

(Москва)

АЛГОРИТМ ПЕРВИЧНОЙ ИДЕНТИФИКАЦИИ
ЛИНИЙ МАСС-СПЕКТРА

Введение. Масс-спектральный анализ органических веществ на приборе с высокой разрешающей способностью при регистрации масс-спектра на фотопластинке требует идентификации нескольких сотен линий, соответствующих определенным комбинациям ионов. Процедура идентификации заключается в измерении расстояния между масс-спектральными линиями, вычислении массы иона, соответствующего данной линии, с точностью, определяемой используемыми средствами измерения. Как правило, точность измерения расстояния находится в пределах ± 5 мкм, и соответственно точность нахождения массы иона составляет $(1-2) \cdot 10^{-3}$ а. е. м.

Идентификация зарегистрированного иона проводится последовательным перебором возможных комбинаций компонентов анализируемой смеси либо с помощью специальных атласов [1]. Процедура идентификации весьма утомительна и трудоемка. Поэтому естественным является желание автоматизировать процесс поиска возможных комбинаций ионов. Известные алгоритмы [2-4] первичной идентификации, т. е. определения элементного состава соединения в соответствии с атомным весом, приписанным масс-спектральной линии, обладают тем недостатком, что они рассчитаны на ограниченное и фиксированное число различных изотопов, участвующих в процессе поиска элементной формулы. В данном сообщении приводится алгоритм первичной идентификации масс-спектральных линий при практически неограниченном числе исходных компонентов.

Описание алгоритма. Пусть какой-либо линии масс-спектра приписан молекулярный вес T , определенный с точностью ε . Требуется найти все наборы N_k , для которых

$$\left| \sum_{k=1}^m N_k W_k - T \right| \leq \varepsilon. \quad (1)$$

Здесь N_k — число изотопов веса W_k , входящих в химическое соединение, которое могло дать идентифицируемую линию. На N_k накладывается ограничение $0 \leq N_k \leq N_{k\max}$. Очевидно, знание всех N_k эквивалентно знанию элементного состава соединения. Суть представленного алгоритма сводится к последовательному перебору всех возможных значений N_k с отбрасыванием заведомо неприемлемых по молекулярному весу вариантов идентификации. Перебор ведется по принципу заполнения ячеек многорядного счетчика: сначала условие (1) проверяется при $N_1 = N_2 = \dots = N_{m-1} = 0$ и $N_m \neq 0$, затем принимается $N_{m-1} = 1, 2, \dots$, далее вводятся отличные от нуля значения N_{m-2} и т. д. Такой подход позволяет проводить поиск вариантов при произвольном m , что эквивалентно организации программы с переменным числом вложенных циклов.

Порядок работы программы. На основе описанного алгоритма составлена программа SPARROW. Программа существует в двух вариантах: на языках МИР-1 (см. листинг в конце сообщения) и ФОРТРАН-4. Рассмотрим второй, более совершенный, вариант. Пример работы программы первичной идентификации приведен в таблице.

Программа запрашивает число составляющих, т. е. число различных изотопов, участвующих в идентификации. Затем для каждой составляющей необходимо указать соответствующие порядковый номер элемента (в периодической таблице), порядковый номер изотопа этого элемента (в сторону возрастания атомных весов) и максимальное количество таких изотопов, участвующих в поиске. Далее вводятся молекулярный вес, приписанный идентифицируемой линии, и абсолютная ошибка его определения ε .

Обсуждение результатов. Главным достоинством программы SPARROW является ее универсальность, т. е. применимость к идентификации любых типов соединений; в этом отношении программа превосходит всевозможные специализированные атласы. Путем введения незначительных дополнений в программу можно предусмотреть расположение вариантов в порядке возрастания молекулярного веса, переход от молекулярных весов к расстояниям, нахождение дисперсии между вариантами, определение необходимой разрешающей способности и т. д. Дальнейшее усовершенствование предполагает привлечение информации об изотопных содержаниях с целью предсказания интенсивности пиков, соответствующих найденным вариантам (аналогично [4, 5]), а также отбрасывание тех вариантов, которые являются неприемлемыми по соображениям химии.

Программа SPARROW предназначена быть составной частью более широкой программы автоматической идентификации сложных масс-спектров; она может быть также использована при установлении необходимой разрешающей способности для масс-спектрального анализа тех или иных классов соединений [2].

