

На правах рукописи



Самохин Андрей Сергеевич

**НОВЫЙ ПОДХОД К ИДЕНТИФИКАЦИИ КОМПОНЕНТОВ
СЛОЖНЫХ СМЕСЕЙ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ С
ИСПОЛЬЗОВАНИЕМ ГАЗОВОЙ ХРОМАТОГРАФИИ/МАСС-
СПЕКТРОМЕТРИИ**

02.00.02 – Аналитическая химия

Автореферат
диссертации на соискание ученой степени
кандидата химических наук

Москва – 2013

Работа выполнена на кафедре аналитической химии
Химического факультета Московского государственного университета
им. М.В. Ломоносова

Научный руководитель:

Доктор химических наук, профессор Ревельский Игорь Александрович
Химический факультет Московского государственного университета
им. М.В. Ломоносова

Официальные оппоненты:

Доктор химических наук Бродский Ефим Соломонович
Институт проблем экологии и эволюции им. А.Н. Северцова РАН

Доктор физико-математических наук Померанцев Алексей Леонидович
Институт химической физики им. Н.Н. Семенова РАН

Ведущая организация:

Санкт-Петербургский государственный университет

Защита состоится 20 марта 2013 года в 13 ч. 00 мин. в конференц-зале на заседании диссертационного совета Д 002.021.01 при Федеральном государственном бюджетном учреждении науки Институте общей и неорганической химии им. Н.С. Курнакова Российской академии наук (ИОНХ РАН) по адресу: 119991, ГСП-1, г. Москва, Ленинский проспект, д. 31.

С диссертацией можно ознакомиться в библиотеке ИОНХ РАН по адресу: г. Москва, Ленинский проспект, д. 31.

С авторефератом можно ознакомиться на сайте ИОНХ РАН: www.igic.ras.ru.

Автореферат разослан 18 февраля 2013 года.

Ученый секретарь диссертационного совета,
кандидат химических наук



Н.Б. Генералова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Идентификация¹ следовых количеств неизвестных² органических соединений, входящих в состав сложных многокомпонентных смесей, является одной из наиболее сложных и важных задач современной аналитической химии. В настоящее время для ее решения главным образом используют метод газовой хроматографии/масс-спектрометрии с электронной ионизацией (ГХ/МС(ЭИ)). Данное обстоятельство объясняется высокой чувствительностью и селективностью метода ГХ/МС, а также наличием баз данных масс-спектров ЭИ, использование которых значительно ускоряет и упрощает процедуру идентификации неизвестных органических соединений. Однако вопросу достоверности такой идентификации не уделяется должного внимания. Обычно вещество считается идентифицированным, если значение величины степени совпадения, рассчитанной между экспериментальным масс-спектром и библиотечным масс-спектром соединения, возглавляющего список возможных кандидатов, превышает установленную исследователем границу. Строгое обоснование такого подхода в литературе отсутствует, а влияние условий регистрации масс-спектров (тип масс-анализатора, способ осуществления калибровки прибора и др.) на результаты поиска по базе данных не учитывается.

Целью работы являлась разработка подхода к повышению достоверности идентификации следовых количеств компонентов сложных смесей органических соединений при использовании метода ГХ/МС(ЭИ).

Для достижения поставленной цели было необходимо:

- изучить вопрос воспроизводимости масс-спектров ЭИ, представленных в базах данных и полученных экспериментально с помощью масс-спектрометров различных типов;
- провести статистическое исследование интенсивности пика молекулярного иона в масс-спектрах ЭИ (представленных в соответствующих базах данных);
- изучить вопрос достоверности идентификации неизвестных органических соединений, основанной на проведении поиска по базам данных масс-спектров ЭИ с использованием коммерчески доступного программного обеспечения;

¹ В настоящей работе под термином «идентификация» понимается отождествление аналита с известным органическим соединением.

² В настоящей работе под термином «неизвестное соединение» понимается известное органическое соединение, возможность присутствия которого в исследуемом образце априорно не рассматривается исследователем.

- рассмотреть возможность увеличения достоверности идентификации, основанной на сопоставлении экспериментальных масс-спектров неизвестного соединения и возможного кандидата (зарегистрированных при одних и тех же условиях) с помощью метода главных компонент;
- рассмотреть возможность применения математических методов для выделения «чистых» и неискаженных масс-спектров соединений, хроматографические профили которых накладываются друг на друга.

