

Письмо в редакцию**Возможность достоверной идентификации неизвестных соединений при использовании программы MS Search и коммерческой базы данных масс-спектров электронной ионизации**

**Андрей Сергеевич Самохин, Александр Игоревич Ревельский,
Дмитрий Александрович Чепелянский, Игорь Александрович Ревельский**

*Московский государственный университет им. М.В. Ломоносова,
химический факультет,
Россия, 119991, Москва, Ленинские горы, д. 1, стр. 3,
E-mail: andrey.s.samokhin@gmail.com*

Поступила в редакцию 11.10.2010 г.; после переработки — 30.10.2010 г.

Ключевые слова: *достоверная идентификация, сравнение масс-спектров, поиск по библиотеке масс-спектров, AMDIS, MS Search, времяпролетный масс-спектрометр.*

Keywords: *reliable identification, mass spectra comparison, library search, AMDIS, MS Search, time-of-flight mass spectrometer.*

Масс-спектры электронной ионизации одного и того же соединения, зарегистрированные на различных приборах при энергии ионизирующих электронов, равной 70 эВ, весьма похожи. Сравнение масс-спектров позволяет подтвердить или установить строение неизвестного соединения. В настоящее время вместо ручного сравнения (например, при помощи атласов масс-спектров [1]) используют автоматический поиск по электронным базам данных масс-спектров [2, 3], который заключается в расчете величины совпадения спектров и ранжировании веществ-кандидатов в порядке уменьшения данной величины. В программе MS Search 2.0 (NIST) помимо величины совпадения спектров (Match) предусмотрен расчет вероятности (Probability) и «уникальности» экспериментального масс-спектра (InLib) [4]. Однако на сегодняшний день не предложены граничные значения указанных параметров, позволяющие говорить о достоверной идентификации или об отсутствии масс-спектра исследуемого соединения в библиотеке. Таким образом, вещество считается идентифицированным с помощью библиотеки, если значение величины Match для первого кандидата в списке превышает установленную исследователем границу. Пороговое значение Match, используемое различными авторами, может изменяться в широких пределах, однако вопрос воспроизводимости величины Match в литературе не рассматривается.

Целью настоящей работы являлось рассмотрение вопроса воспроизводимости величины совпадения (Match) экспериментального (отвечающего нанограммовым количествам аналитов) и библиотечного масс-спектров электронной ионизации, а также рассмотрение возможности проведения достоверной

идентификации неизвестных соединений с помощью программы MS Search и коммерческой базы данных масс-спектров электронной ионизации.

В качестве модельной смеси использовали раствор 10 органических веществ (1-октанол, *n*-ундекан, 2,6-диметилфенол, метиловый эфир октановой кислоты, 2,6-диметиланилин, нафталин, *n*-додекан, 1-деканол, *n*-тридекан и метиловый эфир декановой кислоты) в *n*-гексане. Концентрация компонентов составляла 1.3 нг мкл⁻¹. Масс-спектры электронной ионизации модельных соединений были получены на времяпролетном масс-спектрометре TruTOF HT (Leco), соединенном с газовым хроматографом Agilent 7890A (Agilent Technologies). Ввод пробы объемом 1 мкл осуществляли в режиме без деления — с делением потока (вентиль делителя потока был открыт через 30 с). Температура инжектора хроматографа составляла 250 °С. Разделение модельных соединений осуществляли на капиллярной колонке Restek RTX-5 (10 м × 0.18 мм × 0.2 мкм). Температуру термостата колонок изменяли по следующей программе: изотерма 40 °С в течение 3 мин; нагрев со скоростью 20 °С мин⁻¹ до 220 °С; изотерма 220 °С в течение 3 мин. Используемая температурная программа обеспечивала разделение компонентов смеси до базовой линии. В качестве газа носителя использовали гелий марки 60 (1 мл мин⁻¹). Температура источника ионов составляла 250 °С. Энергия ионизирующих электронов 70 эВ. Диапазон сканируемых значений *m/z* составлял 50–200 Да при скорости сканирования 5 скан с⁻¹.

Модельная смесь была проанализирована 5 раз. С помощью программы AMDIS 2.65 [5] для каждого компонента были выделены 5 масс-спектров элек-

Таблица 1. Средние значения (A_v), стандартные отклонения (StD), минимальные (Min) и максимальные (Max) значения величины $Match$, рассчитанной при сравнении экспериментального и библиотечного масс-спектров. Поиск масс-спектров проводили по библиотеке MainLib (NIST 05).

