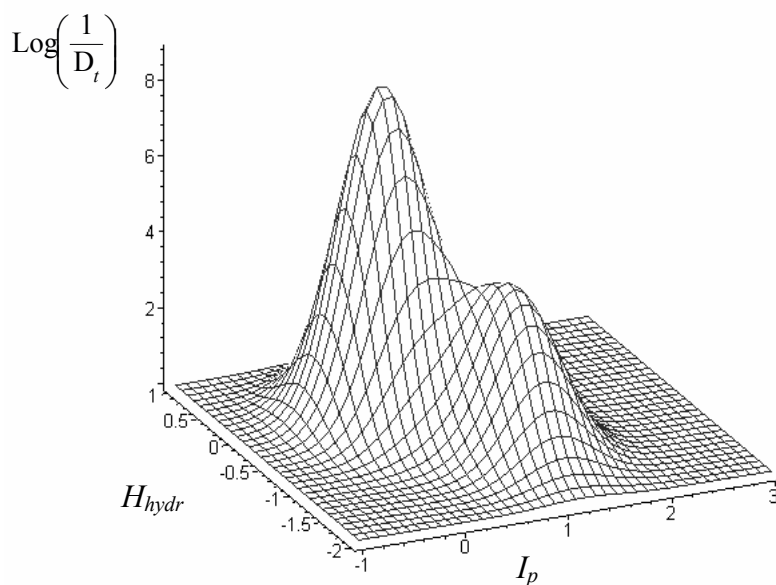


Рисунок 1. Общий вид потенциальной функции для прогнозирования фармакологической активности

Функция аппроксимирует десятичный логарифм обратной среднетерапевтической разовой дозы для взрослого человека. В реальности потенциальная функция всех видов активности, конечно, является  $n$ -мерной, где  $n$  – количество дескрипторов.

Потенциальная функция нормализована для оптимального воспроизведения количественных значений фармакологической активности обучающей выборки. В связи с этим прогнозирование фармакологической активности производится путем расчета значения потенциальной функции в точке метрического пространства дескрипторов,

Для оценки качества модели была использована стандартная ошибка прогноза ( $S_{PRESS}$ ), определяемая методом скользящего контроля. Модель имеет два параметра:  $\sigma$  - сглаживания потенциальной функции и  $m$  - статистического веса молекулярных дескрипторов ( $w_j, j=1, \dots, m$ ).

Так как ошибка прогноза  $S_{PRESS}$  возрастает и убывает вместе с суммой квадратов невязок прогноза ( $PRESS_{cv}$ ), то оптимизация параметров была направлена на её минимизацию (8):

$$PRESS_{cv} = \sum_{i=1}^n \left( \left( \sum_{j=1}^m k_j f_{ij} \right) - A_i \right)^2, \quad j \neq i \quad (8)$$

где  $k_j$  - нормировочный множитель для  $j$ -й функции распределения;  $m$  - количество функций распределения;  $n$  - количество препаратов в обучающей выборке;  $f_{ij}$  - матрица значений функции распределения;  $i$  - номер препарата в обучающей выборке;  $j$  - номер препарата в активной подгруппе обучающей выборки;  $A_i$  - фармакологическая активность  $i$ -го препарата из обучающей выборки.

На рисунке 2 приведен пример поверхности значений PRESScv в пространстве параметров модели (в данн. сл., статистических весов двух дескрипторов).

