

Идея такого алгоритма основывается на дополнении предыдущего алгоритма таким образом, чтобы при поиске он учитывал дополнительные связи между вершинами графа. Для этого в элемент очереди обхода необходимо добавить информацию об уже пройденных вершинах (т. е. уже находящихся в очереди обхода), связь которых с данным элементом очереди нужно проверить. Например, для замкнутого фрагмента, представленного на рис. 7, очередь обхода должна выглядеть так, как показано на рис. 8.

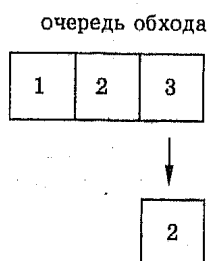


Рис. 8. Очередь обхода для замкнутого фрагмента

При поиске для атома 3 фрагмента совпадающего с ним атома молекулы V_i нужно дополнительно проверить, входит ли в окружение атома молекулы V_i такой атом, к которому уже привязан атом 2 фрагмента. Тем самым проверяется наличие связи между атомами 2 и 3 фрагмента у соответствующих им атомах в молекуле.

Очевидно, что общее количество таких проверок определяется числом дополнительных связей, которые не были учтены при заполнении очереди путем обхода фрагмента в ширину.

* * *

Приведенные алгоритмы поиска фрагмента в молекулярном графе используются в автоматическом ФКСП кодировщике для системы порождения

гипотез с помощью ДСМ-метода [1]. Эксперименты на реальных массивах данных показывают высокую скорость кодирования исследуемых химических соединений, которая составляет 10–20 соединений в секунду, при среднем объеме одной молекулы порядка 60–80 атомов.

СПИСОК ЛИТЕРАТУРЫ

1. Avidon V. V., Pomerantsev A. B. Structure-activity relationship oriented languages for chemical structure representation // J. Chem. Inf. Comput. Sci.— 1982. — Vol. 22, № 4. — P. 207–214.
2. Лейбов А. Е. Автоматическое кодирование химических структур кодом ФКСП // Итоги науки и техники. Сер. Информатика. Т. 15-М: ВИНТИ, 1991. — С. 141–158.
3. Блинова В. Г., Добрынин Д. Я. Языки представления химических структур в интеллектуальных системах для конструирования лекарств // НТИ. сер. 2.— 2000. — № 6. — С. 14–21.
4. Ulman J. R. An algorithm for subgraph isomorphism // J. of the ACM.— 1976. — № 23. — P. 31–42.
5. McKay B. D. Practical graph isomorphism // Congressus Numerantium.— 1981. — № 30. — P. 45–87.
6. Fortin Scott. The graph isomorphism problem // TR 96-20, Dept. of Computer Science.— University of Alberta, 1996.
7. Нечипуренко М. И., Попков В. К., Майнагашев С. М. и др. Алгоритмы и программы решения задач на графах и сетях.— Новосибирск: Наука (СО), 1990.
8. Dalby A., Nourse J. D., Houshell W. D. Description of several chemical structure file formats used by computer programs developed at Molecular Design Limited // J. Chem. Inf. Comput. Sci.— 1992. — № 32. — P. 244–255.

Материал поступил в редакцию 30.10.01.

УДК 004.8:54

М. В. Максин, Н. В. Харчевникова

Квантово-механический модуль системы, реализующей комбинаторно-численный подход к проблеме прогнозирования свойств химических соединений*

Рассматриваются предпосылки к воссозданию системы, реализующей комбинаторно-численный подход к решению проблемы прогнозирования свойств химических соединений. На примере модуля для вычисления энергии образования канцерогенных метаболитов полициклических ароматических углеводородов методами квантовой химии обсуждаются элементы архитектуры системы и требования, предъявляемые к компоненту системы. Детально описывается реализация этого модуля.