Соединение	A_v	StD	Min	Max
1-Октанол	952	5	946	959
<i>n</i> -Ундекан	925	3	919	927
2,6-Диметилфенол	944	2	942	946
Метилвый эфир октановой кислоты	875	9	864	883
2,6-Диметиланилин	941	2	939	943
Нафталин	965	3	961	969
<i>n</i> -Додекан	880	1	879	881
1-Деканол	922	4	915	929
<i>n</i> -Тридекан	921	5	916	927
Метилвый эфир декановой кислоты	818	1	817	818

тронной ионизации низкого разрешения (следует отметить, что при проведении идентификации большинство исследователей ограничивается 1–2 масс-спектрами). Для количественной оценки совпадения масс-спектров использовали величину $Match$, рассчитанную с помощью программы MS Search 2.0 (алгоритм «Identity») [4]. В работе использовали библиотеку MainLib (NIST 05), содержащую только один масс-спектр для каждого представленного в ней соединения. Средние значения, стандартные отклонения, минимальные и максимальные значения величины $Match$ ($n = 5$), рассчитанной при сравнении экспериментального и библиотечного масс-спек-

тров, представлены в табл. 1. Стандартное отклонение не превышает 9, а разница между максимальным и минимальным значением величины $Match$ не превышает 19. Ширина доверительного интервала ($P = 95\%$) величины $Match$ (рассчитанная из предположения подчинения величины $Match$ нормальному распределению) не превышает 11. Средние значения величины $Match$ варьируются в широких пределах 818–965 (несмотря на относительно небольшое число рассматриваемых соединений), что не позволяет установить пороговое значение величины $Match$, достижение которого обеспечивало бы достоверную идентификацию.

Как видно из табл. 2, заданное соединение не во всех случаях занимало первую позицию (rank 1) в списке возможных кандидатов, что можно объяснить неидентичностью условий (тип масс-анализатора, способ ввода вещества в масс-спектрометр, температура ионного источника, количество аналита) регистрации библиотечного и экспериментального масс-спектров. Следует отметить, что информация об условиях регистрации библиотечного масс-спектра отсутствует в базе данных NIST 05. Если заданное соединение не возглавляет список возможных кандидатов, следует говорить о неправильной идентификации за исключением соединений, масс-спектры которых неотличимы друг от друга (например, масс-спектры стереоизомеров или *орто*-, *мета*-, *пара*-изомеров [6]). Два модельных соединения (*n*-додекан и метилвый эфир декановой кислоты) с использованием библиотеки масс-спектров MainLib (NIST 05) были идентифицированы неверно. Следует отметить, что в случае неправильной идентификации разница между значениями величины $Match$

Таблица 2. Разница ($\Delta Match$) между значениями $Match$ заданного соединения и соединения, занимающего вторую/первую* позицию в списке возможных кандидатов (рассматриваются 5 экспериментальных масс-спектров). Поиск масс-спектров проводили по библиотеке MainLib (NIST 05).

Заданное соединение	Соединение, занимающее вторую/первую* позицию в списке возможных кандидатов	Позиция заданного соединения в списке возможных кандидатов	$\Delta Match$				
1-Октанол	1-Циклопропилпентил	1	56	38	45	40	36
<i>n</i> -Ундекан	<i>n</i> -Тридекан	1	25	31	14	19	23
2,6-Диметилфенол	2,5-Диметилфенол	1–2	0	–1**	–6**	–2**	1
Метилвый эфир октановой кислоты	Метилвый эфир гептановой кислоты	1	73	100	97	73	99
2,6-Диметиланилин	2,5-Диметиланилин	1	9	6	8	7	10
Нафталин	Азулен	1	8	11	10	10	9
<i>n</i> -Додекан	<i>n</i> -Тридекан	3	–19**	–19**	–30**	–31**	–34**
1-Деканол	1-Циклопропилнонил	1	30	31	31	30	25
<i>n</i> -Тридекан	2,7,10-Триметилдодекан	1	42	37	44	42	45
Метилвый эфир декановой кислоты	Метилвый эфир ундекановой кислоты	4–6	–27**	–23**	–26**	–26**	–28**

* Соединение занимает вторую позицию в списке возможных кандидатов в случае правильной идентификации заданного соединения ($\Delta Match > 0$); соединение занимает первую позицию в списке возможных кандидатов в случае неправильной идентификации заданного соединения ($\Delta Match < 0$).