: PR, SPRW

PROGRAMME SPARROW-76
CONSTITUENTS NUMBER=7
INPUT ELEMENT & ISOTOPE ORDINAL NUMBERS
FOR EACH CONSTITUENT AND ALSO ITS MAXIMUM QUANTITY
8,3,2
8,2,2
8,1,2
6,2,4
6,1,4
1,2,12
1,1,12
ATOMIC WEIGHT OF THE LINE TO BE IDENTIFIED=46.042
EPSILON=.01

	<u>O</u>	<u>C</u>	<u>H</u>
.460418E+02	0 0 1	0 2	0 6
.460403E+02	0 0 1	0 2	1 4
.460387E+02	0 0 1	0 2	2 2
.460372E+02	0 0 1	0 2	3 0
.460374E+02	0 0 1	1 1	0 5
.460359E+02	0 0 1	1 1	1 3
.460343E+02	0 0 1	1 1	2 1
.460330E+02	0 0 1	2 0	0 4
.460382E+02	0 1 0	0 2	0 5
.460367E+02	0 1 0	0 2	1 3
.460351E+02	0 1 0	0 2	2 1
.460338E+02	0 1 0	1 1	0 4
.460322E+02	0 1 0	1 1	1 2

THE SEARCH IS OVER

PROGRAM RUN 0 HS. 2 MIN. 23SEC.

П Р И Л О Ж Е Н И Е

""9. "ВЫВ"[PROGRAMME], "ПРОБ"1, [SPARROW-76], "СТРО"4; TMAX= $\Sigma(K=1, M, \text{MAXI}[K] \times W[K])$; MAXI[M]=MAXI[M]+1; "ДЛЯ"K=1"Ш"1"ДО"М"ВЫП"N[K]=0; L1.A=0; VAL=TMAX; "ДЛЯ"K=1"Ш"1"ДО"М"ВЫП"("ЕСЛ"K=1"ТО"("ДЛЯ"J=1"Ш"1"ДО"М"ВЫП"ЕСЛ"N[J]<MAXI[J]"ТО"("НА"Л2); "НА"END); L2.VAL=VAL-MAXI[K] \times W[K]; "ЕСЛ"K=M"ТО"(VAL=VAL+W[K]); "ДЛЯ"J=K+1"Ш"1"ДО"М"ВЫП"ЕСЛ"N[J]<MAXI[J]"ТО"("НА"Л3); "ДЛЯ"J=K+1"Ш"1"ДО"М"ВЫП"N[J]=0; "ЕСЛ"K<M"ТО"(N[K]=N[K]+1); L3.A=A+N[K] \times W[K]; L4."ЕСЛ"A>T+EPS"ТО"("ДЛЯ"J=K+1"Ш"1"ДО"М"ВЫП"N[J]=MAXI[J]; "НА"Л4); "ЕСЛ"A<T-EPS-V AL"ТО"("ЕСЛ"N[K]<MAXI[K]"ТО"(N[K]=N[K]+1; A=A+W[K]; "НА"Л4)"ИН"("ДЛЯ"J=K+1"Ш"1"ДО"М"ВЫП"N[J]=MAXI[J]; "НА"Л4)); "ЕСЛ"K=M"ТО"("ДЛЯ"J=1"Ш"1"ДО"М"ВЫП"ЕСЛ"N[J]<10"ТО"("ВЫВ""ЗНАЧ""ПРОБ"3, N[J])"ИН"("ВЫ В""ЗНАЧ""ПРОБ"2, N[J]); "ВЫВ""ЗНАЧ""ПРОБ"7, A, "СТРО"; N[K]=MAXI[K]); "НА"Л1; END. "ВЫВ""СТРО", [THE], "ПРОБ"1, [SEARCH], "ПРОБ"1, [IS], "ПРОБ"1, [OVER]"ГДЕ"

Примечание. M — число составляющих; MAXI[M] — массив $N_{k \max}$; W[M] — массив W_k ; N[M] — массив N_k ; A — молекулярный вес, соответствующий набору N_k ; T — молекулярный вес, приписанный линии масс-спектра; EPS — ошибка в определении T.

Л И Т Е Р А Т У Р А

1. Beynon J. H., Williams A. E. Mass and Abundance Tables for Use in Mass Spectrometry. — Amsterdam — L. — N.-Y.: Elsevier Publ. Co., 1963.
2. Automatic Data Analyzer, N 36006-4: Instructions JMA-IC-O. JEOL, Tokyo, Japan.
3. Тальрозе В. Л., Разникова М. О., Разников В. В., Фридляндский Г. В. Распределение дублетов органических соединений по значениям $\Delta M/M$ и требования к разрешающей способности химических масс-спектрометров. — В кн.: Тез. докл. II Всесоюз. конф. по масс-спектрометрии, Ленинград, 22-24 октября 1974. Л.: Наука, 1974.

4. Горфинкель М. И., Нехорошев С. А. Алгоритм и программа (на входном языке ЭВМ «Наири») для подбора возможных эмпирических формул по заданному точному значению молекулярного веса соединения.— *Автометрия*, 1972, № 3.
 5. Lee J. D. Interpretation of Mass Spectra.— *Talanta*, 1973, vol. 20, p. 1029.