ОПИСАНИЕ ПРОБЛЕМЫ

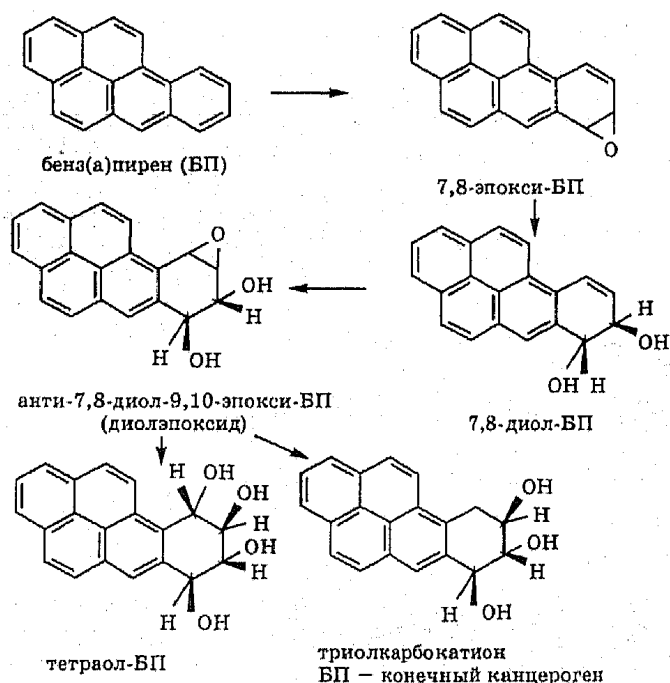
Прогнозирование канцерогенности полициклических ароматических углеводородов (ПАУ) явля-

ется весьма актуальной проблемой, так как большинство продуктов сгорания и, соответственно, загрязнителей атмосферы представляют собой именно ПАУ. Экспертные системы (ЭС) играют не-

* Работа выполнена при поддержке Российского фонда фундаментальных исследований (РФФИ), Грант № 02-01-00537.

маловажную роль в подобного рода исследованиях, так как позволяют заменить лабораторные эксперименты экспериментами компьютерными, т. е. экспериментами с моделью изучаемого явления.

На основе результатов многочисленных экспериментальных и теоретических исследований была разработана модель канцерогенности ПАУ, согласно которой активность углеводородов обусловлена взаимодействием с ДНК конечных метаболитов-карбокатионов, образующихся в результате многостадийного процесса биоактивации ПАУ. Основными этапами этого процесса является образование из исходного соединения эпоксида и диолэпоксида. Затем диолэпоксид атакует критические позиции ДНК по механизму нуклеофильного замещения с образованием в качестве интермедиатов тетраолов и триолкарбокатионов (рисунок). Полагается, что биологическое действие ПАУ — канцерогенное или мутагенное — будет тем сильнее, чем меньше требуется энергии для образования карбокатиона из диолэпоксида. Наиболее опасным является карбокатион с наименьшей энергией образования, т. е. наиболее стабильный относительно диолэпоксида. На практике энергию образования карбокатиона удобно вычислять относительно тетраола, так как при использовании метода Хюккеля энергия тетраола совпадает с энергией диолэпоксида. Таким образом, в рамках этой модели канцерогенная активность в ряду ПАУ зависит от энергии образования наиболее стабильного карбокатиона $\Delta E_{\text{катион}}$ относительно тетраола. Для получения параметров этой зависимости и последующего прогнозирования проводится регрессионный анализ, где $\Delta E_{\text{катион}}$ выступает в качестве независимой переменной, а уровень канцерогенности — в качестве зависимой.



Для проведения квантово-механических расчётов в данном случае может применяться простой метод Хюккеля [1], который значительно снижает размерность задачи и время ее счёта, но обеспечивает, тем не менее, необходимую точность. В методе Хюккеля решение уравнения Шредингера сводится к вычислению собственных чисел и

векторов симметричной матрицы, для чего, в свою очередь, можно использовать метод вращений Якоби [2]. Таким образом, данная модель предлагает довольно эффективный с вычислительной точки зрения метод прогнозирования.

Данный подход был реализован С. И. Маневичем в системе AromWin [3]. Однако в процессе её использования были обнаружены соединения, “выпадающие” из указанной зависимости, т. е. имеющие меньшую активность при большем значении $\Delta E_{\text{катион}}$. Анализ показал, что данной числовой характеристики недостаточно для объяснения проявляемых свойств и необходимо иметь в виду также и структурные свойства соединений. Так возникла идея о том, что для прогнозирования биологической активности надо одновременно использовать логико-комбинаторный подход, а именно ДСМ-метод [4, 5], и числовые параметры — физические характеристики. ДСМ-метод позволяет учитывать всю сложную структуру молекулы и, тем самым, возможность различных процессов биотрансформации и взаимодействия с организмом, в частности процесса детоксикации, а числовые параметры характеризуют отдельные процессы метаболической биоактивации и взаимодействия.

