

Дескрипторы молекулярной структуры в компьютерном дизайне биологически активных веществ

О.А.Раевский

*Институт физиологически активных веществ Российской академии наук
142432 Черногловка Московской области, факс (095)913–2113*

Проанализировано современное состояние проблемы описания дескрипторами молекулярной структуры в компьютерном дизайне биологически активных веществ. Отмечено, что информационное содержание дескрипторов повышается в последовательности: дескрипторы элементного уровня – дескрипторы структурной формулы – дескрипторы электронной структуры – дескрипторы молекулярной формы – дескрипторы межмолекулярных взаимодействий. При этом последующий класс дескрипторов обычно включает информацию, содержащуюся в дескрипторах предыдущего уровня. Подчеркивается практическая невозможность детального описания всех особенностей молекулярной структуры с помощью какого-либо одного дескриптора. Рекомендуются оптимизация числа используемых дескрипторов на основе соответствующих статистических процедур и характеристик построенных на этих дескрипторах моделей структура – свойство.

Библиография — 371 ссылка.

Оглавление

I. Введение	555
II. Классификация дескрипторов	556
III. Дескрипторы элементного уровня	556
IV. Дескрипторы структурной формулы	557
V. Дескрипторы электронной структуры молекул	559
VI. Дескрипторы молекулярной формы	560
VII. Дескрипторы межмолекулярных взаимодействий	561
VIII. Индикаторные дескрипторы	566
IX. Современные компьютерные программы для расчетов дескрипторов	566
X. Необходимые качества дескрипторов	568
XI. Заключение	569
XII. Приложение	569

I. Введение

Эмпирические закономерности изменения свойств веществ с изменением их структуры, установленные к середине XIX в., легли в основу концепции взаимосвязи свойств и структуры химических соединений. Открытый Д.И.Менделеевым Периодический закон можно рассматривать в качестве научной основы концепции взаимосвязи строение – свойство, а предсказание физических свойств и существование неоткрытых новых элементов является первым примером устойчивой предсказательной модели этой взаимосвязи.¹

Систематические исследования взаимосвязи структуры органических соединений и такого их свойства, как биологи-

ческая активность, начались в конце прошлого века^{2–4} и к настоящему времени приобрели массовый характер. Признанным основоположником современного научного направления «количественная связь структура – активность» (quantitative structure – activity relationship, QSAR) является К.Ганч, опубликовавший вместе с Т.Фужитой в 1964 г. соотношение, нашедшее наиболее широкое применение в исследованиях связи между структурой и биологической активностью.² Однако, отдавая дань заслугам этого ученого, необходимо помнить и о многих его предшественниках. В частности, в публикации³ отмечено, что Н.В.Лазарев уже в 1944 г. предложил⁴ использование коэффициента распределения химических соединений в системе масло – вода для установления количественных соотношений структура – активность. Интенсивное развитие и внедрение в практику научных исследований вычислительной техники, в том числе и в рассматриваемой области, способствовало формированию междисциплинарного научного направления — компьютерного молекулярного дизайна биологически активных соединений. Многие концепции этого быстро развивающегося направления модифицируются и совершенствуются. Поэтому необходимо регулярно анализировать состояние новых исследований в данном направлении. Предыдущий обзор на эту тему⁵ печатался в данном журнале в 1988 г.

О.А.Раевский. Доктор химических наук, профессор, заведующий отделом компьютерного молекулярного дизайна, заместитель директора ИФАВ РАН, председатель Российской секции «QSAR and Modelling Society». Телефон: 524 – 5062, e-mail: raevsky@ipac.ac.ru, Internet e-mail: <http://www.ibmh.msk.su/qsar/molpro>

Область научных интересов: компьютерный молекулярный дизайн, методология установления взаимосвязей структура – свойство, компьютерное описание молекулярной структуры, молекулярное распознавание и межмолекулярные взаимодействия.

Дата поступления 16 ноября 1998 г.

В компьютерном дизайне биологически активных веществ с учетом концепции взаимосвязи структура–свойство можно выделить три главные стадии:

- 1) формирование обучающей выборки соединений с заданным свойством (активностью) или набором свойств;
- 2) описание молекулярной структуры соединений рассматриваемой выборки;
- 3) установление взаимосвязи структура–биологическая активность с последующим созданием устойчивой прогностической модели для предсказания свойств новых соединений.

Обзор посвящен анализу современных подходов к количественному описанию молекулярной структуры химических соединений в рамках компьютерного дизайна биологически активных веществ с помощью независимых переменных, характеризующих структурные особенности соединений или их частей (заместителей, фрагментов). Такие переменные называются дескрипторами.

Решающее влияние на развитие рассматриваемого научного направления оказывает прогресс вычислительной техники. Следует отметить, что в компьютерном молекулярном дизайне существует проблема обработки данных о молекулярной структуре в унифицированной форме с максимальным сохранением после ввода в компьютер их информативности.

Для изучения сложных природных процессов и явлений создаются их модели. Преимущества таких моделей по сравнению с оригиналом заключается, например, в большей доступности для изучения, возможности манипулирования и большего контроля, меньших затратах на исследования.⁶ Среди моделей можно выделить иконические (сходство по форме), аналоговые (сходство по форме и функциям) и абстрактные (символические, концептуальные), представляющие взаимосвязь с оригиналом с помощью символов.

Современные представления о молекулярной структуре, конечно, не являются исчерпывающими и в дальнейшем будут совершенствоваться с использованием различных моделей. Здесь уместно привести слова М.Дьюара: «Мы не знаем и, возможно, никогда не узнаем, чем действительно являются молекулы. Наше понимание молекул основано на моделях, которые воспроизводят их свойства достаточно хорошо, чтобы быть полезными» (цит. по лит.⁶).

Описание таких моделей, использующихся в химических исследованиях, приведено в публикациях^{6–8}. В данном обзоре также будут кратко рассмотрены дескрипторы молекулярной структуры с учетом указанной выше классификации моделей.

II. Классификация дескрипторов

Возможны различные способы классификации дескрипторов. Три типа классификации рассмотрены в статье⁹. В первом типе дескрипторы делятся на чистые (описывающие какой-либо один эффект межатомных взаимодействий) и композитные (описывающие два и более эффекта). Среди композитных дескрипторов, в свою очередь, выделяются уникompозитные (описывающие эффекты одного и того же типа) и мультикомпозитные (отображающие сумму эффектов различных типов).

Второй тип классификации основывается на способах оценки дескрипторов — экспериментальные или теоретические.

Третий тип классификации основывается на учете эффекта, который описывает данный дескриптор. Предлагается рассматривать три категории эффектов: электронные (электрические), стерические и межмолекулярные. Частным случаем использования этого типа классификации является классическое уравнение Ганча,^{2, 10–14} которое описывает биологическую активность как функцию констант Гаммета (электронные эффекты), Тафта (стерические эффекты) и



Рис. 1. Информационное содержание дескрипторов молекулярной структуры.

констант гидрофобности (транспортные эффекты за счет межмолекулярных взаимодействий).

В статье¹⁵ для наглядности описания структуры химических соединений использовались геометрические фигуры, соответствующие рассматриваемому уровню структуры: элементный уровень → двумерная структура → трехмерная структура → объемные свойства → стереодинамическая структура → стереоэлектронная структура → взаимодействия с окружением. При этом информационное содержание каждого уровня включает информационное содержание предыдущего уровня. Представляется, что подобный подход в несколько измененном виде можно использовать и в настоящем обзоре при обсуждении дескрипторов. На рис. 1 представлена схема нескольких уровней дескрипторов, на которой последующий уровень включает информационное содержание предыдущего уровня. При таком представлении предполагается, что дескрипторы структурной формулы несут всю информацию, содержащуюся в дескрипторах элементного уровня; дескрипторы электронной структуры включают всю информацию, содержащуюся в дескрипторах структурной формулы, а дескрипторы межмолекулярных взаимодействий содержат всю информацию всех предыдущих уровней. Иногда дескрипторы могут и не содержать информацию предыдущего уровня (например, такие дескрипторы межмолекулярных взаимодействий, как поляризуемость, гидрофобность, часто рассчитываются без учета трехмерной структуры), однако в целом указанный подход к классификации дескрипторов является наглядным, отражает степень их сложности и, следовательно, удобен при анализе. В данном обзоре рассмотрен каждый класс дескрипторов. Следует отметить, что дескриптор любого из указанных уровней может характеризовать как молекулу в целом, так и ее часть (фрагмент, скрин, функциональную группу, заместитель).

III. Дескрипторы элементного уровня

Брутто-формула содержит информацию о сортах атомов, входящих в молекулу, и числе атомов каждого сорта. Единственным свойством, которое может быть точно предсказано на основании брутто-формулы, является молекулярная масса (MW), которую можно рассматривать как дескриптор элементного уровня. В качестве других дескрипторов этого уровня иногда берут числа атомов какого либо сорта.¹⁶

Пример использования таких дескрипторов можно найти в Приложении (см. уравнение (П.1)).

Корректное исследование взаимосвязей структура – свойство должно проводиться для родственных соединений. По этой причине вариация структуры в пределах рассматриваемого ряда соединений подчас сводится к вариации заместителей. Следовательно, различные свойства заместителей могут рассматриваться в качестве структурных дескрипторов. В частности, атомные массы заместителей также могут рассматриваться в качестве дескрипторов элементного уровня. В целом же дескрипторы элементного уровня содержат слишком мало информации о молекулярной структуре и поэтому не могут самостоятельно использоваться для выявления реальных взаимосвязей структура – свойство.

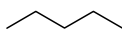
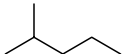
IV. Дескрипторы структурной формулы

Структурная формула молекулы является наиболее распространенным способом описания химического соединения. Как модель молекулярной структуры такая формула содержит элементы иконических, аналоговых и абстрактных моделей⁶ и является основой для конструирования многих видов дескрипторов и прежде всего топологических дескрипторов (рис. 2).

1. Топологические индексы

Топологические дескрипторы (индексы) 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).¹⁷ При этом обычно рассматриваются скелетные атомы (со «стертыми» атомами водорода) и связи между ними. Матричный вид графов используется для оценки топологических индексов. Наиболее часто применяются матрица смежности $A(G)$ и матрица расстояний $D(G)$. Элементы матрицы смежности a_{ij} равны или единице, или нулю в зависимости от того, связана ли ребром вершина i графа G с вершиной j или не связана. При этом число единиц в i -й строке или j -м столбце матрицы равно степени вершины. Каждый элемент в матрице расстояний $D(G)$ представляет число ребер, соединяющих вершину i с вершиной j наикратчайшим путем, и обозначается как d_{ij} . Матрицы смежности и матрицы расстояний для пентана, 2-метилпентана, циклопентана и фурана представлены на схеме 1. Отметим, что для двух последних соединений матрицы смежности и расстояний одинаковы вследствие идентичности молекулярных графов.

Наиболее распространенными топологическими индексами, рассчитываемыми на основе матрицы смежности, являются индекс Плата $F(G)$ (равен сумме степеней каждого

		Схема 1									
Соединение	Молекулярный граф	Матрица смежности					Матрица расстояний				
		1	2	3	4	5	1	2	3	4	5
$\text{CH}_3(\text{CH}_2)_3\text{CH}_3$		1	0	1	0	0	1	0	1	2	3
		2	1	0	1	0	2	1	0	1	2
		3	0	1	0	1	3	2	1	0	1
		4	0	0	1	0	4	3	2	1	0
		5	0	0	0	1	5	4	3	2	1
$\text{CH}_3\text{CH}(\text{CH}_3)(\text{CH}_2)_2\text{CH}_3$		1	0	1	0	0	1	0	1	2	3
		2	1	0	1	0	2	1	0	1	2
		3	0	1	0	1	3	2	1	0	1
		4	0	0	1	0	4	3	2	1	0
		5	0	0	0	1	5	4	3	2	1
		6	0	1	0	0	6	5	4	3	2
Циклопентан		1	0	1	0	1	1	0	1	2	2
		2	1	0	1	0	2	1	0	1	2
		3	0	1	0	1	3	2	1	0	1
		4	0	0	1	0	4	2	1	0	1
		5	1	0	0	1	5	1	2	2	1
		1	0	1	0	1	1	0	1	2	2
		2	1	0	1	0	2	1	0	1	2
		3	0	1	0	1	3	2	1	0	1
		4	0	0	1	0	4	2	1	0	1
		5	1	0	0	1	5	1	2	2	1

ребра в графе G); индекс Гордона – Скэнлберга $Y(G)$ (равен числу путей длины 2); индекс полной смежности $A'(G)$ (равен сумме всех ненулевых элементов матрицы смежности); индекс связности Рандича $\chi(G)$ (будет обсуждаться ниже), индексы заглавной группы $M_1(G)$ и $M_2(G)$. Среди дескрипторов, рассчитываемых на основе матрицы расстояний, можно отметить индекс Винера $W(G)$, индекс Хосойи $Z(G)$, число полярности $P(G)$, индекс суммы расстояний $S(G)$, индекс расстояний между вершинами $VDI(G)$. Формулы для расчета указанных распространенных дескрипторов и других топологических индексов, оцениваемых на основе матрицы смежности и расстояний, а также примеры использования таких дескрипторов в корреляциях структура – свойство приведены в работах^{18–20}.

Указанные матрицы и топологические индексы отражают только топологические особенности молекул без учета типа атомов и реальных расстояний между ними. И уже факт совпадения матриц смежности и расстояний, а следовательно, всех указанных выше топологических индексов для циклопентана и фурана свидетельствует о несовершенстве такого описания структуры соединений.

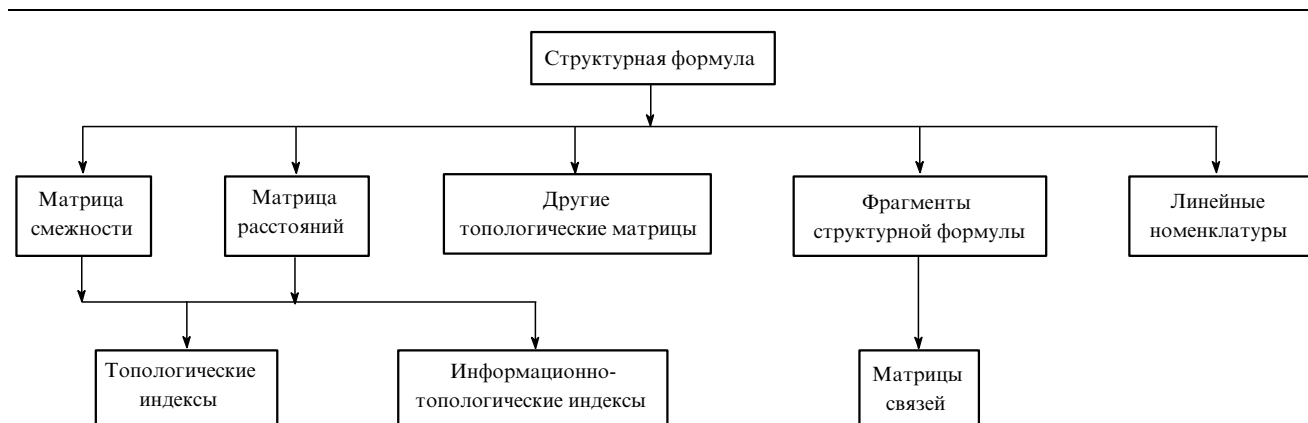


Рис. 2. Элементы описания молекулярной структуры на основе структурной формулы.

Помимо упомянутых матриц смежности и расстояний используют также расширенную дистанционную матрицу E ,²¹ матрицу связности C ,²² матрицу Винера W .²³

Многие исследователи используют различные дескрипторы, оцениваемые на основе матрицы связности C , поэтому кратко рассмотрим их в этом обзоре. В 1975 г. Рандич предложил индекс связности $\chi(G)$ определять следующим образом:²⁴

$$\chi(G) = \sum (v_i v_j)^{-1/2}, \quad (1)$$

где v_i и v_j — соответственно степени вершин i и j в графе G . Суммирование проводится по всем ребрам графа G .

Кир и Холл²⁵ увидели в схеме Рандича основу для развития общего метода описания структур органических соединений с использованием предложенного ими понятия молекулярной связности.²⁵ Первое уравнение для топологического индекса выглядело так:

$$^1X = \sum (\delta_i \delta_j)^{-1/2}, \quad (2)$$

где δ соответствует числу соседних атомов (без учета атомов водорода). Это уравнение идентично уравнению (1). Однако авторы расширили рассматриваемый подход, распространив фрагментацию молекулярного графа не только на связи, но и на атомы ($^0X = \sum (\delta)^{-1/2}$), а также на фрагменты с несколькими связями. Например, при расщеплении молекулярного графа на фрагменты с двумя связями оценивается индекс $^2X = \sum (\delta_i \delta_j \delta_k)^{-1/2}$, при расщеплении молекулярного графа на фрагменты с тремя связями оценивается индекс $^3X = \sum (\delta_i \delta_j \delta_k \delta_l)^{-1/2}$ и т. д. Кроме того, расчет таких индексов стал проводиться и для молекул, содержащих гетероатомы. При этом атомы водорода при гетероатомах (как и ранее атомы водорода при углероде) не принимались в расчет.

Позже Кир и Холл²⁶ предложили общее выражение для оценки δ любого атома

$$\delta_v = Z_v - h, \quad (3)$$

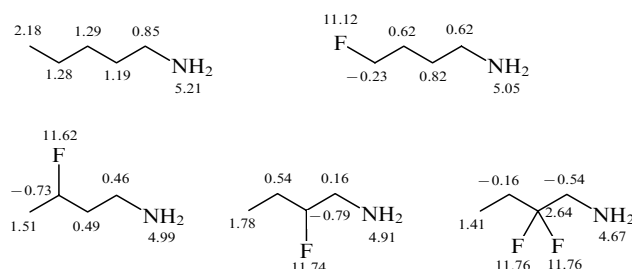
где Z_v — число валентных электронов с учетом электронов неподеленных электронных пар, а h — число атомов водорода при этом атоме (имеющихся в структурной формуле, но отсутствующих в молекулярном графе). Так, значение δ_v для атома углерода бензола равно трем (такой атом имеет четыре валентных электрона и соединен с одним атомом водорода), значение δ_v для атома кислорода в спиртах равно пяти (такой атом имеет шесть валентных электронов и соединен с одним атомом водорода), а в простых эфирах — равно шести (то же число валентных электронов, но отсутствуют атомы водорода при гетероатоме).

В последующих работах Кира и Холла^{27–30} было предложено количественно описывать взаимное влияние атомов на основе индексов электронного состояния (S). Значения S любого атома оцениваются по формуле

$$S = I + \Delta I, \quad (4)$$

где $I = (\delta_v + 1)/\delta$, а $\Delta I = (I_i - I_j)/r_{ij}^2$ (r — число связей между двумя рассматриваемыми атомами). Индексы электронотопологического состояния для некоторых замещенных аминов имеют следующие значения (схема 2).

Схема 2



Индексы Кира и Холла достаточно широко использовались при поиске количественных взаимосвязей структура–свойство (см., например, Приложение, уравнения (П.2) и (П.3)). Их популярность можно объяснить простотой расчета и наглядным представлением структурных формул рассматриваемых соединений. Среди последних работ можно отметить публикации^{31–39}.

Однако описание особенностей электронной структуры числом валентных электронов, а расстояний между атомами — числом связей между ними является слишком упрощенным для выявления реальной молекулярной структуры. Отсюда понятна критика возможности использования указанных индексов для нахождения реальных взаимосвязей структура–биологическая активность.⁴⁰

Распространенным индексом, рассчитываемым из матрицы расстояний, является индекс Винера $W(G)$ — первый топологический индекс, использованный в химии. Он определяется как полусумма элементов матрицы расстояний D . Индекс Винера и родственные ему топологические индексы часто используют в исследованиях структура–свойство для характеристики компактности молекул.^{41–43} Однако необходимо помнить, что в матрице D в качестве расстояния между атомами рассматривается число связей, разделяющих эти атомы. Чтобы приблизить описание структуры с помощью матрицы расстояний к реальной ситуации, предложено использовать в качестве недиагональных элементов матрицы усредненные (типичные) расстояния между атомами в молекулах, помещенных в трехмерную решетку,^{44–47} или расстояния, определенные по экспериментальным данным, либо полученные на основе квантово-химических расчетов.^{48,49}

Молекулярные графы служат источником создания разнообразных топологических индексов. Так, Кир^{50–54} предложил различные каппа-индексы для кодировки формы или конфигурации молекул; описание других новых индексов можно найти в работах^{55–64}. Конкретные корреляции таких топологических индексов с различными физическими свойствами и биологической активностью рассмотрены в публикациях^{65–72}.