Научная новизна. Изучен вопрос воспроизводимости масс-спектров ЭИ, представленных в базах данных и полученных экспериментально с использованием масс-спектрометров различных типов. Показано, что различия между масс-спектрами проявляются в основном не в наборе пиков фрагментных ионов, а в значениях их интенсивностей.

Впервые определено распределение органических соединений в зависимости от интенсивности пика молекулярного иона в масс-спектрах ЭИ (представленных в соответствующих базах данных). Установлено, что относительная интенсивность пика молекулярного иона в масс-спектрах ЭИ не превышает 1% для порядка четверти органических соединений.

Изучен вопрос достоверности идентификации неизвестных органических соединений, основанной на проведении поиска по базам данных масс-спектров ЭИ (с использованием компьютерной программы MS Search). Показано, что в общем случае результаты библиотечного поиска являются ориентировочными.

Показана возможность использования метода главных компонент для дифференциации (основанной на рассмотрении только масс-спектральных данных) следовых количеств органических соединений (*o*-, *m*-, *p*-диметилбензола), обладающих практически идентичными масс-спектрами ЭИ.

Предложен подход, позволяющий значительно повысить достоверность идентификации следовых количеств неизвестных органических соединений при использовании метода ГХ/МС(ЭИ). Подход основан на выборе наиболее вероятного кандидата с помощью баз данных масс-спектров ЭИ (и индексов удерживания) и последующем сопоставлении экспериментальных масс-спектров неизвестного и предполагаемого соединений (зарегистрированных при одних и тех же условиях) с помощью метода главных компонент и *t*-критерия Стьюдента.

Практическая значимость. Рассмотрена возможность использования коммерческих баз данных масс-спектров ЭИ для проведения достоверной идентификации неизвестных органических соединений. Показано, что результаты библиотечного поиска следует рассматривать как ориентировочные.

Показана возможность значительного увеличения достоверности идентификации следовых количеств аналитов, основанной на сопоставлении экспериментальных масс-спектров ЭИ неизвестного соединения и возможного кандидата (зарегистрированных при одних и тех же условиях) при совместном использовании метода главных компонент и t -критерия Стьюдента. Практическое применение предложенного подхода особенно целесообразно в тех случаях, когда необходима высокая достоверность идентификации (например, в судебной химии).

Рассмотрена возможность извлечения «чистых» и неискаженных масс-спектров ЭИ органических соединений, хроматографические профили которых накладываются друг на друга, с использованием различных подходов (метода «вычитания фона», прокрустового анализа) и ряда компьютерных программ (AMDIS, ChromaTOF). Полученные результаты имеют особое значение в случае анализа следовых количеств аналитов и показывают, что при наличии общих значений m/z в масс-спектрах коэлюируемых соединений и высоком содержании мешающего компонента только прокрустов анализ позволяет выделить неискаженные масс-спектры.

Положения, выносимые на защиту.

- Результаты исследования, посвященного межприборной воспроизводимости масс-спектров ЭИ и ее влиянию на достоверность идентификации неизвестных органических соединений.
- Результаты статистического исследования интенсивности пика молекулярного иона в масс-спектрах ЭИ (представленных в соответствующих базах данных), показывающие, что для порядка четверти органических соединений интенсивность пика молекулярного иона в масс-спектрах ЭИ не превышает 1%.
- Результаты, полученные при изучении вопроса достоверности идентификации неизвестных органических соединений, основанной на использовании баз данных масс-спектров ЭИ, и показывающие невысокую достоверность библиотечного поиска.
- Новый подход к достоверному сопоставлению масс-спектров (зарегистрированных при одних и тех же условиях), основанный на совместном использовании метода главных компонент и t -критерия Стьюдента и позволяющий значительно увеличить достоверность идентификации в случае использования метода ГХ/МС(ЭИ).
- Результаты сопоставления ряда программ и математических подходов, предназначенных для выделения «чистых» масс-спектров соединений, хроматографические пики которых разделены неполностью.

Вклад автора состоит в проведении всего объема экспериментальных исследований, обработке данных, интерпретации и обобщении полученных результатов.