Идея такого комбинированного подхода к проблеме прогнозирования канцерогенности ПАУ воплотилась в другой программной системе С. И. Маневича — Hydcarb [6]. В ней в качестве ДСМ-объекта рассматривается объединение структурной компоненты — усовершенствованного ФКСП-кода, и числовой компоненты — энергии образования наиболее стабильного карбокатиона $\Delta E_{\text{катион}}$. Операция “сходство” таких объектов определена как сходство по обоим компонентам, а числовых величин — как принадлежность этих величин одному из кластеров, выделяемых на основе обучающей выборки. Система оправдала возлагавшиеся на неё надежды — в экспериментах с массивом данных, содержащим “отклонения” от регрессионной модели, все примеры были объяснены (по критерию достаточного основания), а результаты прогнозирования признаны адекватными.

После смерти С. И. Маневича исследования в данном направлении приостановились. Система нуждалась в развитии (одно из предполагаемых направлений — использование различных числовых параметров, — химических дескрипторов [7], а также проведение расчётов для более широкого класса соединений) и модернизации (существует лишь 16-битная версия системы), а исходный код был утерян. Встал вопрос о создании системы заново.

Тем временем продолжалось развитие ДСМ-метода, как в его теоретической, так и в прикладной частях, в том числе в области создания экспертных систем для задачи “структура—активность”. Д. А. Добрыниным была воспроизведена и развита версия ФКСП-кодировщика, позволяющая при кодировании циклических структур (в том числе полиароматических углеводородов) учитывать не только максимальную структуру, но и все немаксимальные подструктуры [8]; до этого каждая структура кодировалась своим собственным кодом, и сходство любых двух неидентичных структур было пустым. А. В. Путриным была реализована новая версия ДСМ-ядра, и

на ее базе построена программная система JSM-Explorer/Chemical edition [9], предназначенная для интерактивного проведения экспериментов по прогнозированию свойств химических соединений с помощью ДСМ-метода. Отличительной особенностью этой системы является распределённая архитектура, при которой ключевые подсистемы выделены в отдельные модули с DCOM-интерфейсом. Преимущество такого подхода — это расширяемость системы, с одной стороны, и возможность многократного использования модулей — с другой.

В свете этих достижений воссоздание системы, сочетающей ДСМ-подход и числовые методы, выглядит ещё более перспективным. Во-первых, такая система являлась бы не отдельной программой, а частью программного комплекса, используя и расширяя его возможности и совершенствуясь вместе с ним. Во-вторых, такая система потребовала бы меньших усилий при создании, так как для выполнения общих функций, таких, например, как интерактивный ввод данных и визуализация химических соединений, могла бы использоваться уже существующие модули. Таким образом, для реализации комбинированного подхода достаточно, по большому счёту, реализовать модуль вычисления значений числового параметра и модуль вычисления сходства значений числового параметра. Последний, кстати, мог бы быть общим для различных параметров, т. е. для подключения к анализу дополнительных параметров может оказаться достаточным создать только процедуру вычисления значения, а для вычисления сходства использовать существующий модуль.

ПОСТАНОВКА ЗАДАЧИ

Итак, необходимо создать программный модуль, который для данного полиароматического углеводорода строит его наиболее стабильный карбокатион и вычисляет энергию его образования $\Delta E_{\text{катион}}$ по следующему алгоритму:

1. Найти все внешние циклы (т. е. циклы, имеющие с молекулой одно и только одно общее ребро).

1.1. Для каждого цикла получить соответствующий ему тетраол. Для получения тетраола $M_{\text{тетраол}}^i$ из исходного соединения ароматическая связь в цикле разрывается и к каждому внешнему атому углерода присоединяется группа OH; молекула при этом перестаёт быть “плоской” и положения групп H и OH относительно плоскости цикла (“сверху” — “снизу”) чередуются для соседних атомов C. Вычислить энергию тетраола $E_{\text{тетраол}}^i$ (процедура вычисления энергии ПАУ будет описана ниже).