Для описания структуры получили распространение также информационные топологические индексы. Оценка таких индексов основывается на разбиении множества вершин графа на классы эквивалентности с учетом характера ближайшего окружения вершин.^{18,73} В рамках этого подхода предложен ряд топологических индексов теоретико-информационного типа: 1IC (information content), 1SIC (structural information content), 1BIC (binding information content), 1CIC (complementary information content), рассчитываемые по формулам

$$^1IC = - \sum \frac{n_i}{n} \log_2 \frac{n_i}{n}, \quad (5)$$

$$^1SIC = \frac{^1IC}{\log_2 n}, \quad (6)$$

$$^1BIC = \frac{^1IC}{\log_2 q}, \quad (7)$$

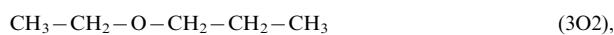
$$^1CIC = \log_2 n - ^1IC, \quad (8)$$

где n_i — число вершин в i -м классе эквивалентности, n — число вершин графа для молекулы, q — число ребер графа. Примеры использования информационных топологических индексов для описания структуры органических соединений при установлении взаимосвязей структура–свойство, в том числе структура–биологическая активность, приведены в работах^{74–80}.

2. Кодирование структурных формул с помощью линейных номенклатур и матриц связей

Необходимость автоматизации обработки информации о структурах химических соединений, объем которой в последние годы резко возрос, обусловила создание таких методов представления структурных формул химических соединений, как различные линейные номенклатуры и матрицы связей, которые можно рассматривать в качестве абстрактных моделей молекул.⁶

В линейной номенклатуре структурная формула кодируется строкой символов, которые в свою очередь кодируют определенные фрагменты структуры. При этом в каждой линейной номенклатуре существует свой строго регламентированный порядок следования символов. Обычно алфавит линейной нотации включает набор стандартных символов. Начальную стадию развития линейных номенклатур для описания структурных формул связывают с именем Висвессера.⁸¹ В качестве примера приведем коды Висвессера для структур некоторых соединений:



Линейная номенклатура Висвессера подробно описана в работе¹⁶.

В нашей стране аналогичный подход выражен Авидоном^{82, 83} с помощью языка ФКСП (фрагментальный код суперпозиции подструктур).

В настоящее время наиболее распространенная линейная номенклатура для представления структурных формул органических соединений содержится в системе SMILES,^{84–86} которая используется, в частности, для поиска структурного сходства соединений.⁸⁷ Заслуживает внимания также линейная номенклатура, предложенная в работе⁸⁸ для описания не только структурной формулы, но и стереоизомерии. В этой номенклатуре код соединения является каноническим и линейным и состоит из двух частей: первая представляет описание структурной формулы, а вторая характеризует стереохимические особенности молекулы.

При описании структурной формулы с помощью матрицы связности каждый атом, каждая связь и тип атомов кодируются прямым образом. Матрицы связности сходны с матрицами смежности и матрицами расстояний. Обычно главная диагональ матриц связности включает коды атомов в структурной формуле, недиагональные элементы представляют собой коды типов связей. Такие матрицы были достаточно распространенными в 70-е и 80-е годы, их описание можно найти в монографии¹⁶. В работе⁸⁹ была предложена матрица связности дескрипторных центров (МСДЦ), в главной диагонали которой коды атомов заменены на коды функциональных групп (дескрипторные центры), а в недиагональных элементах коды связей заменены на число связей между дескрипторами. С целью совершенствования языка описания структуры посредством МСДЦ в последующем было предложено использовать рассчитываемые расстояния между дескрипторными центрами.⁹⁰

3. Структурные фрагменты в качестве дескрипторов

Описанные выше топологические индексы и различные матрицы дают определенную информацию о химических соединениях в целом. Однако, поскольку решающую роль во взаимодействии соединений с другими химическими и биологическими объектами подчас играют только отдельные фрагменты молекул, описание структур соединений рассматриваемого ряда при выявлении взаимосвязей структура–свойство или структура–биологическая активность довольно часто производится пофрагментно. В качестве структур-

ного фрагмента может выступить отдельный атом, заместитель, функциональная группа, конкретное сочетание атомов (с указанием их типов гибридизации).⁹¹ Каждый такой структурный фрагмент может использоваться в качестве самостоятельного дескриптора. Возможна дальнейшая детализация структурного фрагмента за счет указания локальных характеристик атомов, входящих во фрагменты (см., например, классификацию дескрипторных центров на основе их донорно-акцепторной способности).^{92, 93} Существует много подходов к компьютерному конструированию новых веществ на основе структурных фрагментов соединений, обладающих каким-либо общим свойством. В этой связи можно упомянуть различные дескрипторы подструктурного анализа,^{94–97} топологические фармакофоры в процедуре программного комплекса LOGANA,^{98–100} подструктурные дескрипторы,^{101, 102} дескрипторы структурного языка системы DARC,^{103, 104} дескрипторы химической функциональности,¹⁰⁵ дескрипторы дизайна de novo.¹⁰⁶

V. Дескрипторы электронной структуры молекул

В квантово-химических расчетах сложных молекул используются атомные и молекулярные квантово-химические дескрипторы. Методы расчета дескрипторов, их возможности и ограничения достаточно подробно обсуждены в обзоре¹⁰⁷. В данном разделе кратко рассмотрены только квантово-химические дескрипторы, описывающие внутримолекулярные электронные свойства.

Заряды на атомах — наиболее часто используемые квантово-химические дескрипторы для нахождения взаимосвязей структура–свойство. Если разделить химические взаимодействия на электростатические и ковалентные, то заряды на атомах будут характеризовать электростатическую составляющую внутримолекулярных взаимодействий. Полный заряд на атоме можно рассматривать как ненаправленный дескриптор, в то время как σ - и π -электронные плотности характеризуют возможную ориентацию химических взаимодействий и поэтому их можно считать дескрипторами направленного типа.¹⁰⁸ Примеры успешного применения этих дескрипторов приведены в работах^{109–111}. Зная значения зарядов, можно в каждом конкретном исследовании предложить и использовать новый набор дескрипторов этого типа.

Энергии высшей занятой (E_{HOMO}) и низшей незанятой (E_{LUMO}) молекулярных орбиталей — дескрипторы, которые также часто используются в расчетах электронной структуры молекул.^{112–114} Установлено, что в ряде случаев значения E_{HOMO} прямо коррелируют с потенциалом ионизации (I) и характеризуют восприимчивость молекулы к атаке электрофилами. Значения же E_{LUMO} коррелируют с величиной сродства к электрону (A) и характеризуют восприимчивость молекулы к атаке нуклеофилами. Отметим, что I и A используются также в качестве самостоятельных дескрипторов электронной структуры.^{115, 116} Считается, что разность $E_{\text{HOMO}} - E_{\text{LUMO}}$ отражает меру стабильности молекулы.¹¹⁷ Понятия жесткости и мягкости молекул связаны именно с этими дескрипторами электронной структуры.¹¹⁸ Дескриптор молекулярная жесткость (η)

$$\eta = \frac{I - A}{2} \quad (9)$$

также нашел применение в соотношениях структура–свойство (см., например,¹¹⁹).[†]

[†] Жесткость молекулы как меру ее энергетической стабилизации некоторые авторы представляют дескриптором, определяемым как полусумма значений E_{HOMO} и E_{LUMO} (см.¹²⁰) или как взятое со знаком минус значение половины разности E_{HOMO} и E_{LUMO} (см.^{121, 122}).

Дипольный момент и родственные ему дескрипторы. Полярность молекулы описывается ее дипольным моментом (μ). Значение этого дескриптора может быть определено экспериментально или рассчитано с использованием какого-либо квантово-химического метода. Данный дескриптор достаточно широко применяется при установлении взаимосвязи структура – свойство.^{12, 112, 123–125} Однако при векторном сложении дипольных моментов нескольких полярных групп в молекуле может получиться небольшая суммарная величина. Поэтому наряду с μ иногда в качестве дескрипторов полярности используют значения наибольшей разности зарядов между двумя атомами, топологический электронный индекс, локальный дипольный индекс, а также тензор квадрупольного момента и др.¹⁰⁷

Примеры использования дескрипторов электронной структуры в соотношениях структура – биологическая активность приведены в Приложении (уравнения (П.4) и (П.5)).

VI. Дескрипторы молекулярной формы

В настоящее время необходимость учета особенностей пространственной структуры при выявлении взаимосвязей структура – свойства очевидна. Достаточно отметить, что практически все современные автоматизированные системы поиска новых соединений содержат в качестве обязательных элементов различные процедуры сравнения пространственных структур молекул.^{126, 127}

Описание трехмерной структуры молекул каким-либо одним количественным дескриптором (или набором дескрипторов) — достаточно сложная задача. Обсудим простейшую ситуацию, когда рассматриваемые соединения имеют общий каркас и различные заместители в одном и том же положении. В 1976 г. Верлуп с соавт. предложил набор из пяти стерических констант (параметры STERIMOL) для 243 различных заместителей; константы рассчитывались, исходя из стандартных длин связей и валентных углов.¹²⁸ Эти простые параметры нашли достаточно широкое применение в исследованиях структура – свойство.^{129–131} Аналогичный подход с оценкой параметров формы заместителей представлен в работе¹³².

Количественное сопоставление проекций структуры молекул на три ортогональных плоскости впервые было проведено Эймуrom¹³³ при изучении взаимосвязи структуры химических соединений с их запахом. Подход, описывающий форму молекул с помощью шести дескрипторов формы $S_1–S_6$, описан в работе¹³⁴.

Главная проблема, возникающая при сравнении параметров трехмерной структуры молекул различных соединений, связана с их конформационной подвижностью. С появлением дополнительной оси внутреннего вращения экспоненциально возрастает число возможных конформаций, поэтому требуются определенные ограничения, а также специальные процедуры отбора реальных конформаций и сравнения их форм.

В настоящее время существует ряд методов оценки дескрипторов молекулярной формы. Краткое их описание представлено ниже.

Метод дистанционной геометрии (distance geometry approach) развивается с конца 70-х годов и постоянно модифицируется Криппеном с соавт.^{135–141} Последовательные шаги процедуры сравнения пространственной формы молекул могут быть представлены следующими основными этапами.

1. Генерация трехмерной структуры рассматриваемых соединений; первоначальная конструкция строится на основе кристаллографических данных о длинах связей и валентных углах.

2. Приписывание атомам в молекулах определенных физико-химических свойств.

3. Оценка низкоэнергетических областей конформационного пространства для каждого из рассматриваемых соединений.

4. Определение геометрически возможных наложений молекул на базовую структуру. Вначале рассматриваются все комбинации трех атомов в базовой структуре и в сравниваемой молекуле. После нахождения наилучшего соответствия расстояний между какими-либо тремя атомами в сравниваемой паре соединений наложение этих трех пар атомов фиксируется и процесс совмещения в пространстве следующих атомов продолжается.

5. Анализ большого числа «хороших» наложений соединений на базовую структуру позволяет сопоставить молекулярные формы рассматриваемых соединений, при этом «материализуются» представления о дескрипторах молекулярной формы и дается оценка вида связывающей стороны возможного рецептора.

Модель связывающей стороны.^{142–144} Этот подход представляет связывающую сторону возможного рецептора не в виде точек (как в предыдущем методе), а в виде неперекрывающихся областей, занимающих все пространство. При таком представлении пространственной структуры каждый атом всегда должен находиться в одной (и только в одной) области, а модель связывающей стороны должна включать перечень областей (дескрипторов), в которых локализован каждый атом. На рис. 3,а в качестве примера представлена модель связывающей стороны, рассчитанная для соединений, проявляющих конкурентные свойства по отношению к протеину, выделенному из печени мышей,¹⁴⁵ а на рис. 3,б — модель оптимального связывания для бензопирена.

Анализ молекулярной формы. Этот подход развивается в последние 15 лет Хопфингером с соавт.^{146–155} Процедура количественного сравнения пространственной формы молекул различных соединений состоит из шести последовательных операций.

1. Конформационный анализ с учетом внутреннего вращения относительно всех одинарных связей и расчетом энергий для каждого поворота на 10° . При этом используются методы молекулярной механики с силовым полем, включающим вклады дисперсионных, стерических, электростатических взаимодействий и, где это возможно, водородной связи.

2. Отбор индивидуальных конформаций. Реальный набор возможных конформаций для каждого соединения определяется выбранным значением разности энергий для рассматриваемой конформации и глобального минимума энергии.

3. Выбор базового соединения и активной конформации при попарном сравнении форм молекул рассматриваемого ряда соединений.

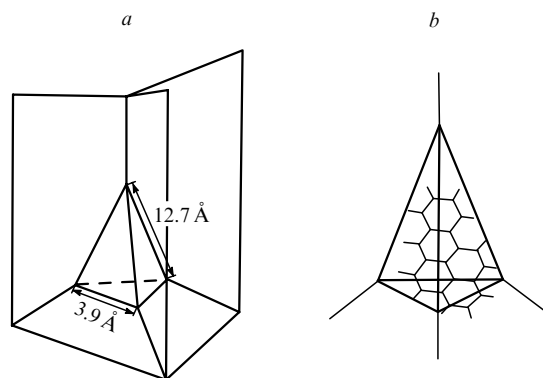


Рис. 3. Модель связывающей стороны для соединений, проявляющих конкурентные свойства по отношению к протеину (а) и модель оптимального связывания для бензопирена (б).¹⁴⁵

4. Наложение молекул каждого соединения рассматриваемого ряда на молекулу базового соединения.

5. Количественное сопоставление молекулярных форм. Предложены следующие дескрипторы сходства молекулярных форм:

общий пространственный объем перекрытия (V_0), определяемый по формуле

$$V_0 = V_u \cap V_v, \quad (10)$$

где V_u и V_v — пространственная занятость молекулами рассматриваемого и базового соединения соответственно;

относительная мера степени сходства в форме функции

$$f_0 = \frac{V_u \cap V_v}{V_u}. \quad (11)$$

Эти дескрипторы не содержат информации о внутримолекулярной устойчивости каждой конформации пары соединений. Поэтому предложен третий дескриптор — индекс общности пространственной формы (I_c):

$$I_c = S_{(u,v,w)} - \frac{\Delta E_u}{[\Delta E_u^* (\Delta E_u^* + \Delta E_v^* + \varepsilon)]^{1/2}} + \frac{\Delta E_v}{[\Delta E_v^* (\Delta E_u^* + \Delta E_v^* + \varepsilon)]^{1/2}}, \quad (12)$$

где $S_{(u,v,w)}$ — это wV_0 или wf_0 , w — параметр сходства формы, ΔE — разность конформационных энергий, $\Delta E_v^* = 1$ ккал·моль⁻¹, ΔE_u^* принимается равным 1, 3, 6 ккал·моль⁻¹ в трех различных расчетах, ε — поправочный коэффициент.

Гипотетическая решетка активной стороны.^{156–160} В основе этого метода лежит предположение о возможности описания молекулы дескрипторами пространства, которое она занимает, и природы атомов, которые занимают это пространство. Трехмерное пространство определяется как набор регулярно расположенных точек в трехмерной решетке. Трансляция молекулы к набору указанных точек выполняется при сохранении тех точек решетки, которые находятся в пределах ван-дер-ваальсовых радиусов каждого атома в молекуле. Результирующая молекулярная решетка, состоящая из точек, несет информацию о типе атомов в молекуле и его свойствах. Тип атома определяется просто: атомам, «богатым» электронами, приписывается значение +1, атомам, «бедным» электронами, приписывается значение -1, а «электронейтральным» атомам — значение 0.

Оценка молекулярных позиций.^{161–163} В этом подходе процесс сопоставления пространственных форм молекул также начинается с поиска низкоэнергетических конформеров и отбора энергетически устойчивых форм для каждого соединения рассматриваемого ряда. Под молекулярной позицией понимается специфическая конформация, связанная с проявлением активности соединения. Молекулярная позиция описывается несколькими дескрипторами.

Взвешенные целостные инвариантные молекулярные (ВЦИМ) дескрипторы конструируются таким образом, чтобы максимально охватить информацию о трехмерной структуре с рассмотрением размера, формы, симметрии молекул и распределения в них атомов.^{164–169} Алгоритм расчета включает анализ главных компонент центрированных молекулярных координат с использованием взвешенной ковариантной матрицы, получаемой из различных взвешенных схем для атомов.

Различные ковариантные матрицы и различные главные оси получаются в зависимости от типа взвешенной схемы. Например, в случае использования в качестве взвешенной

схемы атомных масс направления трех главных осей представляют главные оси инерции.

Для каждой взвешенной схемы рассчитывается набор статистических индексов на атомах, спроецированных на каждую главную компоненту t_m ($m = 1, 2, 3$). Инвариантность рассчитанных параметров по отношению к трансляции обеспечивается процедурой центрирования атомных координат, а инвариантность по отношению к вращению — использованием главных компонент.

Направленные ВЦИМ-дескрипторы представляют собой статистические индексы, рассчитываемые на основании каждой индивидуальной главной компоненты (1, 2, 3). Ненаправленные ВЦИМ-дескрипторы определяются непосредственно из направленных дескрипторов. Ненаправленные дескрипторы характеризуют общую форму молекул, а какая-либо информация, связанная с главными осями, в этом случае исчезает.

В публикациях^{164–169} описаны 11 направленных дескрипторов ($\lambda_1, \lambda_2, \lambda_3, \vartheta_1, \vartheta_2, \gamma_1, \gamma_2, \gamma_3, \eta_1, \eta_2, \eta_3$), характеризующих молекулярные свойства в пределах каждой взвешенной схемы. Таким образом, общее число направленных ВЦИМ-дескрипторов равно 66. В случае плоских структур необходимо учитывать только восемь направленных ($\lambda_1, \lambda_2, \vartheta_1, \vartheta_2, \gamma_1, \gamma_2, \eta_1, \eta_2$) и пять ненаправленных дескрипторов (T, A, V, K, D) для каждой взвешенной схемы.

Анализ молекулярной формы.^{170–174} Этот метод базируется на предположении, что форма электронного облака содержит всю информацию о свойствах соединения. Он был использован прежде всего для анализа молекулярных поверхностей. Поверхность молекулы расчленяется на области, характеризующиеся локальными свойствами кривизны. Выделяют локальные выпуклую D_2 , седлообразную D_1 и вогнутую D_0 области. Соотношения этих дескрипторов определяют группы форм молекул, которые включают двух-, одно- и нольмерные усеченные поверхности, полученные из первоначальной поверхности в результате удаления всех областей данного типа (например, D_2, D_1 или D_0). Количественная мера сходства форм трехмерных молекул может быть найдена при сравнении их числовых кодов форм. Последняя процедура выполняется автоматически с помощью компьютера без субъективного визуального сравнения форм объектов.

Примеры использования дескрипторов молекулярной формы в соотношениях структура–биологическая активность представлены в Приложении (уравнения (П.9)–(П.12)).

VII. Дескрипторы межмолекулярных взаимодействий

Количественно описать межмолекулярные взаимодействия при установлении взаимосвязей структура–биологическая активность можно с помощью разности свободных энергий (ΔG)^{175, 176}

$$\Delta G = G_f - G_i, \quad (13)$$

где G_f и G_i — свободная энергия конечного и исходного состояний соответственно.

Детализация количественного описания межмолекулярных взаимодействий возможна при использовании разностей энтальпии (ΔH) и энтропии (ΔS), связанных с ΔG соотношением

$$\Delta G = \Delta H - T\Delta S, \quad (14)$$

где T — температура реакции. Экспериментально измерив константу равновесия (K), можно определить ΔG

$$\Delta G = \Delta G^\circ - RT \ln K, \quad (15)$$

где ΔG° — разность свободной энергии реакции в стандартных условиях.

