

УДК 544.431.121.2

**МАТРИЦЫ СМЕЖНОСТИ И ГРАФЫ ХИМИЧЕСКИХ ПРЕВРАЩЕНИЙ**

© 2008 Л.А. Грибов\*, В.А. Дементьев, И.В. Михайлов

*Институт геохимии и аналитической химии им. В.И. Вернадского РАН, Москва**Статья поступила 29 мая 2007 г.*

Громадное множество молекулярных объектов и химических превращений исходного продукта, оценить которые даже грубо на основании интуитивных соображений не представляется возможным, наводит на мысль о необходимости разработки простых приемов своеобразного мониторинга реакций. В работе впервые ставится такая проблема и обосновывается возможный путь ее решения.

**Ключевые слова:** мониторинг реакций, матрица смежности, пути реакций.

**ПОСТАНОВКА ЗАДАЧИ**

Необычность постановки самой проблемы требует некоторого исторического экскурса с тем, чтобы все дальнейшее стало понятным.

В конце 60-х годов минувшего века появились первые исследования, положившие начало формированию нового научного направления — математической химии, основанной на формальной логике [ 1 ]. В дальнейшем это привело, во-первых, к созданию так называемых экспертных систем для выявления молекулярных структур ранее вообще неизвестных соединений [ 2 ], а, во-вторых, к развитию также формализованных методов поиска корреляций структура—свойства и априорному выделению структур, желаемыми свойствами обладающих — направление QSAR (качественные соотношения структура—свойства).

Появился целый ряд обобщающих монографий (см., например, [ 3—5 ]) и даже журналов — Journal of Computer Chemistry, Journal of Molecular Modeling, QSAR and Combinatorial Science и др.

Уже на самой ранней стадии выяснилось, что необходимо решить задачу генерации всех возможных молекулярных структур, отвечающих заданной брутто-формуле. То, что одной и той же брутто-формуле может отвечать довольно большое число структурных изомеров, было известно давно. Ошеломляющим оказалось число таких структур. Так, оценки "сверху" для объектов с брутто-формулой  $C_6H_6O$  приводят к числу 2237, с брутто-формулой  $C_8H_6O$  к числу 120427, с брутто-формулой  $C_{10}H_{10}O$  к числу 7288723. Даже для простого случая  $C_6H_6$  получается 217 изомеров! Конечно, сам способ оценки базируется на использовании набора дискретных правил и не учитывает деталей. Поэтому цифры получаются сильно завышенными. Однако есть все основания считать, что если при формальной генерации изомерных структур число их получается равным  $10^6$  для, в общем-то, не очень сложных объектов, то несколько десятков тысяч структур вполне могут существовать!

Эти результаты приводят к выводу, что, несмотря на длительную историю развития химии как науки и на то, что сейчас описано около  $2 \cdot 10^7$  молекул, все равно мы находимся на краю безбрежного океана! Необходимость с самого начала учитывать гигантское число гипотез и потребовало развития средств, широту поиска уменьшающих. Можно возразить, что в химии

---

\* E-mail: gribov@geokhi.ru

уже созданного достаточно, чтобы занять следующие поколения химиков. Стоит ли, поэтому, в безбрежный-то океан выплывать?

Наверное, все-таки стоит. Во-первых, можно считать уже доказанным, что именно в результате локальных цепочечных изомер-изомерных преобразований происходит передача энергии и информации в сложных молекулярных структурах (супра- и биомолекулы и др.). Осознается реальная возможность создания молекулярных информационно-преобразующих наноструктур [6, 7]. Во-вторых, все больший интерес вызывают медленно протекающие (месяцы, годы и более) реакции, играющие большую роль в экологических и геохимических процессах. Химический опыт как в том, так и другом случаях практически отсутствует. Следовательно возникает проблема разработки методов своеобразного химического мониторинга, позволяющего, пусть и грубо, оценить возможные "реакционные связи" во множестве взаимопревращающихся структур (включая димерные) и описать эти связи в форме матрицы смежности и разветвленного графа (дерева) наиболее вероятных химических превращений. Этому и посвящена настоящая работа.

### ПУТИ РЕШЕНИЯ ПРОБЛЕМЫ

Будем рассуждать следующим образом. Для простоты ограничимся молекулярным объектом, обладающим многими структурно-изомерными формами. Пусть эти изомерные формы достаточно стабильны, т.е. ведут себя как устойчивые молекулы. Как в классической, так и в квантовой физике такая устойчивость может существовать только тогда, когда система находится в соответствующей "яме". Реально наблюдаемый переход от одной структурно-изомерной формы молекулярного объекта к другой в теоретической модели связывается с переходом из одной потенциальной «ямы» в соседнюю. При этом не принципиально, что является причиной такого перехода. Сама возможность изомеризации определяется связью между индивидуальными состояниями системы в "ямах".