1.2. Из тетраола строим два производных карбокатиона $M_{\text{катион}}^{ij}$, $j=1,2$, “восстанавливая” ароматическую связь по очереди у каждого из атомов C, принадлежащих общему ребру. Вычисляем энергии карбокатионов $E_{\text{катион}}^{ij}$, $j=1,2$ и энергии их образования $\Delta E_{\text{катион}}^{ij} = E_{\text{катион}}^{ij} - E_{\text{тетраол}}^i$, $j=1,2$. В действительности $\Delta E_{\text{катион}}^{ij} < 0$, что говорит об энергетической выгоды такой трансформации.

2. Найти максимальную по модулю энергию образования: $\Delta E_{\text{катион}} = \max_{i,j} |\Delta E_{\text{катион}}^{ij}|$ и восстановить соответствующий карбокатион $M_{\text{катион}}$, который и считается наиболее стабильным.

Процедура вычисления энергии ПАУ методом Хюккеля выглядит следующим образом:

I. Находятся все атомы C, образующие систему ароматических связей, т. е. π -электронную систему.

II. Строится матрица смежности атомов π -электронной системы

$$\|a_{ij}\|, i, j = \overline{1, n},$$

где

$$a_{ij} = \begin{cases} \alpha_c \equiv 0, & i = j \\ \beta_0 \equiv -1, & |i - j| = 1 \text{ (соседние)} \\ 0, & \text{(несоседние)} \end{cases}$$

Параметры α_c и $\beta_{CC} \equiv \beta_0$ есть эмпирические параметры метода Хюккеля.

III. Находятся собственные значения ε_i , $i = \overline{1, n}$ матрицы $\|a_{ij}\|$, которые являются энергетическими уровнями данной π -электронной системы. Матрица симметричная, так что для решения может использоваться метод вращений Якоби [2].

IV. Вычисляется энергия системы как сумма энергий π -электронов. Электроны (количество электронов n) размещаются на электронных уровнях подвое, начиная с самого нижнего уровня, так

что $E_\pi = \sum_{i=1}^n \varepsilon'_{[\frac{i-1}{2}]+1}$, где ε'_i , $i = \overline{1, n}$ есть собственные значения, упорядоченные по возрастанию, а $[\]$ обозначает взятие целой части числа. С архитектурной точки зрения к модулю предъявляются следующие требования:

1. Использовать стандартные форматы данных, например, MDL-формат для представления химических соединений [10]. Преимущества у этого подхода следующие:

1. Использовать стандартные форматы данных, например, MDL-формат для представления химических соединений [10]. Преимущества у этого подхода следующие:

- такие стандарты являются результатом кропотливой работы и, как правило, представляют собой наилучшее решение проблемы, учитывающее, аспекты, которые на начальном этапе либо не видны разработчику, либо кажутся ему несущественными, но, скорее всего, проявятся со временем;
- позволяют обмениваться данными с другими приложениями, поддерживающими данный стандарт, и, подготовка например файлов с химическими соединениями может производиться в профессиональном редакторе, учитывая все его возможности;
- освобождают разработчика от необходимости документировать выбранный формат данных.

2. Предоставить пользователю возможность точной настройки (fine tuning) модуля, например, с помощью конфигурационного файла. Сюда можно включить и возможность изменения констант предметной области (в нашем случае это, например, эмпирические параметры метода Хюккеля) и параметров методов обработки данных (как, скажем, точность диагонализации матрицы). Поэтому такие параметры должны вычитываться извне, а не быть жёстко закодированными (hardcoded). В то же время необходимо, чтобы система была работоспособной и без такого внешнего источника, т. е. содержала бы в себе конфигурацию “по умолчанию” (default).

3. Возможность трассировки программы с различными уровнями детализации. На низком уровне детализации трассировка может служить протоколом (log) работы программы для пользователя, на высоком — средством отладки программы

для разработчика. Уровень детализации должен быть конфигурационным параметром, отдельным для каждой подсистемы.