Значения ΔG , K , ΔH и ΔS для какой-либо реакции или взаимодействия можно рассматривать в качестве дескрипторов межмолекулярных взаимодействий. Например, при моделировании взаимосвязей структура – активность в качестве дескриптора достаточно часто используются константы диссоциации кислот и оснований (pK_a).^{177–180}

Значительная группа дескрипторов межмолекулярных взаимодействий, оцениваемая из экспериментальных данных по реакционной способности, состоит из различных констант, описывающих электронное влияние заместителей на реакционную способность химических соединений. Основы этого направления были заложены Гамметом, установившим значения параметра σ (константа Гаммета) на основе соотношения¹⁸¹

$$\sigma_X = \lg K_X - \lg K_H, \quad (16)$$

где K_X — соответствующая константа для *мета*- (σ_m) или *пара*-замещенной (σ_p) бензойной кислоты, K_H — константа ионизации для бензойной кислоты в воде при 25°C. К настоящему времени предложено множество констант, описывающих электронное влияние заместителей. Например, широко используются в качестве дескрипторов внутри- и межмолекулярных взаимодействий индукционная (σ_i) и резонансные (σ_r , σ^+ , σ^-) константы, индукционная константа Тафта (σ^*), константа Свена и Лаптона (F).^{182–184}

Дескрипторы реакционной способности, рассчитываемые квантово-химическими методами, составляют важную группу дескрипторов межмолекулярных взаимодействий. Необходимо прежде всего назвать индексы электрофильной (S_e) и нуклеофильной (S_n) суперделокализемости на реакционном центре (r), оцениваемые по формулам:^{185–187}

$$S_e(r) = 2 \sum_k^{\text{vac}} \sum_{\sigma(r)} \frac{(C_{\sigma k})^2}{0.5(E_{\text{НОМО}} + E_{\text{LUMO}}) - E_k}, \quad (17)$$

$$S_n(r) = 2 \sum_i^{\text{occ}} \sum_{\sigma(r)} \frac{(C_{\sigma i})^2}{E_i - 0.5(E_{\text{НОМО}} + E_{\text{LUMO}})}, \quad (18)$$

где числители содержат коэффициенты атомных орбиталей, знаменатели — энергии соответствующих молекулярных орбиталей.

Такие дескрипторы, как «самополяризуемость» $\pi(r)$,¹⁸⁵ отражающая влияние изменения электроотрицательности орбиталей атома на заряд того же атома:

$$\pi(r) = 4 \sum_i^{\text{occ}} \sum_k^{\text{vac}} \sum_{\sigma(r)} \frac{C_{\sigma i}^2 C_{\sigma k}^2}{E_k - E_i}, \quad (19)$$

и атом-атомная поляризуемость $\pi(AB)$,¹⁰⁷ отражающая эффект изменений при одном атоме А на заряд другого атома В

$$\pi(AB) = 4 \sum_i \sum_a \sum_p \sum_r \frac{C_{\pi}^A C_{pa}^A C_{ri}^B C_a^B}{E_i - E_a}, \quad (20)$$

также могут описывать и межмолекулярное взаимодействие.

Эти дескрипторы тоже нашли применение при описании структуры.^{107, 186–188}

Молекулярные рефракция и поляризуемость. При помещении молекулы во внешнее постоянное электрическое поле в ней происходит изменение распределения зарядов. Результирующий дипольный момент определяется соотношением

$$\mu = \alpha E, \quad (21)$$

где α — молекулярная поляризуемость, E — напряженность электрического поля. Экспериментальная оценка поляризуемости может быть произведена на основе ее связи с молекулярной рефракцией (MR)

$$MR = \frac{4\pi N_A \alpha}{3}, \quad (22)$$

(N_A — число Авогадро) и связи рефракции с показателем преломления (n)

$$MR = MV \frac{n^2 - 1}{n^2 + 2}, \quad (23)$$

где MV — молекулярный объем. Оценка MR и α производится на основе либо аддитивности эмпирически установленных вкладов отдельных атомов или фрагментов,¹⁸⁹ либо с помощью квантово-химических расчетов.¹⁹⁰ Примеры использования MR и α можно найти в публикациях^{191–194}.

Липофильность. Самая распространенная группа дескрипторов межмолекулярных взаимодействий для анализа связи структура – активность связана с понятием липофильности, характеризующей транспортные свойства соединений в биологических объектах. В подавляющем числе случаев для количественного описания липофильности берется распределение (P) соединений в модельной системе октанол – вода. В качестве дескриптора этой системы используется значение $\lg P$ для нейтральной формы соединения. Соответствующие дескрипторы для катионной, анионной форм и цвиттер-ионной структур обозначаются индексами $\lg P^+$, $\lg P^-$ и $\lg P^{+/-}$ (см.¹⁹⁵). Распределение ионизированных веществ между октанолом и водой зависит от pK_a самого соединения и pH среды. Для характеристики распределения веществ в биологических средах вместо $\lg P$ используется индекс $\lg D$ с указанием значения pH. В подавляющем числе публикаций $\lg D$ относится к «физиологической» среде с pH 7.4.

Взаимосвязь дескрипторов $\lg D$, $\lg P$, $\lg P^+$, $\lg P^-$ и pK_a определяется соотношениями¹⁹⁶

$$\lg D = \lg P_{\text{XH}} - \lg(1 + 10^{-pK_a + \text{pH}}); \quad (24)$$

для равновесия $\text{X} + \text{H}^+ \rightleftharpoons \text{XH}^+$

$$\lg D = \lg P_{\text{X}} - \lg(1 + 10^{-pK_a - \text{pH}}); \quad (25)$$

для равновесий $\text{X}^- + \text{H}^+ \rightleftharpoons \text{XH} + \text{H}^+ \rightleftharpoons \text{XH}_2^+$

$$\lg D = \lg(A_0 + A_1 + A_2), \quad (26)$$

где

$$A_0 = \frac{P_{\text{X}^-}}{1 + 10^{pK_{a2} - \text{pH}} + 10^{pK_{a1} + pK_{a2} - 2\text{pH}}}, \quad (27)$$

$$A_1 = \frac{P_{\text{XH}}}{1 + 10^{pH - pK_{a2}} + 10^{pK_{a1} - \text{pH}}}, \quad (28)$$

$$A_2 = \frac{P_{\text{XH}_2^+}}{1 + 10^{pH - pK_{a1}} + 10^{2\text{pH} - pK_{a2} - pK_{a1}}}. \quad (29)$$

Методы экспериментальной оценки параметров липофильности подробно описаны в обзорах^{197–199}.

Расчет липофильности осуществляется по различным схемам, включая аддитивность вкладов липофильных констант атомов²⁰⁰ или фрагментов,^{201, 202} конформационно-зависимые подходы,^{203, 204} использование молекулярных дескрипторов,²⁰⁵ различных сольватохромных параметров.²⁰⁶ На распределение соединений в системе октанол – вода влияют в основном два фактора — объем молекул и их способность к образованию водородных связей.^{207, 208}

В работах^{209, 210} для количественного описания вклада молекулярного объема в липофильность использованы такие дескрипторы, как молекулярная масса MW, молекулярный объем MV, молекулярная поверхность SA, интеграл спектров ван-дер-ваальсовых взаимодействий VWI, молекулярная рефракция MR и молекулярная поляризуемость α . Для количественного описания вклада водородной связи исполь-

Таблица 1. Экспериментальные и рассчитанные значения липофильности и вкладов поляризуемости и акцепторной способности при образовании водородных связей.²¹⁰

Соединение	$\sum C_a$ α		$\sum C_d$ α		$\lg P$	
					расчет	эксперимент
Хлорацетамид	3.80	8.00	-2.57	1.99	-0.59	-0.53
Ацетальдегидоксим	2.64	5.59	-1.79	1.39	-0.41	-0.12
Этаноламин	4.16	6.43	-2.82	1.60	-1.23	-1.30
Никотинамид	5.88	13.19	-3.98	3.28	-0.71	-0.37
Гидрохинолин	3.00	11.32	-2.04	2.82	0.77	0.59
Диметилсульфит	1.50	10.44	-1.02	2.60	1.57	1.77
Ацикловир	10.90	21.44	-7.39	5.34	-2.06	-1.52
Сефопиразон	24.10	60.97	-16.34	15.18	-1.17	-0.84
Фторурацил	5.78	10.32	-3.92	2.57	-1.36	-0.85
Гидрофлуметиазид	9.31	25.16	-6.31	6.26	-0.06	0.04
Парацетамол	5.49	16.01	-3.72	3.99	0.26	0.64
Варфарин	7.41	32.83	-5.02	8.17	3.14	2.47

зованы значения факторов доноров и акцепторов свободной энергии водородной связи C_d и C_a , которые подробнее будут обсуждены ниже. Поскольку дескрипторы, описывающие вклад объема молекул, коррелировали между собой, в регрессионное уравнение можно включить только один из них. Наилучший результат был получен при использовании поляризуемости и суммы факторов акцепторной способности к водородной связи (см. Приложение, уравнение (П.14)). Такой подход к количественному описанию липофильности позволяет оценить противоположно направленные вклады поляризуемости и водородной связи и тем самым модифицировать структуру соединений для оптимизации их свойств. В табл. 1 приведены экспериментальные и рассчитанные по уравнению (П.14) значения липофильности некоторых лекарств, а также вклады в липофильность двух указанных параметров.

Анализ многочисленных компьютерных программ, предназначенных для расчета $\lg P$ (см.^{211–213}), показал, что точность расчета липофильности простых нейтральных органических соединений удовлетворительная, а сложных соединений, особенно лекарств, содержащих многие функциональные группы, — весьма невысокая. Это можно объяснить тем, что в большинстве расчетов используется не точная аддитивная схема вкладов коэффициентов распределения соединений в системе октанол–вода (oct), а аддитивность логарифмов вкладов атомов и фрагментов. Кроме того, опубликовано мало результатов экспериментальных определений липофильности соединений, содержащих несколько катионных и(или) анионных групп. В то же время именно такие данные должны составлять основу для точных расчетов липофильности сложных ионизированных соединений.

Применение уравнения (П.14) к расчету значений $\lg P_{oct}$ для различных нейтральных органических соединений привело к установлению следующего соотношения:

$$\lg P_{oct}(\text{эксперимент}) = 1.00 (\pm 0.02) \lg P_{oct}(\text{расчет}), \quad (30)$$

$$N = 2781, R = 0.971, S = 0.32 (\text{см.}^\ddagger).$$

Сравнение рассчитанных по уравнению (П.14) и экспериментальных значений липофильности для соединений, которые уже были тестированы 14 соответствующими коммерческими программами, дали следующие результаты: для 90 простых органических соединений $R = 0.993$; для 48

лекарств $R = 0.935$, а для всех 138 тестируемых соединений $R = 0.971$. Указанные статистические критерии сопоставимы с такими данными для лучших из испытанных программ. Эти результаты, а также возможность рассчитывать полный pH-зависимый профиль для ионизированных соединений свидетельствуют в пользу указанного выше физико-химического подхода к оценке липофильности.

В последние годы для характеристики липофильности наряду с приведенными выше «классическими» дескрипторами широко используется такой дескриптор, как молекулярный потенциал липофильности (Molecular Lipophilicity Potential, MLP),^{214–218} количественно описывающий липофильность в трехмерном пространстве. В данной точке пространства значение потенциала липофильности представляет результат межмолекулярных взаимодействий, закодированных липофильностью всех фрагментов в молекуле. Для расчета потенциала липофильности необходимы две компоненты — значения липофильности фрагментов и функция расстояния, описывающая изменение липофильности в пространстве²¹⁴

$$MLP_k = \sum_{i=1}^N f_i f_{ct}(d_{ik}), \quad (31)$$

где k — метка данной точки пространства, i — метка фрагмента, N — общее число фрагментов в молекуле, f_i — константа липофильности i -го фрагмента, f_{ct} — функция расстояния, d_{ik} — расстояние между фрагментом i и точкой k в пространстве.

Множество примеров успешного применения параметров липофильности в качестве дескрипторов межмолекулярных взаимодействий для установления корреляций структура–свойство приведены в книге¹⁸⁴ и статьях^{214–224}. Примеры использования $\lg P_{oct}$ в регрессионных соотношениях структура–свойство (активность) представлены в Приложении (уравнения (П.14)–(П.16)).

Следует отметить, что большинство процедур расчета как дескрипторов межмолекулярных взаимодействий, так и энергий реакций основывается на предположении об аддитивности вкладов различных типов межмолекулярных взаимодействий¹⁷⁵

$$\Delta G = \Delta G_{\text{solvent}} + \Delta G_{\text{conf}} + \Delta G_{\text{int}} + \Delta G_{\text{motion}}, \quad (32)$$

где $\Delta G_{\text{solvent}}$ — вклад растворителя, ΔG_{conf} — изменение свободной энергии за счет конформационных изменений во взаимодействующих молекулах, ΔG_{int} — изменение свободной энергии за счет специфических взаимодействий, ΔG_{motion} — изменение свободной энергии за счет изменения ее трансляционной, вращательной и колебательной составляющих.

Специфические взаимодействия рассчитываются на основе предположения об аддитивности вкладов взаимодействий различных типов. В настоящее время считается, что основной вклад в специфические взаимодействия вносят стерический (st), электростатический (el) эффекты и водородная связь (hb). В рамках этого предположения свободная энергия специфических взаимодействий может быть представлена следующим образом:²²⁵

$$\Delta G_{\text{int}} = \Delta G_{\text{st}} + \Delta G_{\text{el}} + \Delta G_{\text{hb}}. \quad (33)$$

Ниже рассмотрена параметризация трех указанных типов специфических взаимодействий, проанализированы имеющиеся дескрипторы и описаны компьютерные программы для расчетов энергий этих взаимодействий.

1. Параметризация стерических эффектов

Первый стерический параметр E_s был предложен Тафтом.²²⁶ При изучении реакционной способности эфиров им получено следующее соотношение:

‡ Здесь и далее N , S , R — статистические критерии уравнений: N — число наблюдений, S — стандартное отклонение, R — коэффициент корреляции; используется также критерий Фишера (F).

$$E_s = \lg \frac{k_R}{k_{Me}}, \quad (34)$$

где k_R и k_{Me} — соответственно константы скоростей гидролиза для соединений R^1COOR^2 и $MeCOOR^3$. Предполагается, что в рассматриваемой системе нет какого-либо влияния индукционного и электронного эффектов. В работе приведены таблицы уточненных значений E_s . Примеры успешного применения этого стерического параметра можно найти в статьях^{15, 184}.

В простейшем случае стерические эффекты могут быть охарактеризованы значениями ван-дер-ваальсовых радиусов. Более детальная разработка эмпирических стерических параметров на основе этих радиусов проведена Чартон^{182, 227, 228}. В наиболее простом из предложенных им монопараметрическом методе используется стерический параметр v , который зависит от значений ван-дер-ваальсовых радиусов и выражает стерический эффект первого атома заместителя. В случае атомных группировок, имеющих разные заместители при атомах, через которые «проходит» ось заторможенного внутреннего вращения, для описания стерического эффекта предлагается использовать или набор эффективных значений стерических параметров определенных реакций, или эмпирические по своему характеру уравнения разветвления.

Отметим, что эмпирический подход Чартона включает многие упрощения и требует введения дополнительных эмпирических констант и коэффициентов при рассмотрении новых серий соединений.

Потенциал Леннарда–Джонса учитывает стерическую составляющую суммарной энергии взаимодействия E_{lj} , которая рассматривается как комбинация сил притяжения и отталкивания²²⁵

$$E_{lj} = \frac{A}{d^{12}} - \frac{B}{d^6}, \quad (35)$$

где d — расстояние между парой несвязанных атомов. Когда d мало, член A/d^{12} обуславливает преобладание сил отталкивания и, как следствие, положительные значения E_{lj} . При больших d больше член B/d^6 . Минимальное значение E_{lj} соответствует оптимальному расстоянию несвязанных атомов. Потенциал Леннарда–Джонса может использоваться как самостоятельный дескриптор межмолекулярных взаимодействий, а также для оценки вклада стерических взаимодействий в общую энергию взаимодействия. В частности, как показано в работе²²⁹, для ряда замещенных эфиров уксусной кислоты стерический дескриптор, рассчитанный на основе потенциалов Леннарда–Джонса, хорошо коррелирует с параметром Тафта E_s .

2. Описание электростатических взаимодействий

Для расчета электростатических взаимодействий используются закон Кулона для центрированных на атомах точечных зарядов, а также стандартные подходы молекулярной механики.¹⁷⁵

Практически всегда при сравнении способности молекул к электростатическим взаимодействиям $V(R_i)$ применяют молекулярный электростатический потенциал (molecular electrostatic potential, MEP), который можно рассчитать в каждой точке R пространства, окружающего молекулу,²³⁰

$$V(R_i) = - \int \frac{\delta(R)}{|R_i - R|} dr + \sum \frac{Z_l}{|R_i - R_l|}, \quad (36)$$

где $\delta(R)$ — электронная плотность, r — расстояние, Z_l — заряды ядер, R_l — координаты ядер в молекуле. Результаты расчета MEP можно графически представить с помощью карт. Сравнение карт MEP ряда соединений, имеющих один и тот же тип активности (примеры таких исследований см. в

работах^{231–235}), может оказаться полезным для суждения о характере специфических электростатических взаимодействий этих соединений с биологической мишенью для активности рассматриваемого типа. Однако здесь необходимо помнить о субъективности визуального сравнения.

В последние годы интенсивно развиваются формальные подходы сравнения квантово-химических параметров, связанных с электростатическими взаимодействиями. Прежде всего необходимо упомянуть о мерах молекулярного квантово-химического подобия (molecular quantum similarity measures, MQSM).^{236–239} В этом подходе квантово-химическое подобие двух молекул выявляется при рассмотрении функционала от электронных плотностей ($Z_{AB}(\Omega)$):

$$Z_{AB}(\Omega) = \int \rho_A(\mathbf{r}_1) \Omega(\mathbf{r}_1, \mathbf{r}_2) \rho_B(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \quad (37)$$

где $\rho_A(\mathbf{r}_1)$, $\rho_B(\mathbf{r}_2)$ — электронные плотности молекул А и В, Ω — положительно-определенный оператор. Могут быть выбраны различные операторы. Удобно представить оператор в форме

$$\Omega(\mathbf{r}_1, \mathbf{r}_2) = \delta(\mathbf{r}_1 - \mathbf{r}_2), \quad (38)$$

которая трансформирует интеграл в уравнении (37) в интеграл от произведения распределений плотностей в данной паре молекул

$$Z_{AB} = \int \rho_A(\mathbf{r}) \rho_B(\mathbf{r}) d\mathbf{r}. \quad (39)$$

Если сравниваемые молекулы А и В одинаковы, то получается мера самоподобия

$$Z_{AA} = \int \rho_A^2(\mathbf{r}) d\mathbf{r}. \quad (40)$$

Мера самоподобия используется для построения индексов сходства, которые могут рассматриваться как нормализованные значения MQSM. Первым индексом сходства был предложен индекс Карбо²⁴⁰

$$C_{AB} = \frac{Z_{AB}}{(Z_{AA} Z_{BB})^{1/2}}. \quad (41)$$

Его значения изменяются от нуля (при сравнении несхожих молекул) до единицы (при сравнении идентичных молекул). Примеры использования индекса Карбо для выявления сходства электронных структур приведены в публикациях^{241–244}.

3. Количественное описание способности соединений к образованию водородных связей

Водородная связь имеет определяющее значение в образовании многих молекулярных и ионных комплексов в важных химических и биологических процессах, включая ферментативный катализ. Наиболее яркий пример, показывающий важную роль водородной связи в биологических системах, приведен в монографии²⁴⁵ при описании процесса репликации ДНК. Этот процесс сопровождается новым объединением пуриновых и пиримидиновых оснований путем образования двух (связывание аденина и тимина) или трех (связывание гуанина и цитозина) водородных связей. Проведенный Полингом даже весьма приближенный расчет разности свободных энергий этого комплексообразования (принималось, что одна водородная связь дает выигрыш в энергии в 20 кДж·моль^{–1}) продемонстрировал, что указанная система водородных связей обеспечивает отношение чисел «правильных» и «неправильных» распознаваний как десять миллионов к одному!