Апробация работы. Результаты работы были представлены на III Всероссийской конференции «Масс-спектрометрия и ее прикладные проблемы» (Москва, 2009), на VII Международном зимнем симпозиуме по хемометрике «Modern Methods of Data Analysis» (Санкт-Петербург, 2010), на IV Всероссийской конференции «Фундаментальные вопросы масс-спектрометрии и ее аналитические применения» (Звенигород, 2010), на IV Всероссийской конференции «Масс-спектрометрия и ее прикладные проблемы» (Москва, 2011), на Международной конференции «Conferentia Chemometrica 2011» (Шюмег, Венгрия, 2011), на VIII Международном зимнем симпозиуме по хемометрике «Modern Methods of Data Analysis» (Дракино, 2012), на XIII Международной конференции «Chemometrics in Analytical Chemistry» (Будапешт, Венгрия, 2012), на Всероссийской конференции по аналитической спектроскопии (Краснодар, 2012) и на XIII Международной конференции «European Meeting on Environmental Chemistry» (Москва, 2012).

Публикации. По материалам диссертации опубликовано 7 статей (в рецензируемых журналах) и 12 тезисов докладов.

Структура и объем работы. Диссертация состоит из введения, пяти глав, заключения, выводов и списка литературы. Во введении отражена актуальность темы работы, сформулирована цель, научная новизна, практическая значимость и положения, выносимые на защиту. В первой главе рассмотрены основные подходы к идентификации следовых количеств органических веществ, описанные в литературе и основанные на использовании метода ГХ/МС. Во второй главе описано используемое оборудование, исходные вещества и техника эксперимента. В третьей главе рассмотрена возможность проведения достоверной идентификации неизвестных соединений, основанной на поиске по базам данных масс-спектров ЭИ. В четвертой главе рассмотрена возможность использования метода главных компонент для сравнения масс-спектров, зарегистрированных при одних и тех же условиях. В пятой главе проведено сопоставление подходов, предназначенных для выделения «чистых» масс-спектров коэлюируемых соединений. Работа изложена на 108 страницах, содержит 31 рисунок и 23 таблицы. Список литературы содержит 157 источников.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Оборудование, исходные вещества, техника эксперимента. При проведении анализа использовали ряд приборов:

- времяпролетный масс-спектрометр низкого разрешения «TruTOF HT» (Leco), соединенный с газовым хроматографом «Agilent 7890A» (Agilent Technologies);
- времяпролетный масс-спектрометр низкого разрешения «Pegasus HT» (Leco), соединенный с газовым хроматографом «Agilent 7890A» (Agilent Technologies);
- квадрупольный масс-спектрометр «Trace DSQ» (Thermo Electron), соединенный с газовым хроматографом «Trace GC ultra» (Thermo Electron);
- масс-спектрометр типа ионная ловушка «PolarisQ» (Thermo Electron) с внешним источником ионизации, соединенный с газовым хроматографом «Trace GC ultra» (Thermo Electron).

Разделение проводили на капиллярных хроматографических колонках с неполярной (100%-диметилполисилоксан) или слабополярной (5%-фенил, 95%-диметилполисилоксан) неподвижной фазой. Длина хроматографических колонок составляла 10–30 м, внутренний диаметр – 0,1–0,32 мм, толщина пленки неподвижной фазы – 0,1–0,25 мкм. В качестве газа-носителя использовали гелий марки А (99.95%) или гелий марки 60 (99.9999%). Скорость потока составляла 0,5–1,5 мл мин⁻¹. Температура источника ионов в большинстве случаев была равна 250 °С. Энергия ионизирующих электронов составляла 70 эВ. Скорость сканирования составляла 1–3 спектров с⁻¹ в случае квадрупольного масс-спектрометра и масс-спектрометра типа ионная ловушка и 5–20 спектров с⁻¹ в случае времяпролетных масс-спектрометров.

Для обработки (и моделирования) данных использовали следующее программное обеспечение: MS Search 2.0g (NIST), ChromaTOF 4.32 (Leco), AMDIS 2.70 (NIST), Qual Browser 2.0 (Thermo Fisher), Maple 13 (Maplesoft), Microsoft Excel 2007 (Microsoft Corporation), надстройку Chemometrics Add-In для программы Microsoft Excel (авторы надстройки – А.Л. Померанцев и О.Е. Родионова), а также ряд компьютерных программ, написанных нами на языке программирования С.