Конфигурационные параметры и трассировка программы представляют собой средства дополнительного, "косвенного" ввода и вывода, обеспечивая необходимую системе гибкость для проведения экспериментов.

РЕАЛИЗАЦИЯ МОДУЛЯ

Основными структурами данных, используемых в модуле, являются:

1. Симметричная матрица (SymmetricMatrix), реализованная как частный случай типизированной матрицы (TypedMatrix) с элементами типа double. Реализованы необходимые операции матричной алгебры.

2. Молекула (Molecule), представляющая набор атомов (Atom) и связей (Bond). Атом есть химический элемент (Element), конкретизированный пространственными координатами. Внешним форматом данных является MOL-формат. Также реализованы методы, необходимые для манипуляций с молекулами.

3. Метаболит ПАУ (PAHs::Metabolite) — структура, содержащая молекулу метаболита и значение энергии его образования.

Компонентами, составляющими модуль, являются:

1. Решатель симметричных матриц (SymmetricMatrixResolver), находящий для данной симметричной матрицы её собственные вектора и числа. В настоящий момент реализован один его вариант — решатель Якоби (JacobyResolver).

2. Квантово-механический решатель (Quantum::Resolver), вычисляющий для данной молекулы уровни её энергии и общую энергию при заданном количестве электронов. В настоящий момент реализован простейший решатель — решатель Хюккеля для π -электронных систем (HuckelResolver), использующий для вычислений указанный решатель симметричных матриц. Этот подход (design pattern) известен как Strategy [11].

3. Генератор метаболитов (PAHs::MetabolGenerator), реализующий описанный выше алгоритм поиска наиболее стабильного метаболита с использованием указанного квантово-механического решателя (также подход Strategy) и приведённого ниже алгоритма нахождения крайних циклов в ПАУ.

4. Монитор генерации метаболитов (PAHs::MetabolGenerationTracker), позволяющий отслеживать этапы процесса и выполнять дополнительные действия в ключевых точках. Это решение может рассматриваться как упрощённый вариант подхода Command [11] объектно-ориентированного приемника callback-функций.

5. Конфигуратор (Config) как единый источник конфигурационной информации. Компонента получает значение конфигурационного параметра, указав своё имя и название параметра. Данные могут быть загружены из INI-файла и/или изменены программно. Пример INI-файла и поддерживаемые параметры приведены в *Приложении*. Конфигуратор реализован как Singleton [11].

6. Объект трассировки (Trace), представляющий перенаправляемый поток для вывода текстовой информации (на консоль или в файл) с регулируемым уровнем детализации ("Отключён",

"Ошибки", "Контрольные точки", "Информация", "Данные", "Максимальный"). Направление вывода и уровень детализации задаются отдельно для каждого компонента с помощью конфигурационных параметров TraceTarget и TraceLevel.

7. Для интерактивной версии модуля используется также анализатор командной строки (CommandLineFormat), позволяющий задать требуемый формат, проанализировать командную строку в соответствии с заданным форматом и вывести на экран пользовательскую информацию о формате командной строки. Элементами командной строки (CommandLineItems) могут быть:

- параметры (Parameter), т. е. значения, задаваемые пользователем произвольно, например, имя файла с данными;
- ключи (Option), позволяющие пользователю включить или выключить некоторую дополнительную функциональность, например, /v или -v для более подробного вывода на экран. Ключ может содержать свой список элементов, которые вступают в силу, если ключ выбран;
- переключатели (Switch), задающие набор элементов, из которых пользователь может выбрать только один. Элементом этого набора может быть список элементов.

Формат командной строки вызова модуля приведён в *Приложении*.

Алгоритм нахождения крайних циклов ПАУ выглядит следующим образом.

Для каждого атома углерода рекурсивно находится цепочка его "ароматических" соседей справа и слева; цепочка обрывается на атоме, который имеет число ароматических связей, отличное от двух. Полученная цепочка считается внешним циклом, если её длина равна шести и она замкнута. Проверка на замкнутость может быть отключена (конфигурационным параметром MakeSureChainIsCycle), если в исследуемых соединениях нет подобных цепочек, не являющихся циклами.