Несмотря на столь важную роль водородной связи в формировании свойств веществ, учет этого межмолекуляр-

ного взаимодействия до последнего времени осуществлялся преимущественно на уровне индикаторной переменной.^{12, 13, 246–248} Поэтому в настоящем обзоре целесообразно детально обсудить количественные подходы к описанию водородной связи.

Согласно уравнению (14), для полного термодинамического описания водородной связи требуется пара параметров — энтальпия и энтропия или энтальпия и свободная энергия. Возможны четыре комбинации изменений энтальпии (ΔH) и энтропии (ΔS).^{249–252} Отметим, что авторы большинства работ, посвященных изучению водородных связей в комплексах, изменением энтропии либо пренебрегают вообще, либо считают его пропорциональным изменению энтропии. Огрубленность такого подхода очевидна из примеров, представленных в обзоре²⁵¹.

Развитие эмпирических шкал водородной связи на основе экспериментальных значений термодинамических параметров началось в 1960 годы. Рассмотрим два подхода — аддитивно-мультипликативный и мультипликативный. В случае аддитивно-мультипликативного подхода изменение энтальпии при образовании водородной связи описывается шестью параметрами^{253–255}

$$-\Delta H_{\text{hb}} = e_{\text{AeB}} + c_{\text{AcB}} + t_{\text{AtB}}, \quad (42)$$

где e — параметр, связанный с электростатическим вкладом, c — параметр, связанный с ковалентным взаимодействием, t — параметр, связанный с переносом заряда, индексы A и B относятся соответственно к люисовским кислотам (доноры водородной связи) и основаниям (акцепторы водородной связи). В большинстве исследований корреляций структура–биологическая активность отсутствуют точные сведения о молекулярной структуре одного из партнеров по взаимодействию (например, рецептора, биомишени). Это обстоятельство затрудняет использование уравнения (42). Более широкое распространение получил метод описания термодинамических параметров водородной связи на основе мультипликативного подхода. Энтальпия водородной связи описывается уравнением²⁵⁶

$$\Delta H_{ij} = \Delta H_{11} P_i E_j, \quad (43)$$

где ΔH_{11} — изменение энтальпии при образовании водородной связи между фенолом и диэтиловым эфиром (стандартные партнеры водородной связи) в четыреххлористом углероде,[§] P_i — безразмерный фактор i -донора водородной связи, E_j — фактор j -акцептора водородной связи.

Для описания изменения ΔH_{ij} можно использовать также донорный (E_d) и акцепторный (E_a) факторы энтальпии водородной связи, имеющие противоположные знаки,^{92, 257}

$$\Delta H_{ij} = \Delta H_{11} E_d E_a. \quad (44)$$

Значения E_d и E_a для различных нейтральных и ионизированных соединений оценены в работах^{258–270}.

Аналогичные уравнения были предложены для описания изменения свободной энергии (ΔG) при образовании водородной связи^{271, 272}

$$\Delta G = a C_d C_a + a_0, \quad (45)$$

где a и a_0 — коэффициенты, C_d и C_a — донорные и акцепторные факторы свободной энергии водородной связи; а также констант комплексообразования (K) для водородной связи^{273–276}

$$\lg K = b \alpha \beta + b', \quad (46)$$

где b и b' — коэффициенты, α и β — донорные и акцепторные факторы констант образования водородной связи. В

§ Большинство экспериментов по термодинамике водородной связи проведено в четыреххлористом углероде. Этот растворитель неполярный, поэтому удобен для установления значений описываемых ниже термодинамических факторов.

работе²⁷⁷ оценено большое число значений β для фосфорильных соединений по данным публикаций^{261–263}. С учетом взаимосвязи изменения свободной энергии и констант (см. уравнение (15)) найдено, что различие уравнений (45) и (46) заключено в коэффициенте $2.302RT$. В случае одного и того же массива экспериментальных данных и метода оценки параметров водородной связи должна существовать строгая корреляция между C_d и α , C_a и β . Общей корреляции параметров E_d и C_d , а также E_a и C_a не обнаружено, однако для отдельных групп соединений эти параметры коррелируют.²⁷²

Развитие и успешное применение приведенных выше эмпирических подходов возможно лишь при расширении базы экспериментальных данных по термодинамике водородной связи. Детальное описание обширной базы таких данных представлено в статье²¹⁰. Для расчета значений C_a и C_d были отобраны 414 доноров водородной связи и 1298 акцепторов, для которых имелись экспериментальные термодинамические параметры взаимодействия с несколькими партнерами.²⁵¹ Составленная матрица содержала 414 столбцов и 1298 строк. Это позволило провести строгий расчет значений C_a и C_d .

Приведем значения указанных факторов свободной энергии водородной связи для некоторых соединений.

Соединение	C_d	Соединение	C_a
CF ₃ COOH	–4.78	CCl ₃ CN	0.55
CCl ₃ COOH	–4.75	C ₄ H ₉ Cl	0.60
CHCl ₂ COOH	–4.28	C ₆ H ₅ OCH ₃	1.07
3,5-(NO ₂) ₂ C ₆ H ₃ OH	–4.09	CH ₃ OH	1.76
C ₆ F ₅ COOH	–3.82	CH ₃ C(O)CH ₃	1.92
3-NO ₂ C ₆ H ₄ OH	–3.49	(CH ₃) ₂ SO ₂	2.01
(CF ₃) ₂ CHOH	–3.31	C ₅ H ₅ N	2.45
CF ₃ C(O)NHC(O)CF ₃	–2.98	(C ₂ H ₅) ₃ N	2.57
C ₆ H ₅ OH	–2.50	(CH ₃) ₂ NC(O)N(CH ₃) ₂	2.96
CH(CN) ₂ Br	–2.06	(CH ₃) ₂ SO	3.13
4-NO ₂ C ₆ H ₄ NH ₂	–1.82	(C ₆ H ₅) ₃ PO	3.65
CH ₃ C(O)SH	–1.62	[(CH ₃) ₂ N] ₃ PO	4.00
C ₂ H ₅ OH	–1.42	(C ₄ H ₉) ₂ NO	4.25
C ₆ H ₅ NH ₂	–1.04		
CHCl ₃	–0.81		
2-CH ₃ C ₆ H ₅ NH ₂	–0.50		
C ₂ H ₅ SH	–0.24		
C ₆ H ₅ ≡CH	–0.12		

Важно подчеркнуть, что интервалы значений факторов свободной энергии водородной связи для различных функциональных групп перекрываются (табл. 2). Это свидетельствует о том, что заместители при донорном или акцепторном атоме играют исключительно важную роль в формировании способности этих атомов к образованию водородной связи.

Сравнение экспериментальных значений энтальпии и свободной энергии со значениями, рассчитанными с исполь-

Таблица 2. Минимальные и максимальные значения факторов свободной энергии водородной связи для различных функциональных группировок.

Доноры водородной связи	$C_d^{\min}$	$C_d^{\max}$	Акцепторы водородной связи	$C_a^{\min}$	$C_a^{\max}$
R ¹ R ² R ³ C–H	–2.06	–0.12	R–O–R	0.78	2.01
RS–H	–1.92	–0.24	R–S–R	0.57	1.51
R ¹ R ² N–H	–3.32	–0.50	R ¹ R ² R ³ N	0.58	3.65
PhO–H	–4.24	–0.76	RC≡N	0.55	2.77
AlkO–H	–3.50	–1.06	R ¹ R ² R ³ P=O	2.40	4.02
RC(O)O–H	–4.78	–2.50	R ¹ R ² S=O	1.16	3.25

зованием уравнений (44) и (45), показывает адекватность мультипликативного подхода для количественного описания термодинамики водородной связи²⁵¹

$$\Delta H(\text{расчет}) = -0.49 (\pm 0.29) + 0.99 (\pm 0.08) \Delta H(\text{эксп.}), \quad (47)$$

$$N = 2787, R = 0.970, S = 2.40, F = 44350,$$

$$\Delta G(\text{расчет}) = -0.07 (\pm 0.12) + 1.04 (\pm 0.05) \Delta G(\text{эксп.}), \quad (48)$$

$$N = 3301, R = 0.991, S = 1.12, F = 175000.$$

В случае значительного вклада переноса заряда в водородную связь какого-либо конкретного комплекса возможны отклонения рассчитанных таким способом значений от экспериментальных.²⁷⁸

Зная C_d для 414 доноров и C_a для 1298 акцепторов водородной связи, можно рассчитать значения ΔG_{hb} для 537372 комплексов с водородной связью. При использовании аналогичного подхода в расчетах констант образования водородных связей установлены значения α для 150 доноров и β для 500 акцепторов водородной связи, что дало возможность рассчитать K_{hb} для 75000 комплексов.²⁷⁹

Для мультипликативного описания термодинамических параметров водородной связи предложен целый ряд дескрипторов, включая донорные и акцепторные факторы свободной энергии и энтальпии водородной связи для наиболее сильного донорного и акцепторного центров в молекуле ($C_a^{\max}$, $C_d^{\max}$, $E_a^{\max}$, $E_d^{\max}$), сумму факторов свободной энергии и энтальпии для всех донорных или акцепторных центров в молекуле — $\sum C_a$, $\sum C_d$, $\sum E_a$, $\sum E_d$, значения последних четырех дескрипторов, деленных на молекулярную массу ($\sum C_a/MW$, $\sum C_d/MW$, $\sum E_a/MW$, $\sum E_d/MW$).^{280, 281} Указанные дескрипторы были использованы для моделирования взаимосвязей структура–свойство при изучении таких сложных процессов, как комплексообразование,^{91, 282–286} растворимость в воде,²¹⁰ липофильность,²¹⁰ ингибирование холинэстеразы,^{287, 288} ингибирование системы обратного захвата дофамина,²⁸⁹ определение температур кипения углеводов и кислот,²⁹⁰ инсектицидная активность азот- и фосфорорганических соединений,²⁹¹ ионофорная активность макроциклов,²⁹² фунгицидная активность фенолов и β -адренэргическая активность фенилэтиламина,²⁸⁰ противоспазматическая активность алициклических кислот и фенолов,²⁹³ токсичность по отношению к *Golden orfe fish*,²⁹⁴ проницаемость различных биологических мембран.^{210, 295} Дескрипторы, основанные на факторах α и β , описаны в обзоре²⁷⁹. Сходные индексы водородной связи опубликованы в работе²⁹⁶.

VIII. Индикаторные дескрипторы

При моделировании взаимосвязи структура–свойство (активность) достаточное распространение получили индикаторные дескрипторы (иногда их называют константы de novo). Отнести эти дескрипторы к какому-либо одному определенному классу трудно, поскольку они могут кодировать в неявном виде различные структурные особенности соединений (свойства доноров или акцепторов водородной связи, наличие или отсутствие водородной связи внутри молекул, орто-эффекты, цис–транс- или стереоизомерию, различные фрагменты и т. д.).

В обзоре²⁹⁷ показана целесообразность использования в некоторых случаях индикаторных переменных. Так, для двух выборок папаиновых лигандов были получены сходные двухпараметрические соотношения (см. Приложение, уравнения (П.7) и (П.8)) с достаточно близкими коэффициентами при MR и σ , но с существенно различными значениями свободных членов. Комбинирование указанных уравнений и использование индикаторной переменной I ($I = 1$ для мезиламидов и $I = 0$ для бензамидов) позволили найти общее для

двух рядов соединений соотношение (см. Приложение, уравнение (П.24)).

Другие примеры использования индикаторных дескрипторов можно найти в публикациях^{298, 299}.

Индикаторные дескрипторы особенно полезны на ранних стадиях исследований. С помощью таких дескрипторов могут предварительно формироваться комбинации различных подвыборок соединений с рассматриваемым свойством для получения обучающей выборки, предназначенной для моделирования связи структура–свойство на основе уже более углубленного описания структуры.

При анализе взаимосвязи структура–биологическая активность в качестве дескрипторов структуры часто используются различные молекулярные физико-химические свойства или экспериментальные параметры, полученные различными физическими методами.¹⁵ Использование физико-химических свойств в качестве дескрипторов обсуждалось в предыдущих разделах. Подробно ознакомиться с такими дескрипторами можно в работе³⁰⁰.

IX. Современные компьютерные программы для расчетов дескрипторов

В настоящее время рынок компьютерных программ для химии вообще и для молекулярного дизайна в частности насыщен всевозможными программными продуктами. Поэтому, чтобы быть конкурентноспособной, каждая новая программа должна иметь элементы очевидной новизны или явные преимущества по сравнению с другими.

1. MOLSURF — генератор химических дескрипторов для QSAR.¹⁹³ Эта программа имеет три отличительных качества.

Во-первых, предполагается, что исследователь предпочитает описывать молекулы не квантово-химическими (HOMO, LUMO, электростатический потенциал, силовые поля и т. д.), а химическими терминами — водородные связи, нуклеофилы, электрофилы, основания, кислоты, гидрофобность, поляризуемость, π -взаимодействия и т. д. Поэтому программа MOLSURF рассчитывает дескрипторы, связанные с указанными химическими понятиями.

Во-вторых, программа базируется на корреляциях рассчитываемых значений дескрипторов с экспериментальными данными.

В-третьих, программа отдает предпочтение дескрипторам заместителей по сравнению с молекулярными дескрипторами, поскольку первые более точно отражают изменение свойств при модификации структуры и можно более точно рассчитать их значения.

2. Расчет и использование теоретических дескрипторов в рамках концепции линейных соотношений свободной энергии описаны в публикациях^{301–306}. В этом достаточно простом подходе какое-либо свойство описывается регрессионными уравнениями, содержащими в качестве независимых переменных дескрипторы, связанные со стерическими взаимодействиями, поляризуемостью и водородной связью. Для количественной оценки стерических взаимодействий используют значения молекулярных ван-дер-ваальсовых объемов V_{mc} . Дескриптор, называемый индексом поляризуемости (π_i), рассчитывается как поляризуемость, деленная на объем. При расчете водородной связи авторы выделяют две составляющие — кислотную и основную. Каждая из составляющих в свою очередь делится на ковалентную и электростатическую части. Принимается, что ковалентный вклад определяется энергией высшей занятой молекулярной орбитали (НОМО). Электростатический вклад в основность характеризуется величиной наибольшего формально отрицательного заряда (q_-) в молекуле (предполагается, что наиболее отрицательно заряженный атом должен иметь наибольшее взаимодействие с протоном соседней молекулы). Ковалентная часть кислотной составляющей (e_a) характеризуется энергией низшей незанятой молекулярной

орбитали (берется разность энергий LUMO для соединения и НОМО для воды, результат делится на сто). Меньшее значение ϵ_a указывает на большую тенденцию к образованию водородной связи (составляющая ϵ_b) с водой. Вклад электростатической части в кислотность определяется зарядом наиболее положительно заряженного атома водорода в молекуле (q_+).

При реализации этого подхода какое-либо свойство (P) можно описать следующим шестипараметрическим уравнением:

$$\lg P = \lg P_0 + \frac{AV_{mc}}{100} + B\pi_1 + C\epsilon_b + Dq_- + E\epsilon_a + Fq_+. \quad (49)$$

где A, B, C, D, E, F — коэффициенты.

Указанный подход был применен к описанию свойств и биологической активности химических соединений и дал формально неплохие результаты. Тем не менее представляется, что в данном случае использована упрощенная физическая модель. Кроме того, подобные многопараметрические регрессионные уравнения при небольших обучающих выборках не являются устойчивыми и их использование для предсказания свойств может привести к неверным результатам.

3. В настоящее время наиболее распространенной компьютерной программой, с помощью которой проводятся расчеты энергий межмолекулярных взаимодействий и используются соответствующие дескрипторы для установления взаимосвязи структура–свойство, является программа CoMFA (comparative molecular field analysis).³⁰⁷

Первоначальными рассчитываемыми параметрами являются энергии стерических и электростатических взаимодействий между рассматриваемым соединением и пробным атомом, «размещаемым» в различных узлах правильной трехмерной решетки. Размер этой решетки должен быть достаточно большим, чтобы в ней могла поместиться самая большая молекула рассматриваемой обучающей выборки. Размер ячейки такой решетки обычно принимается равным 2.0 Å, а иногда и 1.0 Å (чем меньше размер решетки, тем больше число дескрипторов и время их расчета). В качестве пробного атома сначала использовался sp^3 -гибридизованный атом углерода с зарядом +1.0, а затем различные атомы и группы (см., например, работу³⁰⁸). Очевидно, что в результате таких расчетов получается чрезвычайно большое число дескрипторов, во много раз превышающее число соединений обучающей выборки. При этом большая часть этих дескрипторов коррелируют друг с другом. В этих условиях создание устойчивой модели возможно лишь при использовании специальной процедуры. Метод PLS (partial least squares)³⁰⁹ позволяет отобрать ортогональные компоненты. В процессе создания устойчивых моделей число этих компонент можно оптимизировать, рассчитав коэффициенты корреляции для предсказательных моделей, получаемых процедурой cross-validation.

В течение десяти последних лет программа CoMFA постоянно совершенствовалась. Например, методологическим аспектам развития программы посвящены статьи^{310, 311} (описание нелинейной CoMFA, использование различных полупериодических методов оптимизации расчета геометрии и свойств соединений). Программа использовалась для изучения многих типов биологической активности различных соединений, например, при исследовании взаимосвязи структур 2-гетарилхинолин-4-аминов и их биологической активности против ВИЧ,³¹² структур фосфорных лигандов и их стерических и электронных эффектов,³¹³ структур разнообразных соединений и их ингибирующей активности,³¹⁴ стероидных ингибиторов,³¹⁵ мелатониновых рецепторных лигандов,³¹⁶ кальциевых блокаторов.³¹⁷

Однако программа CoMFA не лишена некоторых недостатков. Прежде всего необходимо отметить, что она до последнего времени не рассчитывала значения энергий водо-

родных связей, играющих определяющую роль во многих биологических процессах. Кроме того, практически невозможно на основе параметров, выступающих в качестве независимых переменных в конечном регрессионном уравнении, построить какую-либо содержательную физическую модель взаимосвязи структура–свойство, поскольку эти переменные представляют линейную комбинацию многих тысяч дескрипторов. К недостаткам программы необходимо отнести также невозможность ее использования владельцами персональных компьютеров.

4. Определенное сходство с CoMFA имеет программа GRID,²²⁵ в которой также используются различные по природе атомы и функциональные группировки в качестве пробных атомов для оценки межмолекулярных взаимодействий с исследуемыми макромолекулами, «помещенными» в трехмерную решетку. Существенно, что данная программа, в отличие от CoMFA, учитывает не только стерические и электростатические взаимодействия, но и водородные связи (E_{hb}). Вначале для этого использовалась направленно-зависимая 6–4-функция с широким энергетическим минимумом^{225, 318}

$$E_{hb} = \left(\frac{C}{d^6} - \frac{D}{d^4} \right) \cos^m Q. \quad (50)$$

где Q — угол ДНП (Д — донорный атом протеина, Н — атом водорода, П — пробный акцепторный атом); m обычно принимают равным 4.

Когда взаимодействуют два идентичных атома, табличные значения для C и D определяют межатомное разделение d_{min} . Если взаимодействуют разные атомы, используются геометрическое среднее их индивидуальных значений D и арифметическое среднее их разложений d_{min} . Позднее 6–4-функция была заменена на 8–6-функцию^{319–320}

$$E_r = \frac{C}{r^8} - \frac{D}{r^6}, \quad (51)$$

где r — расстояние между пробным атомом и каким-либо атомом рассматриваемого соединения. Энергия водородной связи представляется в этом случае следующей формулой:

$$E_{hb} = E_r E_\tau E_p, \quad (52)$$

где E_τ и E_p — угловые функции, учитывающие взаимное расположение в пространстве неподеленных электронных пар акцептора водородной связи и атома водорода.