При анализе растворов модельных соединений в хроматографическую колонку вводили 0,1–1,3 нг каждого аналита. В работе использовали ряд растворов модельных соединений и образец фармацевтического препарата:

- модельная смесь 10 органических соединений (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 (тестовая смесь № 31, Chrompack);

- индивидуальные растворы *o*-, *m*-, *p*-диметилфенола и этилбензола (PolyScience Corporation) в *n*-гексане;
- индивидуальные растворы 2,4- и 2,5-диметилфенола (PolyScience Corporation) в *n*-гексане;
- модельная смесь 1-октанола, 1-нонанола, 1-деканола, 1-ундеканола, 1-додеканола, 1-тетрадеканола (PolyScience Corporation) в *n*-гексане;
- модельная смесь 64 соединений, используемая в методике EPA-8270;
- фармацевтический препарат Ксеникал (Roche).

Различие между масс-спектрами ЭИ, зарегистрированными в неидентичных условиях. Известно, что условия регистрации (тип масс-анализатора, способ ввода пробы, загрязненность источника ионов и ионной оптики и др.) могут оказывать влияние на вид масс-спектра ЭИ и, как следствие, на результаты библиотечного поиска. К сожалению, публикации, посвященные данной проблеме, в литературе практически отсутствуют. В связи с этим нами было проведено комплексное сопоставление масс-спектров ЭИ, зарегистрированных в неидентичных условиях. Были рассмотрены масс-спектры как полученные нами, так и представленные в базах данных OCAD и NIST'08.

Модельная смесь 10 органических соединений (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) была проанализирована 5 раз на трех различных хромато-масс-спектрометрах (времяпролетном, квадрупольном и ионной ловушке). В хроматографическую колонку было введено порядка 1,3 нг каждого из указанных аналитов. Очищенные масс-спектры были выделены с помощью программы AMDIS. В качестве примера на Рис. 1 представлены масс-спектры 1-октанола, 2,6-диметилфенола, *n*-додекана и метилового эфира декановой кислоты. Как видно из Рис. 1, масс-спектры, зарегистрированные с помощью квадрупольного и времяпролетного приборов, весьма похожи и могут отличаться от спектров, зарегистрированных с помощью ионной ловушки. Различие в масс-спектрах проявляется не в наборе интенсивных пиков фрагментных ионов, а в значениях их интенсивностей. Например, масс-спектры метиловых эфиров органических кислот, зарегистрированные с помощью ионной ловушки, характеризуются более высокими интенсивностями пиков ионов $[M - C_nH_{2n+1}]^+$ (по сравнению с квадрупольным и времяпролетным приборами). В случае ароматических соединений и ациклических предельных углеводородов визуальные различия между масс-спектрами, зарегистрированными на приборах различных типов, практически отсутствуют.