Таким образом, работа модуля состоит в следующем:

1. Создание объекта класса PAHs::MetabolGenerator. Каждый из объектов, создаваемых при работе, запрашивает конфигуратора на предмет необходимых параметров; если параметр не задан, используются значения по умолчанию.

2. Вызов процедуры подготовки генератора для работы с заданным соединением в виде объекта класса Molecule. В интерактивной версии соединение загружается из MOL-файла, указанного в командной строке. В процессе подготовки в соединении находятся внешние циклы и возвращается их число. Если ни один внешний цикл не найден, то данный подход считается неприменимым к указанному соединению и работа прекращается.

3. Вызов процедуры поиска наиболее стабильного метаболита с предоставлением генератору объекта класса Quantum::Resolver (которому, в свою очередь, необходим объект класса SymmetricMatrixResolver). Результатом работы является структура PAHs::Metabolite. В интерактивной версии на экран выводится значение энергии образования метаболита, и сам метаболит сохраняется в MOL-файл, если задан соответствующий параметр командной строки. Промежуточные результаты также могут быть обработаны, если генератору предоставить объект типа

PAHs::MetabolGenerationTracker. В интерактивной версии, если задан соответствующий параметр командной строки, используется SaveToFileTracker, сохраняющий промежуточные соединения в виде MOL-файлов и выводящий информационные сообщения на экран.

Примеры вызова модуля и результирующих сообщений для бенз(а)трапена (соединение № 17 из экспериментального массива) приведены в *Приложении*.

ЭКСПЕРИМЕНТ

Прогонка массива данных из 25 полиароматических углеводородов полностью подтвердила корректность данной реализации. Полученные метаболиты и значения энергии их образования в точности совпали с эталонными. В *Приложении* приведено дерево соединений и значения энергий, полученное с помощью программы для бенз(а)антрацена (соединение № 17 из экспериментального массива). Варьирование конфигурационных параметров позволяет контролировать точность и быстродействие (например, было установлено, что уменьшение точности диагонализации матриц в 100 раз — с $1.0e-6$ до $1.0e-4$ — вызывает увеличение производительности в среднем на 20%, уменьшая точность результата всего на 0.01%). С помощью увеличения уровня детализации трассировочной информации можно получить данные для измерения производительности и проследить путь выполнения, например, процесса поиска внешних циклов и поиска конечного метаболита.

НАПРАВЛЕНИЕ РАЗВИТИЯ

Как было определено ранее, целью данной работы является создание программной системы, объединяющей различные подходы в решении задач прогнозирования свойств химических соединений. Во-первых — это разработка архитектуры такой системы и создание прототипа. Видимо, в основу архитектуры будет положена методология “классной доски” (Blackboard), подробно описанная в [13]. Прототип будет построен на базе системы JSM-Explorer с использованием двух подходов: комбинаторного (в виде ДСМ-метода) и численного (в виде описанного метода прогнозирования канцерогенности ПАУ). Во-вторых, необходимо пополнить “инструментальный ящик”, как за счёт усовершенствования существующих методов и реализаций, так и путем создания новых, например, для других классов химических соединений.

СПИСОК ЛИТЕРАТУРЫ

1. Базилевский М. В. Метод молекулярных орбит и реакционная способность органических молекул. — М.: Химия, 1969.
2. Воеводин В. В., Кузнецов Ю. А. Матрицы и вычисления. — М.: Наука, 1984. — 312 с.
3. Дьячков П. Н., Маневич С. И. Автоматизированная система прогнозирования канцерогенности полициклических углеводородов и их производных методом Хюккеля // НТИ. Сер. 2. — 1996. — № 7.
4. Кузнецов С. О. ДСМ-метод как система автоматического обучения // Итоги науки и техники. Сер. Информатика. Т. 15. — М.: ВИНТИ, 1991.
5. Финн В. К. Правдоподобные рассуждения в интеллектуальных системах типа ДСМ // Итоги науки

и техники. Сер. Информатика. Т. 15. — М.: ВИНТИ, 1991.