С использованием программы GRID успешно построены модели различных биологических рецепторов.^{321–327}

5. В качестве альтернативы применения упомянутых выше программ CoMFA и GRID можно рассматривать подходы, использующие отображение пространственной структуры и способность к межмолекулярным взаимодействиям в виде спектра межатомных расстояний или межатомных взаимодействий.^{251, 270, 328–340} В этих случаях молекулярная структура представляется в форме кривой радиального рассеяния, которая легко может быть найдена для молекулы любой сложности, если известны ее декартовы координаты. Аналогичный подход применяется при изучении строения молекул методом газовой электронографии. Функция радиального распределения ($M(s)$) в приближении сферической симметрии атомов i и j и в отсутствие внутримолекулярных колебаний представляет собой спектр межатомных расстояний, интенсивность которого определяется природой атомов, находящихся на данном расстоянии друг от друга, и динамикой изменения этого расстояния в процессе внутримолекулярных колебаний

$$M(s) = \sum_{i=1}^N \sum_{j>i}^N C_{ij} \frac{\sin sr_{ij}}{sr_{ij}}, \quad (53)$$

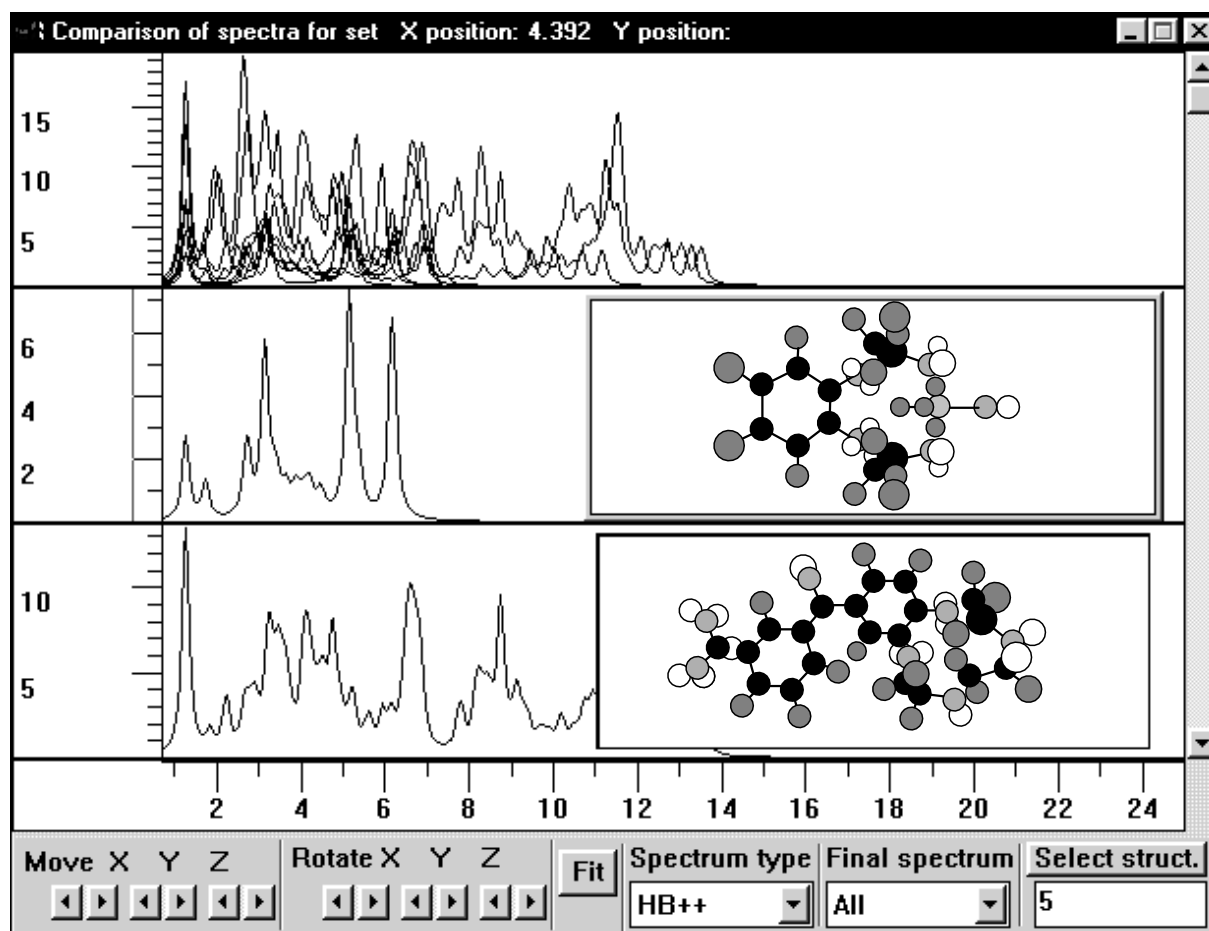


Рис. 4. Общий вид спектра взаимодействий акцепторов водородной связи в некоторых макроциклических полиэфирах.

где C_{ij} — коэффициент, характеризующий рассеивающую способность данной пары атомов, s — угловой параметр: $s = (4\pi/\lambda) \sin(Q/2)$ (λ — длина волны электрона, Q — угол рассеяния), r_{ij} — межатомное расстояние в молекуле.

Можно сказать, что функция радиального распределения, которая получается в процессе электронографического эксперимента, это то, что «видит» электрон, взаимодействующий с молекулой в акте рассеяния. Данная аналогия была распространена на случай взаимодействия рецептора с молекулами введенного в организм вещества.³²⁸ Известно, что взаимодействие рецептора происходит не со всей молекулой реагента, а с частями молекулы, комплементарными активным центрам рецептора. Определяющими факторами являются расстояния между активными центрами и способность к межмолекулярным взаимодействиям атомов, находящихся на этих расстояниях в молекулах реагента. Имеет место своеобразный процесс рассеяния молекул реагента на активных центрах рецептора, т.е. рецептор «видит» молекулу в образе функции радиального распределения. При таком подходе можно модифицировать функцию радиального распределения таким образом, чтобы количественно учесть способность атомов реагента к различным межмолекулярным взаимодействиям. Конкретное воплощение данного подхода осуществлено в работах^{251, 340} и в программном комплексе MOLTRA (Molecular Transform Analysis), с помощью которого рассчитываются спектры стерических взаимодействий атомов, взаимодействий положительно заряженных атомов между собой, отрицательно заряженных атомов между собой, положительно заряженных атомов с отрицательно заряженными, доноров водородной связи между собой, акцепторов водородной связи между

собой, доноров и акцепторов водородной связи. Каждый из этих спектров является суперпозицией полос, соответствующих парам атомов в молекуле. Положение максимумов полос соответствует равному расстоянию этих атомов, а площади под контурами полос пропорциональны произведениям рассматриваемых локальных характеристик для данной пары атомов.

На рис. 4 представлены спектры взаимодействий акцепторов водородной связи в некоторых макроциклических полиэфирах. Каждую точку представленного спектра можно рассматривать в качестве дескриптора межмолекулярных взаимодействий, описывающего взаимодействие акцепторов водородной связи на указанном расстоянии. В интервале расстояний 1.1–20.0 Å при шаге в 0.1 Å для каждого вида спектров получается 190 дескрипторов. Семь видов спектров позволяют описать возможные стерические и кулоновские взаимодействия, а также водородную связь 1340 дескрипторами.

Х. Необходимые качества дескрипторов

К настоящему времени описано множество дескрипторов. Понятно, что углубление представлений о молекулярной структуре побуждает к созданию ее новых моделей и новых дескрипторов, отражающих эти представления. Оправданы также попытки создания новых дескрипторов, которые при вводе в компьютерные программы лучше по сравнению с предыдущими дескрипторами сохраняют и используют информацию о молекулярной структуре.

Тем не менее вводить новые дескрипторы целесообразно только в двух случаях: если ни один из существующих

дескрипторов или их сочетание не обеспечивают создания устойчивых моделей структура–свойство в рассматриваемой обучающей выборке и если новый дескриптор обеспечивает значительное улучшение статистических критериев модели.⁹

Новые дескрипторы должны быть доступны для оценки их значений, удобны в использовании, универсальны, должны полно описывать данную структурную особенность и быть интерпретируемыми.³⁴¹

К сожалению, далеко не всегда указанные выше требования принимаются во внимание.

Эффективность дескрипторов разных классов, используемых в исследованиях взаимосвязей структура–биологическая активность, выявляется при сравнении статистических критериев моделей, построенных на этих дескрипторах. Например, в работах^{290, 293} проведено сравнение моделей классификации агонистов и антагонистов среди замещенных фенилэтиламинов на основе ряда топологических и физико-химических дескрипторов с использованием линейного дискриминантного анализа, дискретно-регрессионной модели²⁹⁰ и кластер-значимого анализа.³⁴² Наилучший результат во всех указанных типах классификации был получен при использовании дескрипторов водородной связи. Другие примеры приведены в работах^{14–16, 337, 343, 344}.

В настоящем обзоре описание дескрипторов проводилось в порядке возрастания содержания в них информации о молекулярной структуре. Однако это не означает, что использование более информативного дескриптора гарантирует лучшие результаты при установлении взаимосвязей структура–свойство. Так, существует множество примеров успешного применения топологических дескрипторов для расчета температур плавления и кипения, энтальпии испарения и других «простых» физических свойств в пределах узких рядов родственных соединений, различающихся, например, длиной цепи алкильного заместителя. В этих случаях статистические критерии регрессионных уравнений, в которых независимыми переменными являются более информативные дескрипторы, подчас уступают статистическим критериям уравнений с простыми дескрипторами, вследствие меньшей точности оценки величин более информативных дескрипторов и содержания в них «лишней» для данной выборки информации. Учитывая сказанное выше, можно сделать следующие выводы: одновременное использование уже существующих дескрипторов разного уровня информативности является оправданным; необходимы разработка новых дескрипторов и совершенствование процедур более точного расчета всех дескрипторов; необходим осознанный подбор дескрипторов для рассматриваемой выборки.[†]

Отметим, что всеобъемлющее описание молекулярной структуры с помощью какого-либо одного дескриптора практически невозможно. Один дескриптор может учесть преимущественно одну особенность электронной или пространственной структуры, либо один вид межмолекулярных взаимодействий. По этой причине описание молекулярной структуры несколькими дескрипторами и построение моделей структура–свойство, основанных на нескольких независимых переменных, представляется разумным.

При создании моделей структура–биологическая активность большое значение в отборе дескрипторов имеют современные представления о процессах в организме, происходящих с участием химических веществ. Например, в настоящее время считается, что сложный механизм воздействия вещества на организм состоит из трех стадий: фармацевтической, фармакокинетической и фармакодинамической.³⁴⁵ Фармацевтическая стадия связана с введением вещества в организм и включает процессы дезинтеграции

готовой формы. Фармакокинетическая стадия определяется транспортом вещества от места введения до конкретной биологической мишени и включает в себя процессы абсорбции, распределения, метаболизма и выделения. На этой стадии существенное значение имеет реакционная способность исходных соединений и их способность к комплексообразованию. Фармакодинамическая стадия характеризуется взаимодействием исходного вещества или его продуктов с биологическими мишенями, обуславливающими конкретный вид биологической активности. На этом этапе большое значение имеет трехмерная структура взаимодействующих партнеров.

Исключительная популярность трехпараметрового уравнения Ганча–Фуджиты связана прежде всего с использованием дескрипторов электронной структуры (константы Гамметта), стерических взаимодействий (стерические константы Тафта) и транспортных свойств (коэффициент распределения в системе октанол–вода). Подобные модели и соответствующие уравнения и в настоящее время успешно используются в молекулярном дизайне биологически активных веществ.^{346–349} Например, в работе²¹⁰ на основе трех дескрипторов получено уравнение для ряда бициклических соединений, описывающее их способность связывать мускариновый рецептор (см. Приложение, уравнение (П.23)). Один из дескрипторов характеризует транспортные возможности соединений, второй связан с растворимостью соединений в воде и их способностью образовывать водородные связи с донорными центрами рецептора, а третий дескриптор указывает на важную роль расстояния между двумя конкретными акцепторами водородной связи при образовании субстрат-рецепторного комплекса.

При создании моделей необходимо соблюдение и формальных критериев. С точки зрения статистики число рассматриваемых соединений должно во много раз превышать число дескрипторов. При этом дескрипторы должны быть достаточно информативными и в то же время между ними не должно быть взаимной корреляции. Достигается это различными процедурами, включая экспериментальный дизайн,³⁵⁰ метод SIMCA/PLS,^{351, 352} главнoкомпонентный и факторный анализы.³⁵³

XI. Заключение

В настоящее время дескрипторы молекулярной структуры являются неотъемлемой частью большинства исследований взаимосвязи структура–свойство (в том числе биологическая активность) и широко используются при оптимизации характеристик соединений в процессе молекулярного дизайна. В Приложении данного обзора приведены примеры успешного использования дескрипторов различного типа для количественного описания разнообразных свойств и видов биологического действия.^{354–371}

Следует подчеркнуть, что в конечном итоге используемые дескрипторы молекулярной структуры должны приводить к созданию моделей структура–свойство, удовлетворяющих определенным требованиям. Такие модели должны находиться в согласии с законами природы, быть достаточно простыми и иметь ясный физический смысл, кроме того, они должны быть устойчивыми и, следовательно, обладать предсказательной способностью, а также обеспечивать возможность идентификации связей, не заданных в явном виде при разработке моделей. Все эти вопросы затрагивают другой важный раздел компьютерного молекулярного дизайна — моделирование взаимосвязей структура–свойство — и требуют специального обсуждения.

XII. Приложение

Примеры успешного использования дескрипторов молекулярной структуры в регрессионных соотношениях структура–свойство (биологическая активность).

[†] Эксперт может опираться не только на общепринятые представления о механизме формирования рассматриваемого свойства, но и на еще не формализованные знания и даже интуицию.

1. Дескрипторы элементного уровня

Растворимость газов в воде, $\lg L^w$ (см.³⁵⁴).

$$\lg L^w = 42.37(\pm 1.11)\text{HDCA}(2) + 0.65(\pm 0.02)[2n(\text{N}) + m(\text{O})] - 0.16(\pm 0.02)\Delta E + 0.12(\pm 0.01)\text{PCWT}^E + 0.82(\pm 0.01)N_{\text{ring}} + 2.65(\pm 0.22), \quad (\text{П.1})$$

где $\text{HDCA}(2)$ — дескриптор водородной связи, $[2n(\text{N}) + m(\text{O})]$ — дескриптор, характеризующий число атомов азота и кислорода в молекулах изученных соединений, $\Delta E = E_{\text{LUMO}} - E_{\text{HOMO}}$, PCWT^E — электронный топологический индекс, N_{ring} — число ароматических колец; $N = 406$, $R = 0.971$, $S = 0.52$, $F = 1300$.

2. Топологические дескрипторы

Липофильность органических соединений (на основе топологических и информационно-топологических индексов), $\lg P$ (см.³⁵⁵).

$$\lg P = 0.19P_{10} - 1.46\text{IC}_0 + 1.09\text{CIC}_2 - 0.77\text{CIC}_3 - 1.36^6X_b + 5.34^0X^V - 3.41^1X^V + 0.55^4X^V - 0.41^3X_C^V + 1.10X_w - 0.17W - 5.60, \quad (\text{П.2})$$

$N = 219$, $R = 0.955$, $S = 0.35$, $F = 194$.

β -Блокаторная активность лекарств, pLD_{50} (см.³⁵⁶).

$$\text{pLD}_{50} = 1.79^1\chi - 1.81^1\chi^V + 0.36^2\chi + 2.37^4\chi_p^V - 2.76^4\chi_p + 1.56, \quad (\text{П.3})$$

где χ — индексы связности Кира и Хола; $N = 17$, $R = 0.905$, $S = 0.208$, $F = 10.1$.

3. Дескрипторы электронной структуры молекул

Связывание эстрогенного рецептора, pEC_{50} (см.³⁵⁷).

$$\text{pEC}_{50} = 3.38(\pm 0.12)D^2 + 3.22(\pm 0.20)\text{Gap} - 44.3(\pm 1.8), \quad (\text{П.4})$$

где D — топологический индекс Балабана, Gap — разность $E_{\text{LUMO}} - E_{\text{HOMO}}$; $N = 33$, $R = 0.99$, $S = 0.30$, $F = 595.0$.

Токсичность по отношению к *Daphnia magna*, EC_{50} (см.³⁵⁸).

$$\text{EC}_{50} = -3.63 \frac{V_{\text{mc}}}{100} - 45.8\pi + 3.71q_- - 2.0q_+ + 11.4, \quad (\text{П.5})$$

где V_{mc} — молекулярный объем, π — индекс поляризуемости, q_- — электростатическая основность, q_+ — электростатическая кислотность; $N = 38$, $R = 0.977$, $S = 0.37$, $F = 176$.

Функцидная активность фосфонатов, pEC_{50} (см.³⁵⁹).

$$\text{pEC}_{50} = 0.296(\pm 0.124)\sum \pi - 0.047(\pm 0.012)\text{MR}(R^2) + 0.543(\pm 0.331)\sum F' + 3.556(\pm 0.215), \quad (\text{П.6})$$

где π — константа гидрофобности, $\text{MR}(R^2)$ — молярная рефракция второго заместителя, F' — электронный параметр Свейна–Лаптона; $N = 57$, $R = 0.851$, $S = 0.290$, $F = 46.32$.

Связывание папаиновых лигандов *N*-мезилглицинфениловыми эфирами (см.²⁹⁷).

$$\lg \frac{1}{K_m} = 0.529(\pm 0.230)\text{MR} + 0.379(\pm 0.200)\sigma + 1.877(\pm 0.130), \quad (\text{П.7})$$

где σ — электронная константа; $N = 13$, $R = 0.935$, $S = 0.105$, $F = 34.51$.

Связывание папаиновых лигандов *N*-бензоилглицинфениловыми эфирами (см.²⁹⁷).

$$\lg \frac{1}{K_m} = 0.771(\pm 0.670)\text{MR} + 0.728(\pm 0.370)\sigma + 3.623(\pm 0.340), \quad (\text{П.8})$$

$N = 7$, $R = 0.971$, $S = 0.148$, $F = 32.85$.

4. Дескрипторы молекулярной формы

Ингибирование лигандного связывания, pIC_{50} (см.³⁶⁰).

$$\text{pIC}_{50} = 10.25(\pm 0.97)V + 1.43(\pm 0.56), \quad (\text{П.9})$$

где V — индекс формы; $N = 26$, $R = 0.912$, $S = 0.462$, $F = 111.1$.

Сладкость производных анилина, $\lg \text{RS}$ (см.³⁶¹).

$$\lg \text{RS} = 0.52L - 1.37W1 + 3.71, \quad (\text{П.10})$$

где L и $W1$ — параметры STERIMOL ; $N = 20$, $R = 0.90$, $S = 0.32$.

Психомиметическая активность производных фенил-2-аминопропана, $\lg \text{MU}$ (см.³⁶²).

$$\lg \text{MU} = 1.99(\pm 0.13) - 1.78(\pm 0.17)\text{MTD} + 0.93(\pm 0.10)S_3 - 0.77(\pm 0.12)q_4, \quad (\text{П.11})$$

где MTD — минимальная топологическая разница, S_3 — электронотопологический индекс атома 3 фенильного кольца, q_4 — заряд на атоме 4 фенильного кольца; $N = 49$, $R^2 = 0.885$, $S = 0.241$, $F = 84.7$.

Ингибирование гепатических рецепторов, pEC_{50} (см.¹⁶⁹).

$$\text{pEC}_{50} = 0.447\text{Tm} + 0.242\text{Ts} - 0.158\text{Vu} + 2.361, \quad (\text{П.12})$$

где Tm , Ts и Vu — ВЦИМ-дескрипторы; $N = 71$, $R = 0.914$, $S = 0.64$.

Ингибирование фотосистемы II, pIC_{50} (см.³⁶³).

$$\text{pIC}_{50} = 1.081 \lg P - 0.57\sigma(\text{Y,Al}) + 0.29L(\text{Y,Al}) - 0.95T(\text{Y,Al}) - 2.21T(\text{Y,Ar}) - 0.32W(\text{X}) + 1.02I_u + 8.16, \quad (\text{П.13})$$

где Y и X — заместители, Al — алкил, Ar — арил, L , W , T — дескрипторы формы, I_u — индикаторная переменная (1 для производных мочевины, 0 — для всех других соединений); $N = 69$, $R = 0.921$.