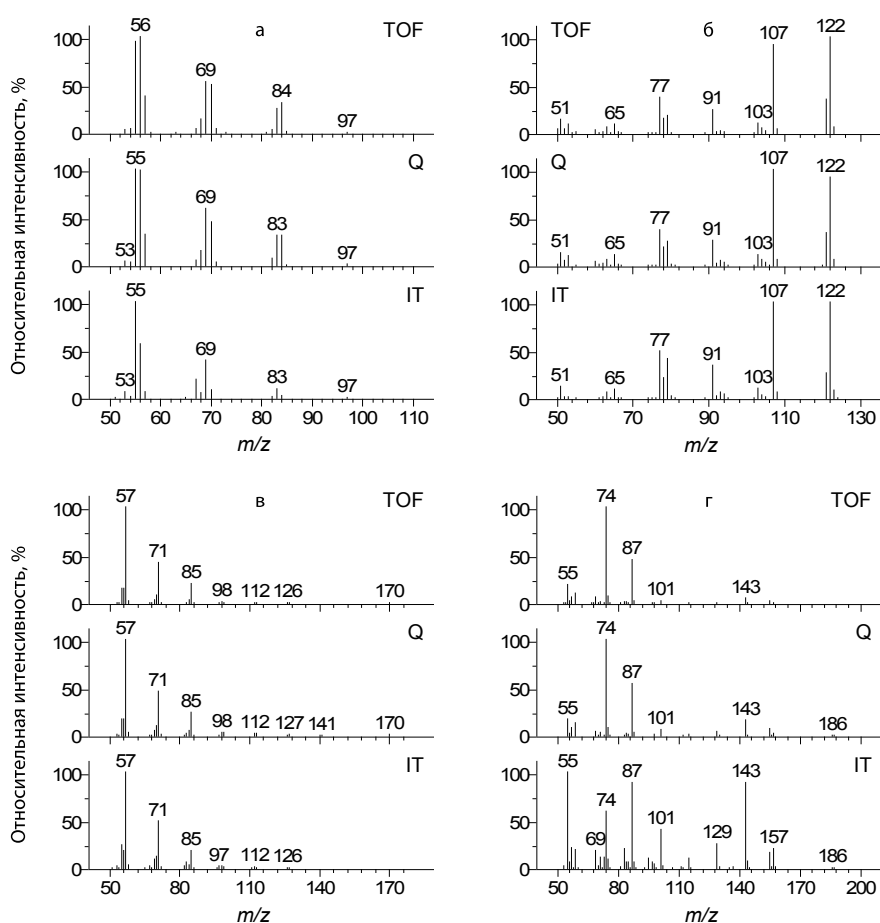


Рис. 1. Масс-спектры 1-октанола (а), 2,6-диметилфенола (б), *n*-додекана (в) и метилового эфира декановой кислоты (г), зарегистрированные с помощью времяпролетного (TOF), квадрупольного (Q) масс-спектрометров и масс-спектрометра типа ионная ловушка (IT). В хроматографическую колонку было введено порядка 1,3 нг аналитов.

Каждый экспериментальный масс-спектр был использован для проведения поиска по базе данных масс-спектров³ NIST'05 (поиск осуществляли с помощью программного обеспечения MS Search). Средние значения величины степени совпадения Match⁴ ($n = 5$) представлены в Табл. 1. В случае времяпролетного и квадрупольного приборов для каждого компонента смеси были получены приемлемые результаты (заданное соединение занимало первую или вторую позицию

³ Поиск по базе данных заключается в расчете величины степени совпадения между экспериментальным и библиотечными масс-спектрами и ранжировании веществ-кандидатов в порядке уменьшения данной величины.

⁴ Величина степени совпадения Match (рассчитанная с использованием алгоритма «Identity», реализованного в программном обеспечении MS Search) представляет собой сумму двух слагаемых. Первое слагаемое отражает степень «похожести» масс-спектров, рассчитанную по оптимизированному методу скалярного произведения (координаты вектора в N -мерном пространстве задаются как: $\sqrt{I_m} \cdot m^2$, где m – определенное значение m/z , I_m – интенсивность соответствующего масс-спектрального пика). При расчете второго слагаемого учитывается различие в отношениях интенсивностей соседних пиков в рассматриваемом масс-спектре и масс-спектре сравнения. Величина степени совпадения Match может изменяться в диапазоне от 0 (когда масс-спектры не содержат общих значений m/z) до 999 (когда масс-спектры полностью совпадают).

в списке возможных кандидатов; средние значения величины Match находились в диапазоне 916–960). В случае ионной ловушки только ароматические соединения одновременно характеризовались как высоким значением величины Match (более 900), так и первой (или второй) позицией в списке возможных кандидатов. Для ациклических предельных углеводов и первичных спиртов неудовлетворительный результат (заданное соединение не входило в список первых 10 кандидатов) мог быть объяснен невысокой «уникальностью» масс-спектров (молекулярные ионы не были зарегистрированы, а значения m/z фрагментных ионов характерны для большого числа органических соединений).

Табл. 1. Средние значения величины степени совпадения Match ($n = 5$), полученные при сравнении экспериментальных и библиотечных (NIST'05) масс-спектров. В скобках указаны значения стандартного отклонения.

Название соединения	Величина степени совпадения Match		
	Времяпролетный масс-спектрометр	Квадрупольный масс-спектрометр	Масс-спектрометр типа ионная ловушка
1-октанол	952 (6)	942 (6)	723 (16)**
<i>n</i> -ундекан	930 (3)	938 (11)	826 (8)***
2,6-диметилфенол	953 (4)*	940 (11)*	908 (6)*
метилвый эфир октановой кислоты	932 (9)	916 (10)	790 (8)
2,6-диметиланилин	960 (2)	952 (6)	921 (9)*
нафталин	954 (1)	959 (4)*	918 (5)*
<i>n</i> -додекан	935 (7)	946 (6)	815 (6)**
1-деканол	934 (9)	927 (9)	773 (16)**
<i>n</i> -тридекан	923 (6)	947 (7)	826 (9)**
метилвый эфир декановой кислоты	926 (2)	922 (9)	780 (9)