** Отрицательные значения $\Delta Match$ указывают на то, что заданное соединение не занимает первую позицию в списке возможных кандидатов:

- 2,6-диметилфенол занимает 2 позицию (в трех из 5 случаев) в списке возможных кандидатов;
- *n*-додекан занимает 3 позицию в списке возможных кандидатов;
- метилвый эфир декановой кислоты занимает 4–6 позицию в списке возможных кандидатов.

первого кандидата и заданного соединения составила 19–34 единицы (табл. 2). Таким образом, первая позиция в списке возможных кандидатов также не гарантирует достоверную идентификацию даже в том случае, когда разница между значениями величины Match первого и второго кандидата превышает 30 единиц.

Один из подходов, позволяющих увеличить вероятность попадания масс-спектра заданного соединения на первую позицию в списке возможных кандидатов, заключается в использовании баз данных, содержащих несколько масс-спектров одного и того же соединения, зарегистрированных в различных условиях [7, 8]. Использование в нашем исследовании наряду с библиотекой MainLib (NIST 05) библиотеки RepLib (NIST 05), содержащей дополнительные масс-спектры соединений, уже представленных в MainLib, позволило провести правильную идентификацию *n*-додекана и метилового эфира декановой кислоты (заданные соединения занимали первую позицию в списке возможных кандидатов). Однако в последних редакциях библиотек масс-спектров NIST [2] и Wiley [3] для большинства соединений (более 85 %) содержится только один масс-спектр. Задача включения в базу данных всевозможных вариаций масс-спектра даже в случае только одного соединения неосуществима. К тому же неизвестно, насколько зарегистрированный в «новых» условиях масс-спектр должен отличаться от масс-спектра (или масс-спектров), уже занесенного в библиотеку.

Использование информации о молекулярной массе соединения (при проведении библиотечного поиска) позволило провести правильную идентификацию *n*-додекана и метилового эфира декановой кислоты (заданные соединения занимали первую позицию в списке возможных кандидатов). В общем случае молекулярная масса может быть определена при использовании масс-спектрометрии с химической ионизацией. Следует отметить, что совместное использование масс-спектра электронной ионизации и информации о молекулярной массе позволяет увеличить достоверность идентификации. Однако в случае программы MS Search использование информации о молекулярной массе никак не влияет на расчет величины Match, а позволяет только сократить число возможных кандидатов.

Из полученных данных можно сделать вывод, что использование алгоритма библиотечного поиска «Identity» (реализованного в программе MS Search 2.0) в сочетании с использованием коммерческой библиотеки масс-спектров MainLib (NIST 05) не позволяет сделать однозначное заключение о достоверной идентификации. Результаты поиска по библиотеке масс-спектров (позиция в списке возможных кандидатов и/или значение величины Match) следует рассматривать как ориентировочные. Для достоверной идентификации необходимо проводить сравнение масс-спектров неизвестного и предполагаемого соединения, зарегистрированных в одинаковых условиях. Регистрация масс-спектра химической ионизации (наряду с масс-спектром электронной ионизации) позволяет более надежно установить молекулярную массу и также увеличивает достоверность идентификации. К сожалению, в общепринятой практике эти этапы зачастую отсутствуют.

Благодарности

Авторы выражают благодарность компании Leco за предоставление времяпролетного масс-спектрометра TruTOF HT (Leco) для проведения исследований.

Список литературы

1. Abrahamsson S., Stenhagen E., McLafferty F.W. Atlas of mass spectral data New York : Wiley, 1969.
2. NIST/EPA/NIH Mass spectral library (National Institute of Standards and Technology) 2008.
3. Wiley Registry of mass spectral data, 9th electronic ed. 2009.
4. NIST MS Search user guide. National Institute of Standards and Technology, Gaithersburg, 2008. 49 p.
5. Stein S.E. An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data // *J. Am. Soc. Mass Spectrom.* 1999. Vol. 10, N 8. P. 770–781.
6. McLafferty F.W., Turecek F. Interpretation of mass spectra Mill Valley : University Science Books, 1993. 371 p.
7. McLafferty F.W., Stauffer D.B. Retrieval and interpretative computer programs for mass spectrometry // *J. Chem. Inf. Comput. Sci.* 1985. Vol. 25, N 3. P. 245–252.
8. McLafferty F.W., Stauffer D.B. An improved comprehensive data base for matching unknown mass spectra // *J. Mass Spectrom. Ion Proc.* 1984. Vol. 58. P. 139–149.