5. Дескрипторы межмолекулярных взаимодействий

Липофильность различных соединений, $\lg P$ (см.²⁰⁹).

$$\lg P = -0.01(\pm 0.08) + 0.249(\pm 0.005)\alpha - 0.68(\pm 0.02)\sum C_a, \quad (\text{П.14})$$

где α — молекулярная поляризуемость, $\sum C_a$ — сумма факторов акцепторной способности к водородной связи; $N = 234$, $R = 0.961$, $S = 0.34$, $F = 1383$.

Эффект протеинового связывания, $\lg \text{HSAI}$ (см.³⁶⁴).

$$\lg \text{HSAI} = 0.59(\pm 0.03)\lg P + 0.27(\pm 0.08), \quad (\text{П.15})$$

$N = 57$, $R = 0.92$, $S = 0.28$, $F = 283$.

Токсичность по отношению к *Daphnia*, $\lg(1/\text{IC}_{50})$ (см.³⁶⁵).

$$\lg \frac{1}{\text{IC}_{50}} = 0.809 \lg P + 1.64, \quad (\text{П.16})$$

$N = 23$, $R^2 = 0.84$, $S = 0.30$, $F = 118$.

Антибактериальная активность, pMIC (см.³⁶⁶).

$$\text{pMIC} = -0.35 \text{pK}_a + 0.68, \quad (\text{П.17})$$

$N = 8, R = 0.973, S = 0.063$

Лиганд-рецепторное связывание, ΔG (см.³⁶⁷).

$$\Delta G = 0.06(\pm 0.01) E_L - 0.05(\pm 0.02) \Delta E_{\text{solv}} + 7.74(\pm 0.57), \quad (\text{П.18})$$

где E_L — изменение внутримолекулярной энергии лиганда при связывании, ΔE_{solv} — изменение энергии сольватации при связывании; $N = 10, R = 0.92, F = 20.61$.

Ингибирующая способность арилзамещенных пиперазина и пиперидина, $\lg(1/\text{EC}_{50})$ (см.³⁶⁸).

$$\lg \frac{1}{\text{EC}_{50}} = 0.62(\pm 0.04) \lg P - 0.39(\pm 0.05) \sum C_a - 2.16(\pm 0.18), \quad (\text{П.19})$$

$N = 17, R^2 = 0.90$.

Токсичность карбоксильных кислот по отношению к *Xenopus embryos*, pLC_{50} (см.³⁵⁸).

$$\text{pLC}_{50} = 0.696 \lg P - 0.358 \text{pK}_a - 0.541, \quad (\text{П.20})$$

$N = 44, R = 0.947$.

Мембранная проницаемость (human red cell basal), $\lg \text{BP}$ (см.³⁶⁹).

$$\lg \text{BP} = -0.70(\pm 0.64) + 1.08(\pm 0.16) \sum C_d, \quad (\text{П.21})$$

$N = 10, R = 0.983, S = 0.433$.

Мембранная проницаемость (cells of the alga *Chara ceratophylla*), $\lg \text{Per}$ (см.³⁶⁹).

$$\lg \text{Per} = 0.83(\pm 0.57) + 0.59(\pm 0.12) \sum C_d, \quad (\text{П.22})$$

$N = 27, R = 0.903, S = 0.49$.

Связывание мускаринового рецептора бициклическими лигандами, $\lg K$ (см.²¹⁰).

$$\lg K = 1.28(\pm 0.15) \lg P - 0.09(\pm 0.06) \sum C_a + 0.27(\pm 0.08) \text{HBA5.3A} - 4.58(\pm 0.48), \quad (\text{П.23})$$

где HBA5.3A — дескриптор, характеризующий взаимодействие двух акцепторов водородной связи, находящихся на расстоянии 5.3 Å; $N = 27, R = 0.918, S = 0.38, F = 41.1$

6. Индикаторные дескрипторы

Связывание папаиновых лигандов *N*-метилглицин- и *N*-бензоил-глицинфениловыми эфирами, $\lg(1/K_m)$ (см.²⁹⁷). (Комбинация уравнений (П.7) и (П.8) с помощью индикаторной переменной I ($I = 1$ для мезиламидов и $I = 0$ для бензамидов).)

$$\lg \frac{1}{K_m} = 0.57(\pm 0.26) \text{MR} + 0.56(\pm 0.19) \sigma - 1.92(\pm 0.15) I + 3.74(\pm 0.17), \quad (\text{П.24})$$

$N = 20, R = 0.990, S = 0.148, F = 272.04$.

Ингибирование HIV-RT, pIC_{50} (см.³⁷⁰).

$$\text{pIC}_{50} = 1.47 I + 0.41 \mu + 2.69 \lg \text{MW} - 4.36, \quad (\text{П.25})$$

где I — индикаторная переменная, MW — молекулярная масса, $N = 14, R = 0.94, S = 0.42, F = 27.4$.

Цитотоксическая активность, pIC_{50} (см.³⁷¹).

$$\text{pIC}_{50} = 0.73(\pm 0.46) I_2' + 1.14(\pm 0.38) I_5' + 1.12(\pm 0.43) I_{xy} + 3.15, \quad (\text{П.26})$$

где I_2' , I_5' и I_{xy} — индикаторные переменные; $N = 22, R = 0.915, S = 0.32, F = 30.78$.

Литература

1. C.Hansch. *Drug Metabolism Rev.*, **15**, 1279 (1984-85)
2. C.Hansch, T.Fujita. *J. Am. Chem. Soc.*, **86**, 1616 (1964)
3. R.Lipnick, V.A.Filov. *Trends Pharm. Sci.*, **13**, 56 (1992)
4. Н.В.Лазарев. *Неэлектролиты*. ВММА, Ленинград, 1944
5. О.А.Раевский, А.М.Сапегин. *Успехи химии*, **57**, 1565 (1988)
6. S.Nikolic, N.Trinajstic. *Croat. Chem. Acta*, **70**, 777 (1997)
7. О.А.Раевский. *Рос. хим. журн.*, **39**, 109 (1995)
8. J.Tomasi. *J. Mol. Struct. (Theochem)*, **179**, 273 (1988)
9. M.Charton. *News. Int. QSAR Soc.*, **1**, 3 (1991)
10. C.Hansch. *Acc. Chem. Res.*, **26**, 147 (1993)
11. C.Hansch, D.Hoekman, H.Gao. *Chem. Rev.*, **96**, 1045 (1996)
12. E.J.Lien. *SAR. Side Effects and Drug Design*. Marcel Dekker, New York, 1987
13. E.J.Lien. *Chin. Pharm. J.*, **42**, 93 (1990)
14. P.C.Jurs, S.L.Dixon, L.Egolf. In *Chemometric Methods in Molecular Design. Vol. 2*. (Ed. H.Waterbeemd). VCH, Dusseldorf, 1995. P. 15
15. B.Testa, L.B.Kier. *Med. Res. Rev.*, **11**, 35 (1991)
16. Э.Стьюпер, У.Брюер, П.Джурс. *Машинный анализ связи химической структуры и биологической активности*. Мир, Москва, 1982. 235 с.
17. *Применение теории графов в химии*. (Под ред. Н.С.Зефинова, С.И.Кучанова). Наука, Новосибирск, 1988. 306 с.
18. М.И.Станкевич, И.В.Станкевич, Н.С.Зефинов. *Успехи химии*, **57**, 337 (1988)
19. Ю.Г.Папулов, В.Р.Розенфельд, Т.Г.Кеменова. *Молекулярные графы*. ТГУ, Тверь, 1990
20. Д.Рувре. В кн. *Химические приложения топологии и теории графов*. (Под ред. Р.Кинга). Мир, Москва, 1987. С. 181
21. S.S.Tratch, M.I.Stankevitch, M.S.Zefirov. *J. Comput. Chem.*, **11**, 899 (1990)
22. L.B.Kier, L.H.Hall. *Molecular Connectivity in Structure – Activity Analysis*. Wiley, New York, 1986
23. M.Randic, X.Guo, T.Oxley, H.Krishnapriyan. *J. Chem. Inf. Comput. Sci.*, **33**, 709 (1993)
24. M.Randic. *J. Am. Chem. Soc.*, **97**, 6609 (1975)
25. L.B.Kier, L.H.Hall. *J. Pharm. Sci.*, **64**, 1971 (1975)
26. L.B.Kier, L.H.Hall. *J. Pharm. Sci.*, **65**, 1806 (1976)
27. L.B.Kier, L.H.Hall. *Pharm. Res.*, **8**, 801 (1990)
28. L.H.Hall, B.Mohney, L.B.Kier. *Quantum Struct. – Act. Relat.*, **10**, 43 (1991)
29. L.H.Hall, L.B.Kier. *J. Chem. Inf. Comput. Sci.*, **35**, 1039 (1995)
30. L.H.Hall, L.B.Kier, B.B.Brown. *J. Chem. Inf. Comput. Sci.*, **35**, 1074 (1995)
31. A.K.Saxena. *Quantum Struct. – Act. Relat.*, **14**, 31 (1995)
32. A.K.Saxena. *Quantum Struct. – Act. Relat.*, **14**, 142 (1995)
33. D.L.Dowdy, T.E.McKone. *Environ. Toxicol. Chem.*, **16**, 2448 (1997)
34. L.H.Hall, C.T.Story. *J. Chem. Inf. Comput. Sci.*, **36**, 1004 (1996)
35. M.T.D.Cronin. *Toxicol. Vitro*, **10**, 103 (1996)
36. P.-Z.Lang, X.-F.Ma, G.-H.Lu, Y.Wang, Y.Bian. *Chemosphere*, **32**, 1547 (1996)
37. M.Soskis, D.Plavsic, N.Trinajstic. *J. Chem. Inf. Comput. Sci.*, **36**, 829 (1996)
38. R.Garcia-Domenech, F.J.Garcia-March, R.M.Soler, J.Galvez, G.M.Anton-Fos, J.V.de Julian-Ortiz. *Quantum Struct. – Act. Relat.*, **15**, 201 (1996)
39. R.R.A.Abon-Shaaban, H.A.Al-Khamees, H.S.Abou-Auda, A.P.Simonelli. *Pharm. Res.*, **13**, 129 (1996)
40. H.Kubinyi. *Quantum Struct. – Act. Relat.*, **14**, 149 (1995)
41. S.Nikolic, N.Trinajstic, Z.Mihalic. *Croat. Chem. Acta*, **68**, 105 (1995)
42. I.Lukovits. *J. Chem. Inf. Comput. Sci.*, **34**, 1079 (1994)
43. A.A.Dobrynin, A.A.Kochetova. *J. Chem. Inf. Comput. Sci.*, **34**, 1082 (1994)
44. M.Randic. *Stud. Phys. Theor. Chim.*, **54**, 101 (1988)
45. M.Randic. *Int. J. Quantum Chem. Quantum Biol. Symp.*, **15**, 201 (1988)

46. M.Randic, A.F.Kleiner, L.M.DeAlba. *J. Chem. Inf. Comput. Sci.*, **34**, 277 (1994)
47. M.Randic, M.Razinger. *J. Chem. Inf. Comput. Sci.*, **35**, 140 (1995)
48. O.G.Mekenyan, D.Peitchev, D.Bonchev, N.Trinajstic, I.Bangov. *Drug. Design*, **36**, 176 (1986)
49. M.V.Diudea, D.Horvath, A.Graovac. *J. Chem. Inf. Comput. Sci.*, **35**, 129 (1995)
50. L.B.Kier. *Quantum Struct. – Act. Relat.*, **4**, 109 (1985)
51. L.B.Kier. *Quantum Struct. – Act. Relat.*, **5**, 1 (1986)
52. L.B.Kier. *Quantum Struct. – Act. Relat.*, **5**, 7 (1986)
53. L.B.Kier. *Quantum Struct. – Act. Relat.*, **6**, 8 (1987)
54. L.B.Kier. *Quantum Struct. – Act. Relat.*, **6**, 117 (1987)
55. Y.-Y.Yao, Lu Xu, Y.-Q.Yang, X.-S.Yuan. *J. Chem. Inf. Comput. Sci.*, **33**, 590 (1993)
56. G.Rucker, C.Rucker. *J. Chem. Inf. Comput. Sci.*, **33**, 683 (1993)
57. H.P.Schultz, T.P.Schultz. *J. Chem. Inf. Comput. Sci.*, **33**, 240 (1993)
58. H.P.Schultz, E.B.Schultz, T.P.Schultz. *J. Chem. Inf. Comput. Sci.*, **33**, 863 (1993)
59. H.P.Schultz, E.B.Schultz, T.P.Schultz. *J. Chem. Inf. Comput. Sci.*, **34**, 1154 (1994)
60. M.V.Diudea, M.Topan, A.Graovac. *J. Chem. Inf. Comput. Sci.*, **34**, 1072 (1994)
61. T.Pisanski, A.Zithik, A.Graovac, A.Baumgartner. *J. Chem. Inf. Comput. Sci.*, **34**, 1090, (1994)
62. J.Galvez, R.Garcia, M.T.Salabert, R.M.Soler. *J. Chem. Inf. Comput. Sci.*, **34**, 520 (1994)
63. E.Estrada. *J. Chem. Inf. Comput. Sci.*, **35**, 701 (1995)
64. E.Estrada. *J. Chem. Inf. Comput. Sci.*, **35**, 708 (1995)
65. B.Llorente, N.Rivero, R.Carrasco, R.M.Martinez. *Quantum Struct. – Act. Relat.*, **13**, 419 (1994)
66. I.Lukovits. *Quantum Struct. – Act. Relat.*, **9**, 227 (1990)
67. G.G.Gash, J.J.Breen. *J. Chem. Inf. Comput. Sci.*, **33**, 275 (1993)
68. A.R.Katritzky, E.V.Gordeeva. *J. Chem. Inf. Comput. Sci.*, **33**, 835 (1993)
69. J.Galvez, R.Garcia-Domenech, J.V. de Julian-Ortiz, R.M.Soler. *J. Chem. Inf. Comput. Sci.*, **35**, 272 (1995)
70. B.Lucic, S.Nikolic, N.Trinajstic, D.Juretic, A.Juric. *Croat. Chem. Acta*, **68**, 435 (1995)
71. J.Niederfellner, D.Lenoir, G.Matuschek, F.Rehfeldt, H.Utschick, R.Bruggemann. *Quantum Struct. – Act. Relat.*, **16**, 38 (1997)
72. J.V.de Julian-Ortiz, R.Garcia-Domenech, J.Galvez Alvarez, R.Soler Roca, F.J.Garcia-March, G.M.Anton-Fos. *J. Chromatogr., A*, **719**, 37 (1996)
73. В.Магнусон, Д.Харрис, С.Бейсак. В кн. *Химические приложения топологии и теории графов*. (Под ред. Р.Кинга). Мир, Москва, 1987. С. 206
74. G.Bazylak. *J. Planar. Chromatogr.*, **7**, 202 (1994)
75. S.C.Basak, S.Bertelsen, G.D.Grunwald. *J. Chem. Inf. Comput. Sci.*, **34**, 27 (1994)
76. W.J.Boecklen, G.J.Niemi. *SAR QSAR Environ. Res.*, **2**, 79 (1994)
77. A.T.Balaban, S.C.Basak, T.Colburn, G.D.Grunwald. *J. Chem. Inf. Comput. Sci.*, **34**, 1118 (1994)
78. S.C.Basak, G.D.Grunwald. *J. Chem. Inf. Comput. Sci.*, **35**, 366 (1995)
79. T.Wieland. *Arzheim.-Forsch. Drug. Res.*, **46**, 273 (1996)
80. A.T.Balaban, T.S.Balaban. *J. Chim. Phys. Phys.-Chim. Biol.*, **89**, 1735 (1992)
81. W.J.Wisswesser. *The Wisswesser Line-Formula Chemical Notation*. McGraw-Hill, New York, 1968
82. В.В.Авидон. *Хим.-фармацевт. журн.*, **8**, 22 (1974)
83. В.В.Авидон, В.С.Аралович. *НТИ. Сер. 2. Информ. процессы и системы*, (5), 26 (1975)
84. D.Weininger. *J. Chem. Inf. Comput. Sci.*, **28**, 31 (1988)
85. D.Weininger. *J. Chem. Inf. Comput. Sci.*, **29**, 97 (1989)
86. D.Weininger. *J. Chem. Inf. Comput. Sci.*, **30**, 237 (1990)
87. A.Dreifahl, M.Reinhard. *J. Chem. Inf. Comput. Sci.*, **33**, 886 (1993)
88. K.K.Agarwal, H.L.Gelernter. *J. Chem. Inf. Comput. Sci.*, **34**, 463 (1994)
89. В.В.Авидон, В.Е.Голендер, А.Б.Розенблит. В кн. *Методы представления и обработки структурной информации для анализа связи структура – активность*. ИОС АН ЛатвССР, Рига, 1981. С. 8
90. А.Б.Розенблит, В.Е.Голендер. *Логико-комбинаторные методы в конструировании лекарств*. Зинатне, Рига, 1983. 451 с.
91. О.А.Раевский, А.М.Сапегин, В.В.Чистяков, В.П.Соловьев, Н.С.Зефирова. *Координац. химия*, **16**, 1175 (1990)
92. О.А.Раевский, В.В.Авидон, В.П.Новиков. *Хим.-фармацевт. журн.*, **16**, 968 (1982)
93. V.P.Avidon, O.A.Raevsky, V.S.Arolovich, R.I.Zhdanov. In *Bioactive Spin Labels*. (Ed. R.I.Zhdanov). Springer-Verlag, Berlin, 1992. С. 549
94. Г.М.Баренбойм, А.Г.Маленков. *Биологически активные вещества*, Наука, Москва, 1986. С. 90
95. I.Paeva, E.Golovinsky. *Quantum Struct. – Act. Relat.*, **9**, 16 (1990)
96. A.Ormerod, P.Willett, D.Bawden. *Quantum Struct. – Act. Relat.*, **8**, 115 (1989)
97. A.Ormerod, P.Willett, D.Bawden. *Quantum Struct. – Act. Relat.*, **9**, 302 (1990)
98. W.J.Streich, R.Franke. *Quantum Struct. – Act. Relat.*, **4**, 13 (1985)
99. R.Franke, W.J.Streich. *Quantum Struct. – Act. Relat.*, **4**, 51 (1985)
100. R.Franke, W.J.Streich. *Quantum Struct. – Act. Relat.*, **4**, 63 (1985)
101. V.K.Gombar, E.P.Jaeger, P.C.Jurs. *Quantum Struct. – Act. Relat.*, **7**, 225 (1988)
102. R.A.Lewis, A.C.Good, S.D.Pickett. In *Computer-Assisted Lead Finding and Optimization. Current Tools for Medicinal Chemistry*. (Eds H.van de Waterbeemd, B.Testa, G.Folkers). VCH, Weinheim, 1997. P. 137
103. J.E.Dubois, M.Loukianoff, C.Mercier. *J. Chim. Phys.*, **89**, 1493 (1992)
104. C.Mercier, G.Trouiller, J.-E.Dubois. *Quantum Struct. – Act. Relat.*, **7**, 88 (1990)
105. E.J.Martin, J.M.Blaney, M.A.Siani, D.C.Spellmeyer, A.K.Wong, W.H.Moos. *J. Med. Chem.*, **38**, 1431 (1995)
106. H.-J.Bohm. In *Computer-Assisted Lead Finding and Optimization. Current Tools for Medicinal Chemistry*. (Eds H.van de Waterbeemd, B.Testa, G.Folkers). VCH, Weinheim, 1997. P. 125
107. M.Karelson, V.S.Lobanov, A.R.Katritzsky. *Chem. Rev.*, **96**, 1027 (1996)
108. O.Kikuchi. *Quantum Struct. – Act. Relat.*, **6**, 179 (1987)
109. H.Bauknecht, A.Zell, H.Bayer, P.Levi, M.Wagener, J.Sadowski, J.Gasteiger. *J. Chem. Inf. Comput. Sci.*, **36**, 1205 (1996)
110. S.Karabunarliev, O.G.Mekenyan, W.Karcher, C.L.Russom, S.P.Bradbury. *Quantum Struct. – Act. Relat.*, **15**, 302 (1996)
111. M.A.Ordocica, M.L.Velazquez, J.G.Ordocica, J.L.Escobar, P.A.Lehmann. *Quantum Struct. – Act. Relat.*, **12**, 246 (1993)
112. F.T.Hatch, M.E.Colvin, E.T.Seidl. *Environ. Mol. Mutagen.*, **27**, 314 (1996)
113. A. Inoue, T.Kawai, M.Iimura, H.Sugimoto, H.Kawakami. *J. Med. Chem.*, **39**, 4460 (1996)
114. M.T.D.Cronin, T.W.Schultz. *Chemosphere*, **32**, 1453 (1996)
115. M.Cocchi, M.C.Menziani, P.G.De Benrdrtti, G.Cruciani. *Chemometr. Intell. Lab. Syst.*, **14**, 209 (1992)
116. K.Tuppurainen. *J. Mol. Struct. (Theochem)*, **306**, 49 (1994)
117. B.Balasubramanian. *SAR QSAR Environ. Res.*, **2**, 59 (1994)
118. R.G.Parr, R.Pearson. *J. Am. Chem. Soc.*, **105**, 7512 (1983)
119. Z.Medven, H.Gusten, A.Subljic. *J. Chemometr.*, **10**, 135 (1996)
120. C.Graham, R.Gealy, O.T.Macina, M.H.Karol, H.S.Rosenkranz. *Quantum Struct. – Act. Relat.*, **15**, 224 (1996)
121. W.H.J.Vaes, C.Hamwijk, E.U.Ramos, H.J.Mverhaar, J.L.M.Hermens. *Anal. Chem.*, **68**, 4458 (1996)
122. D.F.W.Lewis, B.G.Lake. *Toxicology*, **125**, 31 (1998)
123. R.C.Young, C.Graham, M.L.Roantree. In *QSAR in Drug Design and Toxicology*. (Eds D.Hadzy, B.Jerman-Blazic). Elsevier, Amsterdam, 1987. P. 91
124. T.I.Oprea, L.Rurunczi, E.E.Moret. In *Trends in QSAR and Molecular Modelling 92*. (Ed. C.G.Wermuth). ESCOM, Leiden, 1993. P. 398
125. S.Gibson, R.McGuire, D.C.Rees. *J. Med. Chem.*, **39**, 4065 (1996)
126. D.Chapman. *J. Comput.-Aid. Mol. Design*, **10**, 501 (1996)
127. J.H.van Drie. *J. Chem. Inf. Comput. Sci.*, **37**, 38 (1997)
128. A.Verloop, W.Hoogenstraaten, J.Tipker. In *Drug Design*. (Ed. J.Ariens). Academic Press, New York, 1976. P. 165
129. J.Jurgens, H.Schwedletzky, P.Heising, J.K.Seydel, B.Wiedemann, U.Holzgrabe. *Arch. Pharm. Med. Chem.*, **329**, 179 (1996)
130. A.Tanaka, H.Fujiwara. *J. Med. Chem.*, **39**, 5017 (1996)
131. W.Draber. *Z. Naturforsch., C. Biosci.*, **51**, 1 (1996)
132. M.M.Davila-Jimenez, M.P.Elizalde-Gonzales, S.Garcia-Diaz. *J. Liq. Chrom. Rel. Technol.*, **21**, 1629 (1998)