Поступило в редакцию 10 августа 1979 г.;
 окончательный вариант — 25 февраля 1980 г.

УДК 578.5

Ю. П. ЧЕРНОВ

(Новосибирск)

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СИГНАЛОВ ЭРИТРОЦИТОВ В КОНДУКТОСИСТЕМАХ

Необходимость исследований количественных характеристик сигналов кондуктометрических датчиков (СКД) связана с проблемой использования КД для получения и обработки распределений эритроцитов человека по размерам. Сильная деформированность эритроцитов приводит к сложной зависимости СКД от размеров, ориентации и степени деформации ξ эритроцитов.

Для модели непроводящего эритроцита — сильно деформированного сплюснутого сфероида (т. е. для $\xi = c/a \ll 1$, где полуоси сфероида $a = b > c$ жестко связаны с осями координат x, y, z) — в приближении внешнего однородного поля [4] собственные значения β^i ($i = x, y, z$) постоянного тензора β_{ij} в СКД

$$\delta\varphi(r) = \sum_{ij} \beta_{ij} r_i E_j r^{-3} \quad (1)$$

могут быть разложены в ряд по параметру ξ :

$$\beta^x = \beta^y = -V(4\pi)^{-1} [1 + (\pi/4)\xi - (1 - \pi^2/16)\xi^2 + \dots], \quad (2)$$

$$\beta^z = -(2/3) S_a^{3/2} \pi^{-5/2} [1 + 4\pi^{-1}\xi + (16\pi^{-2} - 3/2)\xi^2 + \dots], \quad (3)$$

где $S_a = \pi a^2$ — площадь поперечного сечения сфероида. При направлении внешнего поля вдоль полуоси с СКД становится пропорциональным не объему V сфероида, а степени S_a в главном члене разложения (3), но и в этом случае также существует достаточно сложная зависимость СКД от ξ .

Метод численного расчета полей при потенциальном обтекании тел произвольных конфигураций в электролитических средах [2] позволяет моделировать СКД для реальных эритроцитов. Результаты приводимых ниже расчетов были получены для модели КД [2] с радиусом цилиндрического канала $R = 50$ и длиной $L = 100$. СКД регистрировался по разности потенциалов δU в двух точках, расположенных на оси канала и симметрично удаленных относительно его центра на расстояние $r = 200$. Размеры приведены в относительных единицах. Исследовалась модель непроводящего реального эритроцита, представленного в виде тора радиусами R_G и R_S с перепонкой толщиной H .

Как и в случае сфероидальных клеток [2], СКД для эритроцита, проходящего вдоль оси симметрии канала, имеет плавную колоколообразную форму, амплитудное значение достигается в центре канала, точка перегиба почти совпадает с границей канала. Формы СКД приведены на рис. 1 для модели эритроцита (a) размерами $R_G = 5$, $R_S = 2.5$, $H = 4$ и углом ориентации $\alpha = 0^\circ$, b и v — для $R_G = 2.5$, $R_S = 1.25$, $H = 1$, $\alpha = 0^\circ$ и $\alpha = 90^\circ$ соответственно. (В случае $\alpha = 0^\circ$ оси симметрии канала КД и эритроцита совпадают). Амплитуда и фронты СКД зависят от размеров эритроцита и его ориентации в канале (пример определения размеров частиц по фронтам СКД можно найти в [3]).

СКД для эритроцитов подобных форм (т. е. сохраняющих фиксированными некоторые геометрические пропорции), как и СКД для сфероидальных клеток, с большой точностью пропорционален объему. Результаты расчетов представлены прямоми на рис. 2. Прямые отражают линейную зависимость амплитуд СКД (АСКД) от объема V для подобных эритроцитов: k и m — с углом $\alpha = 0^\circ$, l и n — с $\alpha = 90^\circ$. Прямая S соответствует АСКД для сфер. Исследовались модели эритроцитов, для которых фиксировались отношения $R_G/R_S = 2$, $H/R_G = 0.6$ для прямых k и l , а также 2.6 и 6/13 для прямых m и n . Эти пропорции отвечают реальным отношениям диаметра к толщине для нативных эритроцитов, которые равны соответственно 3 и 3.6. Факторы форм (ФФ) для первой группы подобных эритроцитов принимают значения $f_1 = f_2 = 1.32$, $f_3 = 2.73$, а для второй группы — 1.31 и 3.17 соответственно. По результатам, полученным экспериментально в работе [4], нижняя оценка границы ФФ эритроцитов составляет 1.27 для $f_1 = f_2$ и 2.38 для f_3 .

Отметим эффект экранирования тором эритроцита его внутренней перепонки и возникающие отсюда трудности принципиального характера при экспериментальном