6. Маневич С. И., Харчевникова Н. В., Дьячков П. Н. Прогнозирование контрпродуктивных свойств химических соединений при комбинированном использовании структурных формул и численных энергетических параметров // НТИ. Сер. 2. — 2000. — № 5.

7. Маневич С. И. Расчет дескрипторов соединений органической химии // НТИ. Сер. 2. — 1996. — № 5-6.

8. Блинова В. Г., Добрынин Д. А. Язык ФКСР описания химической структуры соединения // НТИ. Сер. 2. — 2001. — № 6. — С. 14-21.

9. Путрин А. В. Диссертация канд. техн. наук. — М.: ВИНТИ, 2000.

10. CTfile Formats. MDL Information Systems, Inc. 1999.

11. Gamma E., Helm R., Johnson R., Vlissides J. Design Patterns CD: Elements of Reusable Object-Oriented Software. — Addison Wesley Longman. Inc., 1998.

12. Дьячков П. Н. Квантово-химические расчеты в изучении механизма действия и токсичности чужеродных соединений // Итоги науки и техники. Сер. Токсикология. Т. 16. — М.: ВИНТИ, 1990.

13. Buschman F., Meunier R., Rohnert H., Sommerlad P., Stal M. Pattern-Oriented Software Architecture, Volume 1: A System of Patterns — John Wiley & Son Ltd, 1996.

ПРИЛОЖЕНИЕ

1. Пример INI-файла и описание поддерживаемых конфигурационных параметров.

[Trace]

Секция Trace содержит значения по умолчанию параметров трассировки работы компонента, то есть если параметр не найден в секции, соответствующей компоненту, то будет использовано значение параметра из этой секции.

Параметр TraceTarget задаёт направление вывода информации. Возможные значения:

&cout — Стандартная консоль вывода (standard output); &cerr — Стандартная консоль вывода ошибок (standard error output); &clog — Псевдоним для cerr; <file> — Любое другое значение рассматривается как имя файла. TraceTarget = PAHsMetabol.trc.

Параметр TraceLevel определяет уровень детализации выводимой информации. Возможные значения: 0 — Трассировка отключена; 1 — Выводить только ошибки; 2 — Сообщения уровня 1 плюс прохождение контрольных точек; 3 — Сообщения уровня 2 плюс дополнительная информация; 4 — Сообщения уровня 3 плюс данные (объём может быть огромен); 10 — Вся возможная информация. TraceLevel = 0.

[PAHsMetabolGenerator]

Если параметр MakeSureChainIsCycle равен true, то в процессе поиска внешних циклов найденные цепочки углеродов проверяются на замкнутость (то есть, что они действительно являются циклами). Если в анализируемом соединении нет углеродных цепочек длиной равной шести, не являющихся циклами, то значение может быть false — это немного увеличит производительность. MakeSureChainIsCycle = false.

[HuckelResolver]

α определяет значение кулоновского параметра для углерода $\alpha_C = 0.0$; β определяет значение резонансного параметра $\beta = -1.0$. [JacobyResolver]

Program Precision определяет точность диагонализации матрицы. Точнее, он задаёт пороговое значение отношения суммы квадратов недиагональных элементов к сумме квадратов диагональных элементов; если это отношение меньше указанного порогового значения, то матрица считается диагональной.

Precision = $1e-5$.

2. Формат командной строки интерактивной версии модуля
PAHsMetabol.exe filename [/s [/a] prefix] [/v] [/i [inifile]].

Параметры: filename — имя MOL-файла исследуемой молекулы ПАУ; prefix задаёт префикс, который будет использоваться при формировании имён выходных файлов; /s сохраняет молекулу конечного метаболита в MOL-файле; имя файла образуется как префикс_met.mol; /a сохраняет также промежуточные соединения в MOL-файлах: тетраол — в файле префикс_номер цикла_tet.mol, карбокатион 1 — в файле префикс_номер цикла_cc1.mol, карбокатион 2 — в файле префикс_номер цикла_cc2.mol; /v включает режим вывода подробной информации (в обычном режиме на экран выводится только энергия образования конечного метаболита или сообщение об ошибке); /i загружает конфигурационную информацию из указанного INI-файла; inifile имя INI-файла, значение по умолчанию — PAHsMetabol.ini.