133. J.E.Amoore. *Ann. N. Y. Acad. Sci.*, **116**, 457 (1964)
134. R.H.Rohrbaugh, P.C.Jurs. *Anal. Chim. Acta*, **199**, 99 (1987)
135. G.M.Crippen. *J. Med. Chem.*, **22**, 988 (1979)
136. G.M.Crippen. *J. Med. Chem.*, **23**, 599 (1980)
137. G.M.Crippen. *J. Med. Chem.*, **24**, 198 (1981)
138. A.K.Ghose, G.M.Crippen. *J. Med. Chem.*, **25**, 892 (1982)
139. A.K.Ghose, G.M.Crippen. *J. Med. Chem.*, **27**, 901 (1984)
140. A.K.Ghose, G.M.Crippen. *J. Med. Chem.*, **28**, 333 (1985)
141. A.K.Ghose, G.M.Crippen, G.R.Revankar, P.A.McKernan, D.F.Smee, R.K.Robins. *J. Med. Chem.*, **32**, 746 (1989)
142. G.M.Crippen. *J. Comput. Chem.*, **8**, 943 (1987)
143. G.M.Crippen. *J. Comput. Chem.*, **10**, 673 (1989)
144. M.P.Bradley, G.M.Crippen. *J. Med. Chem.*, **36**, 4171 (1993)
145. L.G.Bouli, G.M.Crippen, H.A.Barton, H.Kwon, M.A.Marletta. *J. Med. Chem.*, **33**, 771 (1990)
146. A.J.Hopfinger. *J. Am. Chem. Soc.*, **102**, 7196 (1980)
147. M.Mabilia, R.A.Pearlstein, A.J.Hopfinger. *Eur. J. Med. Chem.*, **20**, 1133 (1985)
148. M.G.Koehler, K.Rowberg-Schaefer, A.J.Hopfinger. *Arch. Biochem. Biophys.*, **266**, 152 (1988)
149. B.J.Burke, A.J.Hopfinger. *J. Med. Chem.*, **33**, 274 (1990)
150. A.J.Hopfinger, B.J.Burke, W.J.Dunn III. *J. Med. Chem.*, **37**, 3768 (1994)
151. B.J.Burke, W.J.Dunn III, A.J.Hopfinger. *J. Med. Chem.*, **37**, 3775 (1994)
152. P.I.Nagy, J.S.Tokarski, A.J.Hopfinger. *J. Chem. Inf. Comput. Sci.*, **34**, 1190 (1994)
153. J.S.Tokarski, A.J.Hopfinger. *J. Med. Chem.*, **37**, 3639 (1994)
154. K.-B.Rhyu, H.C.Patel, A.J.Hopfinger. *J. Chem. Inf. Comput. Sci.*, **35**, 771 (1995)
155. J.W.Dunn, A.J.Hopfinger, C.Catana, C.Duraiswami. *J. Med. Chem.*, **39**, 4825 (1996)
156. A.M.Doweyko. *J. Med. Chem.*, **31**, 1396 (1988)
157. A.M.Doweyko. *J. Math. Chem.*, **7**, 273 (1991)
158. A.M.Doweyko, W.B.Mattes. *Biochemistry*, **31**, 9388 (1992)
159. A.M.Doweyko. *J. Med. Chem.*, **37**, 1769 (1994)
160. J.J.Kaminsky, A.M.Doweyko. *J. Med. Chem.*, **40**, 427 (1997)
161. A.N.Jain, T.G.Dietterich, R.H.Lathrop, D.Chapman, R.E.Critchlow, B.E.Bauer, T.A.Webster, T.Lozañj-Perez. *J. Comput.-Aid. Mol. Design*, **8**, 635 (1994)
162. A.N.Jain, K.Koile, D.Chapman. *J. Med. Chem.*, **37**, 2315 (1994)
163. A.N.Jain, N.L.Harris, J.Y.Park. *J. Med. Chem.*, **38**, 1295 (1995)
164. R.Todeschini, M.Lasagni, E.Marengo. *J. Chemometr.*, **8**, 263 (1994)
165. R.Todeschini, P.Gramatica, R.Provenzani. *Chemometr. Intell. Lab. Syst.*, **27**, 221 (1995)
166. R.Todeschini, M.Vigni, R.Provenzani, A.Finizio, P.Gramatica. *Chemosphere*, **32**, 1527 (1996)
167. R.Todeschini, C.Bettoli, G.Giurin, P.Gramatica, P.Miana, E.Argese. *Chemosphere*, **33**, 71 (1996)
168. R.Todeschini, P.Gramatica. *Quantum Struct. – Act. Relat.*, **16**, 113 (1997)
169. R.Todeschini, P.Gramatica. *Quantum Struct. – Act. Relat.*, **16**, 120 (1997)
170. P.G.Mezey. *J. Chem. Inf. Comput. Sci.*, **32**, 650 (1992)
171. P.G.Mezey. *Shape in Chemistry. Introduction to Molecular Shape and Topology*. VCH, New York, 1992
172. P.D.Walker, P.G.Mezey. *J. Am. Chem. Soc.*, **115**, 12423 (1993)
173. P.D.Walker, P.G.Mezey. *J. Am. Chem. Soc.*, **116**, 12022 (1994)
174. P.G.Mezey, Z.Zimpel, P.Warburton, P.D.Walker, D.G.Irvine, S.L.Dixon, B.Greenberg. *J. Chem. Inf. Comput. Sci.*, **36**, 602 (1996)
175. A.Murcko, M.A.Murcko. *J. Med. Chem.*, **38**, 4953 (1995)
176. M.Charton. In *Lipophilicity in Drug Action and Toxicology*. (Eds V.Pliska, B.Testa, H.Waterbeemd). VCH, Weinheim, 1996. P. 387
177. C.Dauwe, B.Sellergren. *J. Chromatogr., A*, **753**, 191 (1996)
178. P.G.Tratnyek. In *Perspectives in Environmental Chemistry*. (Ed. D.L.Macalady). Oxford University Press, New York, 1998. P. 167
179. L.Quadri, A.Cerri, P.Ferrari, E.Folpini, M.Mabilia, P.Melloni. *J. Med. Chem.*, **39**, 3385 (1996)
180. Z.Nikolovska-Coleska, L.Suturkova, K.Dorevski, A.Krbavcic, T.Solmajer. *Quantum Struct. – Act. Relat.*, **17**, 7 (1998)
181. L.P.Hammet. *J. Am. Chem. Soc.*, **59**, 66 (1937)
182. M.Charton. *Prog. Phys. Org. Chem.*, **13**, 119 (1981)
183. C.Hanch, A.Leo, R.W.Taft. *Chem. Rev.*, **91**, 165 (1991)
184. C.Hansch, A.Leo, D.Hoekman. *Exploring QSAR. Hydrophobic, Electronic and Steric Constants*. American Chemical Society, Washington, DC, 1995
185. G.Schuurman. *Quantum Struct. – Act. Relat.*, **9**, 326 (1990)
186. O.G.Mekenyan, G.D.Veith. *SAR QSAR Environ. Res.*, **2**, 129 (1994)
187. S.Karabunarliev, O.G.Mekenyan, W.Karcher, C.L.Russom, S.P.Bradbury. *Quantum Struct. – Act. Relat.*, **15**, 311 (1996)
188. B.W.Clare, C.T.Supuran. *J. Pharm. Sci.*, **83**, 768 (1994)
189. K.J.Miller. *J. Am. Chem. Soc.*, **112**, 8533 (1990)
190. R.M.Smith, P.S.Pearlman. In *QSAR and Molecular Modelling. Concepts, Computational Tools and Biological Applications*. (Eds F.Sanz, J.Giraldo, F.Manaut). J.R.Prous Science Publishers, Barcelona, 1995. P. 222
191. J.Jaerv, K.Sak, M.Eller, P.Ek, A.Engstroem, I.Engstroem. *Bioorg. Chem.*, **24**, 159 (1996)
192. J.D.Hirst. *J. Med. Chem.*, **39**, 3526 (1996)
193. P.Sjoberg. In *Computer-Assisted Lead Finding and Optimization. Current Tools for Medicinal Chemistry*. (Eds H.van de Waterbeemd, B.Testa, G.Folker). VCH, Weinheim, 1997. P. 83
194. S.-S.So, M.Karplus. *J. Med. Chem.*, **39**, 5246 (1996)
195. V.Pliska, B.Testa, H.Waterbeemd. In *Lipophilicity in Drug Action and Toxicology*. (Eds V.Pliska, B.Testa, H.Waterbeemd). VCH, Weinheim, 1996. P. 1
196. A.Avdeef. In *Lipophilicity in Drug Action and Toxicology*. (Eds V.Pliska, B.Testa, H.Waterbeemd). VCH, Weinheim, 1996. P. 109
197. J.C.Dearden, G.M.Bresnen. *Quantum Struct. – Act. Relat.*, **7**, 133 (1988)
198. L.-G.Danielsson, Y.-H.Zhang. *Trends Anal. Chem.*, **15**, 188 (1996)
199. M.L.C.Montanari, C.A.Montanari, D.P.Veloso, A.E.Beezer, J.C.Mitchell, P.L.O.Volpe. *Quantum Struct. – Act. Relat.*, **17**, 102 (1998)
200. V.N.Viswanadhan, A.K.Ghose, G.R.Revankar, R.K.Robins. *J. Chem. Inf. Comput. Sci.*, **29**, 163 (1989)
201. R.F.Rekker, R.Mannhold. *Calculation of Drug Lipophilicity*. VCH, Weinheim, 1992
202. G.Klopman, J.-K.Wang, M.J.Dimayuga. *J. Chem. Inf. Comput. Sci.*, **34**, 752 (1994)
203. E.Kellogg, D.Abraham. *J. Comput.-Aid. Mol. Design*, **5**, 545 (1991)
204. H.Waterbeemd, H.Karajannis, H.Kansy, D.Obrecht, K.Muller, C.Lehmann. In *QSAR and Molecular Modelling. Concepts, Computational Tools and Biological Applications*. (Eds F.Sanz, J.Giraldo, F.Manaut). J.R.Prous Science Publishers, Barcelona, 1995. P. 78
205. N.Bodor, M.J.Huang. *J. Pharm. Sci.*, **81**, 272 (1992)
206. M.H.Abraham, H.S.Chadra. In *Lipophilicity in Drug Action and Toxicology*. (Eds V.Pliska, B.Testa, H.Waterbeemd). VCH, Weinheim, 1996. P. 311
207. H.Waterbeemd, B.Testa. *Adv. Drug. Res.*, **16**, 87 (1987)
208. N.El Tayar, B.Testa. In *Trends in QSAR and Molecular Modelling 92*. (Ed. C.G.Wermuth). ESCOM, Leiden, 1993. P. 101
209. O.A.Raevsky, K.-J.Shaper, J.K.Seydel. *Quantum Struct. – Act. Relat.*, **14**, 433 (1995)
210. O.A.Raevsky. In *Computer-Aided Lead Finding and Optimization. Current Tools for Medicinal Chemistry*. (Eds H.Waterbeemd, B.Testa, G.Folkers). VCH, Wienheim, 1997. P. 367
211. H.Waterbeemd, R.Mannhold. In *Lipophilicity in Drug Action and Toxicology*. (Eds V.Pliska, B.Testa, H.Waterbeemd). VCH, Weinheim, 1996. P. 401
212. R.Mannhold, K.Dross. *Quantum Struct. – Act. Relat.*, **15**, 403 (1996)
213. H.Waterbeemd, R.Mannhold. *Quantum Struct. – Act. Relat.*, **15**, 410 (1996)
214. P.A.Carrupt, P.Gaillard, F.Billois, P.Weber, B.Testa, C.Meyer, S.Perez. In *Lipophilicity in Drug Action and Toxicology*. (Eds V.Pliska, B.Testa, H.Waterbeemd). VCH, Weinheim, 1996. P. 195
215. C.Altomare, A.Carotti, V.Carta, S.Kneubihler, P.A.Carrupt, B.Testa. In *QSAR and Molecular Modelling. Concepts, Computational Tools and Biological Applications*. (Eds F.Sanz, J.Giraldo, F.Manaut). J.R.Prous Science Publishers, Barcelona, 1995. P. 463