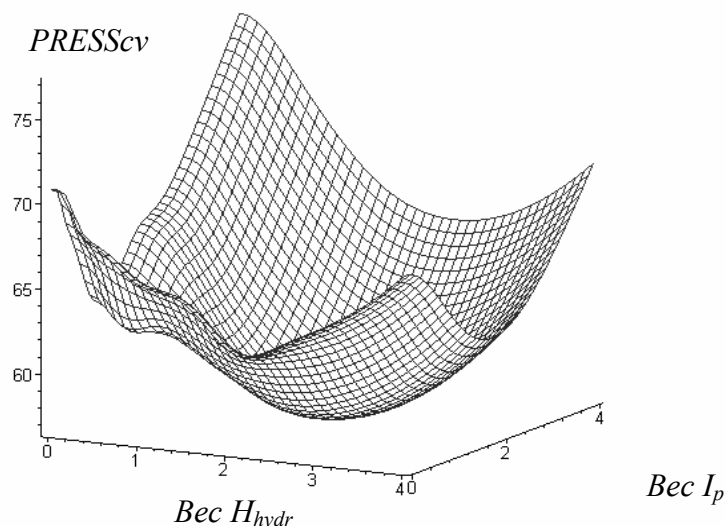


Рисунок 2. Поверхность PRESScv в пространстве параметров модели (веса дескрипторов  $H_{hydr}$  и  $I_p$ )

Поверхность имеет выраженный минимум, соответствующий оптимальной величине параметров. Поиск точки минимума в пространстве параметров проводился при помощи стандартной итерационной процедуры, основанной на методе Ньютона (метод касательных).

Проблема взаимной корреляции МД, включенных нами в основу алгоритма, исключается за счет того, что данная схема является геометрическим методом распознавания [9].

Образец конечного результата работы прогностического алгоритма представлен в таблице 1. Нулевой элемент матрицы соответствует тождеству совокупности МД для одной и той же молекулы. Ненулевые элементы соответствуют относительному расстоянию от одной молекулы (соотв. совокупности МД) до другой.

Таблица 1. Пример результата работы прогностического алгоритма

Shortname	tirofiba	valacicl	vesnarin	vigabatr	yohimbin	zuclopen
yohimbin	5.402	3.704	2.677	6.866	0	3.103
trimetop	6.068	3.385	3.734	5.434	1.902	4.001
lobelinu	5.241	3.456	2.915	6.894	2.024	2.932
cholecal	4.691	4.163	2.277	7.830	2.099	2.719
ajmaline	6.465	3.809	3.737	6.061	2.213	3.685
pentoxyv	4.991	3.199	3.491	6.577	2.225	4.204
psorcuta	5.247	3.625	2.872	7.685	2.304	3.700
trifluop	5.782	4.392	2.638	8.101	2.317	2.742
ciproflo	5.072	3.126	3.220	6.029	2.317	3.962
oxprenol	6.168	3.465	4.238	5.349	2.351	4.492
hydroxyz	6.072	3.919	3.490	6.837	2.358	3.640
indometa	4.594	4.120	2.237	7.118	2.368	3.150
butiroxa	4.627	3.400	2.721	6.661	2.372	3.219
butamira	5.268	3.258	3.883	5.595	2.377	4.694
trihexyp	6.453	3.867	4.313	5.946	2.427	4.414
cyclodol	6.326	3.552	4.175	5.913	2.432	4.360

Для практической апробации нами были выбраны (случайным образом) три пары препаратов, входящие в «первую сотню» пар, продемонстрировавших минимальные взаимные расстояния в пространстве молекулярных дескрипторов: дибунол-пропофол; клопидогрель-триметоприм и метилурацил- ацетилсалициловая кислота.

Далее было проведено сравнительное изучение противоэкссудативной, антиоксидантной и мембраностабилизирующей активности препаратов бутилокситолуол (дибунол) и пропофол в условиях экспериментальной патологии. В качестве экссудативного воспаления использовали модель острого каррагенинового отека задней лапы крыс. В качестве индуктора процессов перекисного окисления липидов и дестабилизации мембран эритроцитов крыс был избран клеточный яд тетрахлорметан. Исследования проводились по заказу Пятигорской государственной фармацевтической академии. Экспериментальная работа проводилась на белых половозрелых крысах. При сравнительной оценке действия бутилокситолуола и пропофола на острое воспаление, установлено следующее: бутилокситолуол и пропофол проявили выраженный противоэкссудативный эффект. При этом на мембраны эритроцитов было оказано выраженное стабилизирующее и антиоксидантное действие. Эффективность обоих препаратов, не имела достоверно значимых различий.