3. Примеры вызова модуля и результирующих сообщений для бенз(а)антрацена (соединение № 17 из экспериментального массива).

Основной результат — энергия образования наиболее стабильного карбокатиона:

M:\FinnLab\>PAHsMetabol.exe Molecule.lib./17.mol $\Delta E = 0.848049$.

Дополнительный вывод информации о промежуточных соединениях (ключ/v) и сохранение структуры конечного метаболита в MOL-файле (ключ/s):

M:\FinnLab\>PAHsMetabol.exe Molecule.lib./17.mol /s 17 /v

Cycle 1, Tetraol: Energy is -19.4482

Cycle 1, Cation 1: Energy is -20.1847

Energy of formation is 0.754902

Cycle 2, Tetraol: Energy is -19.3137

Cycle 2, Cation 1: Energy is -20.1618

Energy of formation is 0.848049

Cycle 2, Cation 2: Energy is -20.0834

Energy of formation is 0.769689

Most stable metabolite: Saved in file 17-met.mol

Energy of formation is 0.848049.

Сохранение промежуточных соединений в MOL-файлах (ключ/a) и загрузка нестандартной конфигурации модуля (ключ /i):

M:\FinnLab\>PAHsMetabol.exe Molecule.lib./17.mol /s /a 17 /i /v

Configuration is loaded from file PAHsMetabol.ini

Cycle 1, Tetraol: Saved in file 17-1-tet.mol

Energy is -19.4482

Cycle 1, Cation 1: Saved in file 17-1-cc1.mol

Energy is -20.1847

Energy of formation is 0.736447

Cycle 1, Cation 2: Saved in file 17-1-cc2.mol

Energy is -20.2031

Energy of formation is 0.754902

Cycle 2, Tetraol: Saved in file 17-2-tet.mol

Energy is -19.3137

Cycle 2, Cation 1: Saved in file 17-2-cc1.mol

Energy is -20.1618

Energy of formation is 0.848049

Cycle 2, Cation 2: Saved in file 17-2-cc2.mol

Energy is -20.0834

Energy of formation is 0.769689

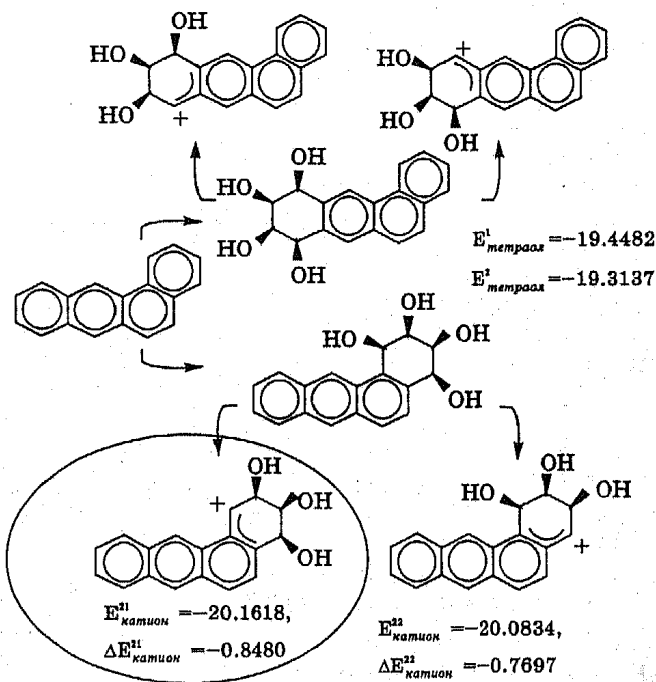
Most stable metabolite: Saved in file 17-met.mol

Energy of formation is 0.848049.

4. Полученное дерево метаболизма для бенз(а)антрацена (наиболее стабильный, т. е. конечный метаболит обведён кружком):

$E_{\text{катион}}^{11} = -20.1847$, $\Delta E_{\text{катион}}^{11} = -0.7364$

$E_{\text{катион}}^{12} = -20.2031$, $\Delta E_{\text{катион}}^{12} = -0.7549$



Материал поступил в редакцию 22.01.02.