216. M.Laguerre, M.Saux, A.Carpy. In *QSAR and Molecular Modelling. Concepts, Computational Tools and Biological Applications*. (Eds F.Sanz, J.Giraldo, F.Manaut). J.R.Prous Science Publishers, Barcelona, 1995. P. 588
217. I.Rozas, M.Martin. *J. Chem. Inf. Comput. Sci.*, **36**, 872 (1996)
218. S.Immel, F.W.Lichtenhaler. *Leibigs. Ann. Chem.*, **39** (1996)
219. T.Fujita. *Quantum Struct. – Act. Relat.*, **16**, 107 (1997)
220. B.Testa, P.-A.Carrupt, P.Gaillard, F.Billois, P.Weber. *Pharm. Res.*, **13**, 335 (1996)
221. C.Hansch, D.Hoekman, A.Leo, L.Zhang, P.Li. *Toxicol. Lett.*, **79**, 45 (1995)
222. M.T.D.Cronin, J.C.Dearden. *Quantum Struct. – Act. Relat.*, **14**, 1 (1995)
223. M.T.D.Cronin, J.C.Dearden. *Quantum Struct. – Act. Relat.*, **14**, 117 (1995)
224. M.T.D.Cronin, J.C.Dearden. *Quantum Struct. – Act. Relat.*, **14**, 329 (1995)
225. P.J.Goodford. *J. Med. Chem.*, **28**, 849 (1985)
226. R.W.Taft. In *Steric Effects in Organic Chemistry*. (Ed. M.S.Newman). Wiley, New York, 1956. P. 556
227. M.Charton. *Environ. Health Perspect.*, **61**, 229 (1985)
228. M.Charton. In *Similarity Models in Organic Chemistry, Biochemistry and Related Fields*. (Eds R.I.Zalewski, T.M.Krygowski, J.Shorter). Elsevier, Amsterdam, 1991. P. 629
229. K.H.Kim, Y.C.Martin. In *QSAR, Rational Approaches to the Design of Bioactive Compounds*. (Eds C.Silipo, A.Vittoria). Elsevier, Amsterdam, 1991. P. 151
230. F.Sanz, M.Martin, F.Lapena, F.Manaut. *Quantum Struct. – Act. Relat.*, **5**, 54 (1986)
231. I.Alkorta, H.O.Villar. *J. Med. Chem.*, **37**, 210 (1994)
232. D.B.Kireev, O.A.Raevsky, V.I.Fetisov. In *Trends in QSAR and Molecular Modelling 92*. (Ed. C.G.Wermuth). ESCOM, Leiden, 1993. P. 331
233. L.Bonati, E.Fraschini, M.Lasagni, E.P.Madoni, D.Pitea. *J. Mol. Struct. (Theochem)*, **340**, 83 (1995)
234. R.Bureau, J.C.Lancelot, H.Prunier, S.Rault. *Quantum Struct. – Act. Relat.*, **15**, 373 (1996)
235. A.K.Bhattacharjee, J.M.Karle. *J. Med. Chem.*, **39**, 4622 (1996)
236. R.Carbo, B.Calabuig. *Int. J. Quantum Chem.*, **42**, 1681 (1992)
237. R.Carbo, B.Calabuig. *J. Chem. Inf. Comput. Sci.*, **32**, 600 (1992)
238. E.Besalu, L.Amat, X.Fradera, R.Carbo. In *QSAR and Molecular Modelling. Concepts, Computational Tools and Biological Applications*. (Eds F.Sanz, J.Giraldo, F.Manaut). J.R.Prous Science Publishers, Barcelona, 1995. P. 396
239. X.Fradera, L.Amat, E.Besalu, R.Carbo-Dorca. *Quantum Struct. – Act. Relat.*, **16**, 25 (1997)
240. R.Carbo, L.Leyde, M.Arnau. *Int. J. Quantum Chem.*, **17**, 1185 (1980)
241. A.C.Good, S.-S.So, W.G.Richard. *J. Med. Chem.*, **36**, 433 (1993)
242. R.Carbo, B.Calabuig. *J. Mol. Struct. (Theochem)*, **254**, 517 (1992)
243. R.Carbo, E.Besalu, B.Calabuig, L.Vera. *Adv. Quant. Chem.*, **25**, 253 (1994)
244. D.J.Wild, P.Willett. *J. Chem. Inf. Comput. Sci.*, **36**, 159 (1996)
245. Л.Полинг, П.Полинг. *Химия*. Мир, Москва, 1978. С. 458
246. P.Magee. *Sci. Total Environ.*, **109/110**, 155 (1991)
247. N.Froloff, A.Faurion, P.M.Leod. *Chem. Sens.*, **21**, 425 (1996)
248. E.Forgas, T.Cserhati. *J. Liq. Chromatogr. Relat. Technol.*, **19**, 1849 (1996)
249. О.А.Раевский. *Теорет. и эксперим. химия*, **22**, 450 (1986)
250. О.А.Раевский. *Успехи химии*, **59**, 375 (1990)
251. О.А.Раевский. *J. Phys. Org. Chem.*, **10**, 405 (1997)
252. О.А.Раевский, В.Ю.Григорьев, В.П.Соловьев, И.В.Мартынов. *Докл. АН СССР*, **299**, 1433 (1988)
253. R.S.Drago, B.V.Wayland. *J. Am. Chem. Soc.*, **87**, 3571 (1965)
254. R.S.Drago. *Structure and Bonding*. Springer-Verlag, Heidelberg, 1973. 227 p.
255. M.K.Kroeger, R.S.Drago. *J. Am. Chem. Soc.*, **103**, 3250 (1981)
256. А.В.Иогансен. *Теорет. и эксперим. химия*, **7**, 302 (1971)
257. О.А.Раевский, В.П.Новиков. *Хим.-фармацевт. журн.*, **16**, 583 (1982)
258. И.В.Мартынов, О.А.Раевский. *Докл. АН СССР*, **265**, 664 (1982)
259. И.В.Мартынов, О.А.Раевский. *Вест. АН СССР*, (7), 93 (1983)
260. И.В.Мартынов, О.А.Раевский. *Журн. Всесоюз. хим. о-ва им. Д.И.Менделеева*, **28**, (1983)
261. О.А.Раевский, М.М.Гилязов, Я.А.Левин. *Журн. общ. химии*, **53**, 563 (1983)
262. О.А.Раевский, М.М.Гилязов, Я.А.Левин. *Журн. общ. химии*, **53**, 1720 (1983)
263. О.А.Раевский, М.М.Гилязов, Я.А.Левин. *Журн. общ. химии*, **53**, 1724 (1983)
264. О.А.Раевский, В.Ю.Григорьев, В.П.Соловьев. *Хим.-фармацевт. журн.*, **18**, 578 (1984)
265. Т.Г.Валеева, Я.А.Левин, О.А.Раевский. *Журн. общ. химии*, **55**, 1590 (1985)
266. О.А.Раевский, В.Ю.Григорьев, В.П.Соловьев, А.И.Иванов, В.Б.Соколов, И.В.Мартынов. *Журн. общ. химии*, **57**, 2073 (1987)
267. О.А.Раевский, В.Ю.Григорьев, В.П.Соловьев, И.В.Мартынов. *Докл. АН СССР*, **298**, 1166 (1988)
268. О.А.Раевский, А.Ф.Солотнов, В.П.Соловьев. *Журн. общ. химии*, **57**, 1240 (1987)
269. А.М.Сапегин, О.А.Раевский, В.В.Чистяков, И.В.Мартынов. *Хим.-фармацевт. журн.*, **21**, 578 (1987)
270. О.А.Raevsky. In *QSAR in Drug Design and Toxicology*. (Eds D.Hadzy, V.Jerman-Blazic). Elsevier, Amsterdam, 1987. P. 31
271. О.А.Раевский, В.Ю.Григорьев, В.П.Соловьев. *Хим.-фармацевт. журн.*, **23**, 1294 (1989)
272. О.А.Raevsky, V.Ju.Grigo'ev, D.B.Kireev, N.S.Zefirov. *Quantum Struct. – Act. Relat.*, **11**, 49 (1992)
273. R.W.Taft, D.Gurka, L.Loris, P.R.Schleyer, J.W.Rakshys. *J. Am. Chem. Soc.*, **91**, 4801 (1969)
274. J.-L.M.Abboud, L.Bellon. *Ann. Chim.*, **5**, 63 (1970)
275. M.H.Abraham, P.E.Grellier, D.V.Prior, R.W.Taft, J.J.Moris, P.J.Taylor, C.Lourence, M.Bertholot, R.M.Doherty, M.J.Kamlet, J.-L.M.Abboud, K.Sraidi, G.Guiheneuf. *J. Am. Chem. Soc.*, **110**, 8534 (1988)
276. M.H.Abraham. *Chem. Soc. Rev.*, **22**, 73 (1993)
277. R.W.Taft, W.J.Shuely, R.H.Doherty, M.J.Kamlet. *J. Org. Chem.*, **53**, 1737 (1988)
278. P.-C.Maria, J.-F.Gal, J.de Franceschi, E.Fargin. *J. Am. Chem. Soc.*, **109**, 483 (1987)
279. M.H.Abraham. In *Quantitative Treatments of Solute/Solvent Interactions. Theoretical and Computational Chemistry. Vol. 1*. (Eds P.Polizer, J.S.Murray). Elsevier, Amsterdam, 1994. P. 83
280. О.А.Raevsky, V.Ju.Grigo'ev, V.P.Solov'ev. In *QSAR Rational Approaches to the Design of Bioactive Compounds*. (Eds C.Silipo, A.Vittoria). Elsevier, Amsterdam, 1991. P. 135
281. О.А.Раевский, А.М.Сапегин. *Хим.-фармацевт. журн.*, **24**, 43 (1990)
282. О.А.Раевский, А.М.Сапегин, В.В.Чистяков, А.Н.Раздольский, Н.С.Зефирев. *Докл. АН СССР*, **309**, 623 (1989)
283. H.-J.Schneider, V.Rüdiger, O.Raevsky. *J. Org. Chem.*, **58**, 3648 (1993)
284. H.-J.Schneider, T.Blatter, A.Eliseev, V.Rüdiger, O.A.Raevsky. *Pure Appl. Chem.*, **65**, 2329 (1993)
285. О.А.Раевский, Г.И.Шнайдер. *Журн. орг. химии*, **31**, 1793 (1995)
286. V.P.Solov'ev, N.N.Strakhova, O.A.Raevsky, V.Rudiger, H.-J.Schneider. *J. Org. Chem.*, **61**, 5221 (1996)
287. О.А.Раевский, В.В.Чистяков, Р.С.Агабекян, А.М.Сапегин, Н.С.Зефирев. *Биоорг. химия*, **16**, 1509 (1990)
288. I.I.Kitova, O.A.Raevsky, V.G.Blinova, N.S.Zefirov. In *QSAR and Molecular Modelling. Concepts, Computational Tools and Biological Applications*. (Eds F.Sanz, J.Giraldo, F.Manaut). J.R.Prous Science Publishers, Barcelona, 1995. P.160
289. С.О.Бачурин, Н.В.Лукоянов, Л.Н.Петрова, Л.С.Соляков, С.Е.Ткаченко, О.А.Раевский. *Докл. АН*, **346**, 549 (1996)
290. O.Raevsky, A.Sapegin, N.Zefirov. *Quantum Struct. – Act. Relat.*, **13**, 412 (1994)
291. А.Н.Раздольский, А.М.Сапегин, О.А.Раевский, Е.В.Шавелева, Г.И.Смирнова, Н.Г.Рожкова. *Хим.-фармацевт. журн.*, **26** (9–10), 92 (1992)
292. У.З.Мирходжаев, Н.В.Лукоянов, И.И.Китова, Х.Ф.Абдулаев, О.А.Раевский, А.К.Ташмухамедова. *Узб. биол. журн.*, (1), 8 (1991)
293. O.Raevsky, V.Ju.Grigo'ev, D.B.Kireev, N.S.Zefirov. *J. Chim. Phys. Phys.-Chim. Biol.*, **89**, 1747(1992)
294. О.А.Raevsky, V.Ju.Grigo'ev, E.Mednikova. In *Trends in QSAR and Molecular Modelling 92*. (Ed. C.G.Wermuth). ESCOM, Leiden, 1993. P. 116

295. H. Waterbeemd, G. Gamenisch, G. Folkers, O. A. Raevsky. *Quant. Struct.-Act. Relat.*, **15**, 480 (1996)
296. J. Li, Y. Zhang, A. J. Dallas, P. W. Carr. *J. Chromatogr.*, **550**, 101 (1991)
297. H. Kubinyi. In *Burger's Medicinal Chemistry and Drug Discovery*. (Ed. M. E. Wolff), Wiley, New York, 1995. P. 497
298. J. J. Hostynek, P. S. Magee. *Quantum Struct. - Act. Relat.*, **16**, 473 (1997)
299. J. Halova, O. Strouf, P. Z. A. Sochozova, N. Uchida, T. Yuzyri, K. Sakakibara, M. Hirota. *Quantum Struct. - Act. Relat.*, **17**, 37 (1998)
300. W. H. Streng. *Drug Design Today*, **2**, 415 (1997)
301. L. Y. Wilson, G. R. Famini. *J. Med. Chem.*, **34**, 1668 (1991)
302. G. R. Famini, V. P. Loubnev, E. K. Frykman, L. Y. Wilson. *Quantum Struct. - Act. Relat.*, **17**, 558 (1998)
303. G. R. Famini, W. P. Ashman, A. P. Miciewicz, L. Y. Wilson. *Quantum Struct. - Act. Relat.*, **11**, 162 (1992)
304. G. R. Famini, L. Y. Wilson, C. A. Penski. *J. Phys. Org. Chem.*, **5**, 395 (1992)
305. A. H. Lowrey, C. J. Cramer, J. J. Urban, G. R. Famini. *Comput. Chem.*, **19**, 209 (1995)
306. N. A. Chester, G. R. Famini, M. V. Haley, C. W. Kurnas, P. A. Sterling, L. Y. Wilson. *ACS Symp. Ser.*, **643**, 191 (1996)
307. R. D. Cramer III, D. E. Patterson, J. D. Bunce. *J. Am. Chem. Soc.*, **110**, 5959 (1988)
308. R. J. Vaz. *Quantum Struct. - Act. Relat.*, **16**, 303 (1997)
309. W. Lindberg, J. A. Persson, S. Wold. *Anal. Chem.*, **55**, 643 (1983)
310. K. Hasegawa, T. Kimura, K. Funatsu. *Quantum Struct. - Act. Relat.*, **16**, 219 (1997)
311. C. Navajas, A. Poso, K. Tuppurainen, J. Gynther. *Quantum Struct. - Act. Relat.*, **15**, 189 (1996)
312. D. B. Kireev, J. R. Chretien, O. A. Raevsky. *Eur. J. Med. Chem.*, **30**, 395 (1995)
313. W. E. Steinmetz. *Quantum Struct. - Act. Relat.*, **15**, 1 (1996)
314. M. Akamatsu, Y. Ozoe, T. Ueno, T. Fujita, K. Mochida, T. Nakamura, F. Matsumura. *Pest. Sci.*, **49**, 319 (1997)
315. T. I. Oprea, A. E. Garcia. *J. Comput.-Aid. Mol. Design*, **10**, 186 (1996)
316. S. Sicsic, I. Serraz, J. Andrieux, B. Bremont, M. Mathe-Allanmat, A. Poncet, S. Shen, M. Langlois. *J. Med. Chem.*, **40**, 739 (1997)
317. F. Corelli, F. Manetti, A. Tafi, G. Campiani, V. Nacci, M. Botta. *J. Med. Chem.*, **40**, 125 (1997)
318. D. N. A. Boobbyer, P. J. Goodford, P. M. McWhinnie, R. C. Wade. *J. Med. Chem.*, **32**, 1083 (1989)
319. R. C. Wade, K. J. Clark, P. J. Goodford. *J. Med. Chem.*, **36**, 140 (1993)
320. R. C. Wade, P. J. Goodford. *J. Med. Chem.*, **36**, 148 (1993)
321. G. Cruciani, P. J. Goodford. *J. Mol. Graph.*, **12**, 116 (1994)
322. A. M. Davis, N. P. Gensmantel, E. Johansson, D. P. Marriot. *J. Med. Chem.*, **37**, 963 (1994)
323. E. E. Polymeropoulos, N. Hofgen. *Quantum Struct. - Act. Relat.*, **16**, 231 (1997)
324. T. Langer. *Quantum Struct. - Act. Relat.*, **15**, 469 (1996)
325. M. Pastor, G. Cruciani. *J. Med. Chem.*, **38**, 4637 (1995)
326. M. Itzstein, J. C. Dyason, S. W. Oliver, H. F. White, W.-Y. Wu, G. B. Kok, M. S. Pegg. *J. Med. Chem.*, **39**, 388 (1996)
327. S. Clementi, G. Cruciani, P. Fifi, D. Riganelli, R. Valigi. *Quantum Struct. - Act. Relat.*, **15**, 108 (1996)
328. В. П. Новиков, О. А. Раевский. *Хим.-фармацевт. журн.*, **16**, 574 (1982)
329. I. Csorvassy, L. Tozser. In *QSAR. Rational Approaches to the Design of Bioactive Compounds*. (Eds C. Silipo, A. Vittoria). Elsevier, Amsterdam, 1991. P. 193
330. J. W. King, R. Kassel, B. B. King. *Int. J. Quantum Chem. Quantum Biol. Symp.*, **17**, 27 (1990)
331. J. W. King, R. J. Kassel. *Int. J. Quantum Chem. Quantum Biol. Symp.*, **18**, 289 (1991)
332. J. W. King, R. J. Kassel. *Int. J. Quantum Chem. Quantum Biol. Symp.*, **19**, 179 (1992)
333. J. W. King. *Int. J. Quantum Chem. Quantum Biol. Symp.*, **20**, 139 (1993)
334. G. R. Famini, R. J. Kassel, J. W. King, L. Y. Wilson. *Quantum Struct. - Act. Relat.*, **10**, 344 (1991)
335. J. W. King. *Int. J. Quantum Chem. Quantum Biol. Symp.*, **21**, 209 (1994)
336. Б. И. Кучкаев, Б. А. Князев, В. К. Курочкин, В. А. Петрунин. *Хим.-фармацевт. журн.*, **26** (6), 60 (1992)
337. Б. И. Кучкаев, Б. А. Князев, В. К. Курочкин, В. А. Петрунин. *Хим.-фармацевт. журн.*, **31** (8), 30 (1997)
338. J. Schuur, P. Selzer, J. Gasteiger. *J. Chem. Inf. Comput. Sci.*, **36**, 334 (1996)
339. J. Gasteiger, J. Sadowski, J. Schuur, P. Selzer, L. Steinhauer, V. Steinhauer. *J. Chem. Inf. Comput. Sci.*, **36**, 1030 (1996)
340. O. Raevsky, L. Dolmatova, V. Grigor'ev, I. Lisyansky, S. Bondarev. In *QSAR and Molecular Modelling. Concepts, Computational Tools and Biological Applications*. (Eds F. Sanz, J. Giraldo, F. Manaut). J. R. Prous Science Publishers, Barcelona, 1995. P. 241
341. R. Franke. *Theoretical Drug Design Method*. Springer-Verlag, Berlin, 1984. 412 p.
342. J. McFarland, D. Gan. *Drug Inform. J.*, **24**, 705 (1990)
343. J. McFarland, D. Gan. *Quantum Struct. - Act. Relat.*, **13**, 11 (1994)
344. H. Matter. *J. Med. Chem.*, **40**, 1219 (1997)
345. E. J. Ariens, A. M. Simonis. *Top. Curr. Chem.*, **52**, 1 (1974)
346. K.-t. Chung, L. Kirkovsky, A. Kirkovsky, W. P. Purcell. *Mutat. Res.*, **387**, 1 (1997)
347. C. A. Montanari, M. S. Tute. *Quantum Struct. - Act. Relat.*, **16**, 480 (1997)
348. Y. Nezu, N. Wada, F. Yoshida, M. Miyazawa, T. Shimizu, T. Fujita. *Pestic. Sci.*, **52**, 343 (1998)
349. H. Kubinyi. *Drug Design Today*, **2**, 457 (1997)
350. V. Austel. In *Chemometric Methods in Molecular Design. Vol. 2*. (Ed. H. Waterbeemd). VCH, Dusseldorf, 1995. P. 49
351. W. Dunn, S. Wold. In *Chemometric Methods in Molecular Design. Vol. 2*. (Ed. H. Waterbeemd). VCH, Dusseldorf, 1995. P. 179
352. S. Wold. In *Chemometric Methods in Molecular Design. Vol. 2*. (Ed. H. Waterbeemd). VCH, Dusseldorf, 1995. P. 195
353. R. Franke, A. Gruska. In *Chemometric Methods in Molecular Design. Vol. 2*. (Ed. H. Waterbeemd). VCH, Dusseldorf, 1995. P. 115
354. A. R. Katritzky, L. Mu, M. Karelson. *J. Chem. Inf. Comput. Sci.*, **36**, 1162 (1996)
355. S. C. Basak, B. D. Gute, G. D. Grundwald. *J. Chem. Inf. Comput. Sci.*, **36**, 1054 (1996)
356. R. Carcia-Domenich, C. de Grigoria Alapont, J. V. de Julian-Ortiz, J. Galvez, L. Popa. *Bioorg. Med. Chem. Lett.*, **7**, 567 (1997)
357. S. P. Bradbury, O. G. Mekenyan, G. T. Ankley. *Environ. Toxicol. Chem.*, **15**, 1945 (1996)
358. D. A. Dawson, T. W. Schultz, R. S. Hunter. *Teratogen Carcinogen Mutagen*, **16**, 109 (1996)
359. N. K. Roy, E. S. J. Nidiry, K. Vasu, S. Bedi, B. Lalljee, B. Singh. *J. Agric. Food Chem.*, **44**, 3971 (1996)
360. A. Cappelli, M. Anzini, S. Vomero, P. G. De Benedetti, M. C. Menziani, G. Giorgi, C. Manzoni. *J. Med. Chem.*, **40**, 2910 (1997)
361. J. Polanski, J. Gasteiger, M. Wagoner, J. Sadowski. *Quantum Struct. - Act. Relat.*, **17**, 27 (1998)
362. M. Mracec, S. Muresan, M. Mracec, Z. Simon, G. Naray-Szabo. *Quantum Struct. - Act. Relat.*, **16**, 459 (1997)
363. R. Vaz. *Weed. Sci.*, **44**, 718 (1996)
364. M. Rowley, J. J. Kulagowski, A. P. Watt, D. Rathbone, G. I. Stevenson, R. W. Carling, R. Baker, G. R. Marshall, J. A. Kemp, A. C. Foster, S. Grimwood, R. Hargreaves, C. Hurley, K. L. Saywell, M. D. Tricklebank, P. D. Leeson. *J. Med. Chem.*, **40**, 4053 (1997)
365. Y. H. Zhao, M. T. D. Cronin, J. C. Dearden. *Quantum Struct. - Act. Relat.*, **16**, 131 (1997)
366. M. J. B. Mengelers, P. E. Hougee, L. H. M. Janssen, A. S. J. P. A. M. van Miert. *J. Vet. Pharmacol. Ther.*, **20**, 276 (1997)
367. J. S. Tokarski, A. J. Hopfinger. *J. Chem. Inf. Comput. Sci.*, **37**, 792 (1997)
368. P. Chiba, M. Hitzler, E. Richter, M. Huber, C. Tmej, E. Giovagnoni, G. Ecker. *Quantum Struct. - Act. Relat.*, **16**, 361 (1997)
369. O. A. Raevsky, K.-J. Schaper. *Eur. J. Med. Chem.*, **33**, 799 (1998)
370. K. C. S. C. Liu, S.-S. Lee, M.-T. Lin, C.-W. Chang, C.-L. Liu, J.-Y. Hsu, S. Ren, E. J. Lien. *Med. Chem. Res.*, **7**, 168 (1997)
371. R. Garg, S. P. Gupta. *J. Enzyme Inhib.*, **11**, 171 (1997)

DESCRIPTORS OF MOLECULAR STRUCTURE IN COMPUTER-AIDED DESIGN OF BIOLOGICALLY ACTIVE COMPOUNDS

O.A.Raevsky

Institute of Physiologically Active Compounds, Russian Academy of Sciences

142432 Chernogolovka, Moscow Region, Russian Federation, Fax +7(095)913-2113

The current state of description of molecular structure in computer-aided molecular design of biologically active substances by means of descriptors is analysed. The information contents of descriptors increases in the following sequence: element-level descriptors—structural formulae descriptors—electronic structure descriptors—molecular shape descriptors—intermolecular interaction descriptors. Every subsequent class of descriptors normally covers information contained in the previous-level ones. It is emphasised that it is practically impossible to describe all the features of a molecular structure in terms of any single class of descriptors. It is recommended to optimise the number of descriptors used by means of appropriate statistical procedures and characteristics of structure—property models based on these descriptors.

Bibliography — 371 references.

Received 16th November 1998