В квантовой механике эта связь возникает, если существует заметно отличное от нуля перекрытие волновых функций, отвечающих состояниям подсистем, и резонанс уровней энергии, при котором достигается максимальное "смешивание" комбинирующих состояний, что и создает условия, необходимые для резонансного туннелирования.

Если ограничиться не слишком большими внешними воздействиями, необходимыми для стимуляции соответствующих преобразований, то получим, что "смешивание" должно происходить между состояниями подсистем, расположенными по энергиям не очень уж далеко от минимумов ям. Поскольку каждая из волновых функций в основном локализована внутри своей ямы, то отсюда непосредственно следует, что химическое превращение возможно только тогда, когда минимумы ям в некотором пространстве не чрезмерно удалены друг от друга.

Высказанные соображения сразу указывают направление дальнейших шагов: надо, во-первых, как-то охарактеризовать сами "ямы", а, во-вторых, найти сдвиг одной "ямы" по отношению к другой. Первое легко сделать, если решить задачу о колебаниях атомов структуры для каждой "ямы" и найти систему нормальных координат. В самом деле, именно в системе нормальных координат полностью отражаются как геометрические, так и динамические характеристики молекулярного объекта. Если с самого начала ограничить всю задачу именно мониторингом и не требовать высокой точности, то решения необходимых колебательных задач может быть выполнено в автоматическом режиме с использованием приближенных величин, сведенных в банки данных о геометрических параметрах молекул и силовых постоянных.

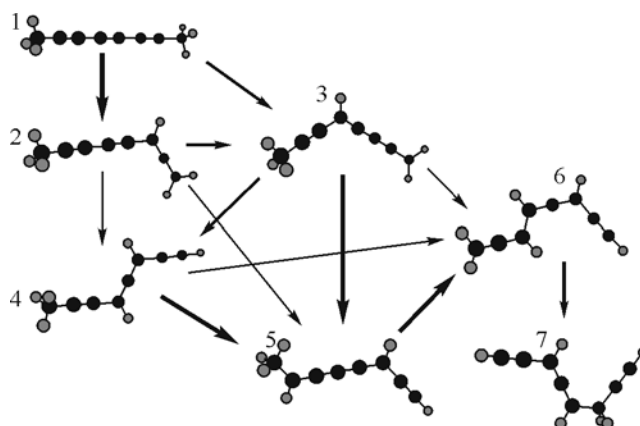
Вся необходимая для этого техника хорошо разработана (см., например, [8, 9]). Время расчета для одной достаточно сложной структуры даже на ПК составляет доли секунды. Этого вполне достаточно для работы с очень большим числом структур.

Зная характеристики "ям" для отдельных структур, можно оценить сдвиг "ям" комбинирующих состояний относительно друг друга. Для этого надо найти соотношение между совокупностями нормальных координат первой структуры  $Q_1$  и второй —  $Q_2$ .

В матричной форме это соотношение, найденное также довольно давно [10], имеет вид:

$$Q_2 = " + b.$$

Пути взаимопревращения изомерных структур с общей брутто-формулой  $C_8H_6$ . Степень "жирности" стрелок характеризует вероятности переходов



Здесь  $Q_1$  и  $Q_2$  — матрицы-столбцы;  $A$  — квадратная матрица поворотов и изменений масштабов координат и  $b$  — матрица-столбец (вектор) в многомерном пространстве  $Q_1$  сдвига искомого дна одной "ямы" по отношению к другой.

Сейчас авторами статьи созданы программы, позволяющие найти как матрицу  $A$ , так и вектор сдвига  $b$  для молекул любой сложности также за очень малое время при простом и, в принципе, допускающем полную автоматизацию управления. Тем самым, проблема "размещения" "ям" изомерных структур в виде точек в многомерном пространстве нормальных координат исходного изомера решается полностью.