Проведено сравнительное исследование клопидогреля и триметоприма на гемореологию и гемограмму животных *in vivo* в условиях экспериментальной нормы. Исследование проводилось зарегистрированными методами: оценка параклинических показателей крови и определение времени начала и конца свертывания крови после курсового (7 суток) введения препаратов. В эксперименте были задействованы животные: белые крысы линии Wistar массой  $220 \pm 10$  г. В результате эксперимента установлено: клопидогрель и триметоприм проявили выраженный ингибирующий эффект на состояние периферической крови, угнетая образование и функционирование форменных элементов крови: лейкоцитов, эритроцитов, тромбоцитов. Клопидогрель ингибирует преимущественно функцию и количество тромбоцитов. Триметоприм преимущественно ингибирует количество и профиль лейкоцитов. Угнетение количества и функции форменных элементов крови показало выраженную эффективность обоих препаратов.

Проведено сравнительное исследование аспирина и метилурацила на асептическое воспаление в эксперименте *in vivo*. Исследование проводилось зарегистрированным методом: асептическое воспаление cotton pellet (7 суток). В эксперименте были задействованы животные: белые крысы линии Wistar массой  $220 \pm 10$  г. В результате эксперимента установлено: аспирин и метилурацил проявили выраженный противовоспалительный эффект. Аспирин ингибирует преимущественно стадию экссудативного воспаления. Изучение антипролиферативных свойств аспирина в сравнении с метилурацилом показало выраженную эффективность обоих препаратов.

Более подробное описание методик проведения фармакологического экспериментов можно получить непосредственно у его авторов: д.м.н. М.Н. Ивашева и д.б.н. А.В.Сергиенко (Пятигорская государственная фармацевтическая академия) по адресу, указанному в начале сообщения.

Таким образом, все фармакологические эксперименты прямо или косвенно подтвердили эффективность использования алгоритма прогнозирования новых видов действия известных лекарственных препаратов, основанного на анализе пространства молекулярных дескрипторов.

## Литература

1. *Погребняк А.В.* Программа для ЭВМ «Mspace». Свидетельство об официальной регистрации №2003612547, 21.11.2003, РОСПАТЕНТ.
2. *Погребняк А.В., Глушко А.А., Оганесян Э.Т.* Matrix – новый алгоритм прогнозирования биологического действия органических молекул основанный на многомерном анализе физико-химических дескрипторов современных лекарственных препаратов. I. Общие принципы. // Журнал органической химии.- 2002.- Т.38.- Вып.11.- С.1618-1629.
3. *Berkert U., Allinger N.L.* Molecular Mechanics.- American Chemical Society Monograph 177, Washington, D.C., 1982.
4. *Dewar M.J.S., Zoebisch E.G., Stewart J.J.P.* AM1: A new general purpose quantum mechanical molecular model // J.Am.Chem.Soc.- 1985.- Vol.107.- P.3902-3909.
5. *Orozco M., Bachs M., Luque F.J.* Development of Optimized MST/SCRF Methods for Semiempirical Calculations. The MNDO and PM3 Hamiltonians // J.Comp.Chem.- 1995.- №16.- P.563-585.
6. *Tomasi J., Persico M.* Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent // Chem.Rev.- 1994.- Vol.94.- P.2027-2094.
7. *Allen M.P., Tildesley D.J.* Computer Simulation of Liquids.- Clarendon Press: Oxford, 1987.- Ch.9.- P.261.
8. *Тетко И.В., Танчук В.Ю., Васильев С.А., Хиля В.П., Пода Г.И., Луйк А.И.* Эволюционное программирование для выявления закономерностей "структура-активность" в ряду производных 3-феноксихромона и 3-фенокси-4-гидроксикумарина // Биоорганическая химия.- 1995.- Т.21.- №10.- С. 809-815.
9. *Горелик А.Л., Скрипкин В.А.* Методы распознавания.- М.: Высшая школа, 1984. - 208с.