Количественные оценки вероятностей химических превращений могут быть проведены методами, описанными в [ 11 ]. В общем случае вероятность реакции определяется интегралом перекрытия  $S_{(1)}^{(2)}$  собственных функций резонирующих состояний комбинирующих подструктур. Этот интеграл равен  $S_{(1)}^{(2)} = k \exp(-b^2)$ . Здесь  $k < 1$  — коэффициент, зависящий от характеристик (параметров кривизны) "ям" и  $b^2$  — норма вектора  $b$  соответствующих волновых функций. Для оценки "сверху" можно принять  $k = 1$ . Если теперь значения  $\exp(-b^2)$  использовать в качестве элементов матрицы, порядок следования столбцов и строк которой принят соответствующим выбранным номерам изомеров в общем их множестве, то получается матрица, которую можно назвать матрицей смежности для совокупности химических превращений. Если, далее, распределить такие значения по трем интервалам (реакции быстрые (1), средние (2) и медленные (3)), то можно построить систему графов.

Линии, соединяющие номера изомерных структур, будут иметь смысл ребер графа, а диагональные элементы матрицы — смысл вершин графа. В качестве примера приведена матрица смежности для изомер-изомерных превращений ряда простых соединений с общей брутто-формулой  $C_8H_6$  и вид соответствующего графа (см. рисунок).

	2	3	4	5	6	7
1	9,21 $1,0 \cdot 10^{-4}$	15,25 $2,38 \cdot 10^{-7}$	41,16 $1,33 \cdot 10^{-18}$	37,26 $6,88 \cdot 10^{-17}$	43,03 $2,05 \cdot 10^{-19}$	35,2 $5,16 \cdot 10^{-16}$
2		9,12 $1,1 \cdot 10^{-4}$	21,3 $5,62 \cdot 10^{-10}$	20,01 $2,04 \cdot 10^{-9}$	39,7 $2,16 \cdot 10^{-18}$	45,36 $2,0 \cdot 10^{-20}$
3			19,08 $5,17 \cdot 10^{-9}$	10,55 $2,62 \cdot 10^{-5}$	21,81 $3,37 \cdot 10^{-10}$	42,75 $1,08 \cdot 10^{-19}$
4				10,28 $3,43 \cdot 10^{-5}$	23,33 $7,38 \cdot 10^{-11}$	34,41 $1,14 \cdot 10^{-15}$
5					9,39 $8,36 \cdot 10^{-5}$	39,91 $4,65 \cdot 10^{-18}$
6						20,91 $3,04 \cdot 10^{-9}$

Последовательность строк и столбцов в матрице смежности отвечает изомерным структурам, показанным на рисунке. В результате можно получить наглядную и информативную картину возможных химических превращений и выбрать наиболее вероятные пути преобразования исходной структуры в любую из желаемых. Важно, что, хотя для построения матрицы смежности и требуются не дискретные правила, а вычисления, однако практическая их реализация вполне может быть доведена до высокой степени автоматизации и не требует больших усилий и времени. Следовательно возникает, перспектива создания такой экспертной системы для поиска путей реакций, в которой предлагаемый подход может составить основу первого блока, осуществляющего грубую фильтрацию исходных гипотез.

Работа выполнена при финансовой поддержке гранта Президента Российской Федерации (№ НШ-5203.2006.3).

#### СПИСОК ЛИТЕРАТУРЫ

1. *Химические приложения теории графов и топологии*. Ред. Р. Кинг – М.: Мир, 1987.
2. *Gribov L.A., Elyashberg M.E.* // *J. Mol. Struct.* – 1970. – **5**. – P. 179 – 198.
3. *Nendza M.* *Structure-Activity Relationships in Environmental Sciences.* – Dordrecht: Kluwer, 1997.
4. *Miertus S., Fassina G.* *Combinatorial Chemistry and Technology.* – N.Y.: Marcel Dekker, 1999.
5. *Jung G.* *Combinatorial Chemistry: Synthesis, Analysis, Screening.* – Weinheim: Wiley-VCH, 2000.
6. *Грибов Л.А.* // *Известия АН (сер. хим.)*. – 2002. – № 2. – С. 213 – 219.
7. *Грибов Л.А.* // *Вестник РАН*. – 2002. – **72**, № 7. – С. 611 – 617.
8. *Волькенштейн М.В., Грибов Л.А., Ельяшевич М.А., Степанов Б.И.* *Колебания молекул.* – М.: Наука, 1972.
9. *Грибов Л.А., Дементьев В.А.* *Методы и алгоритмы вычислений в теории колебательных спектров молекул.* – М.: Наука, 1981.
10. *Грибов Л.А.* // *Докл. АН СССР*. – 1987. – **292**, № 5. – С. 1161 – 1165.
11. *Грибов Л.А., Баранов В.И.* *Теория и методы расчета молекулярных процессов: спектры, химические превращения и молекулярная логика.* – М.: КомКнига, 2006.