* заданное соединение в ряде случаев занимало вторую позицию в списке возможных кандидатов, что объяснялось похожестью масс-спектров веществ кандидатов: (2,6-диметилфенол, 2,3-диметилфенол, 2,5-диметилфенол), (2,6-диметиланилин, 2,3-диметиланилин, 2,5-диметиланилин), (нафталин, азулен);

** заданное соединение не входило в список первых 10 веществ-кандидатов;

*** заданное соединение не входило в список первых 5 веществ-кандидатов.

Также нами был рассмотрен ряд масс-спектров, представленных в базах данных NIST'08 (подбиблиотека *replib*⁵) и OCAD v.8⁶. Нами были выбраны соединения, для которых в базах данных было представлено наибольшее число масс-спектров. В случае базы данных OCAD были рассмотрены 128 соединений (для каждого из

⁵ В случае базы данных NIST масс-спектры представлены в двух подбиблиотеках: *mainlib* и *replib*. Подбиблиотека *mainlib* является основной и содержит только по одному масс-спектру для каждого соединения. Подбиблиотека *replib* содержит дополнительные масс-спектры соединений, уже представленных в *mainlib*.

⁶ OCAD является базой данных Организации по запрещению химического оружия.

которых было представлено как минимум 3 масс-спектра). В случае базы данных NIST'08 были рассмотрены 358 соединений (для каждого из которых в базе данных содержалось как минимум 4 масс-спектра). Общее число рассматриваемых масс-спектров составило 430 для базы данных OCAD и 1485 для базы данных NIST'08.

Библиотечные масс-спектры, отвечающие одному и тому же соединению, различались в основном интенсивностями отдельных масс-спектральных пиков (тогда как набор значений m/z практически не менялся). При этом, чем дальше два масс-спектральных пика отстояли друг от друга, тем сильнее могло меняться их отношение в масс-спектрах, зарегистрированных в неидентичных условиях. Например, для соседних масс-спектральных пиков ($\Delta m/z = 1$) вероятность того, что отношение $\frac{I_{m/z}^{(1)}}{I_{m/z+\Delta m/z}^{(1)}}$ превысит 1,5, составляла всего 6%; однако, если расстояние между рассматриваемыми пиками превышало 100 единиц шкалы m/z , то эта вероятность возрастала до 39%.

Интенсивность пика молекулярного иона в масс-спектрах ЭИ. Молекулярный ион является наиболее значимым ионом в масс-спектрах ЭИ органических соединений. Наличие пика молекулярного иона позволяет определить молекулярную массу неизвестного соединения и увеличивает достоверность идентификации. Общепринято, что отсутствие пика молекулярного иона в масс-спектрах ЭИ снижает достоверность идентификации неизвестных соединений, основанной на библиотечном поиске. В соответствии с результатами, полученными группой профессора Амирава⁷, порядка 30% соединений (из базы данных NIST'93) имеют малоинтенсивный (менее 1%) пик молекулярного иона в соответствующих масс-спектрах ЭИ. Насколько нам известно, эти данные не были подтверждены другими исследователями, а более подробная информация в литературе не представлена. В связи с этим нами было проведено статистическое исследование, целью которого являлось определение распределения органических соединений в зависимости от интенсивности пика молекулярного иона в масс-спектрах ЭИ. В работе было рассмотрено порядка 187000 соединений из базы данных масс-спектров NIST'08 и около 182000 соединений из базы данных масс-спектров Wiley 8th.

На Рис. 2 представлено распределение органических соединений в зависимости от интенсивности пика молекулярного иона ($I_{[M]}$), нормированной на основной пик. Распределения, полученные для соединений, представленных в базах данных NIST'08 и Wiley 8th, весьма близки (в случае интегрального распределения различие не

⁷ Информация об интенсивности пика молекулярного иона получена в рамках кандидатской диссертации (Dagan S. Mass spectrometry in supersonic molecular beams, PhD Thesis. Tel-Aviv University. 1996) и представлена на сайте: <http://www.tau.ac.il/chemistry/amirav/smbms.shtml>.

превышает 5%). Только порядка 26% соединений характеризуются интенсивным пиком молекулярного иона ($I_{[M]} > 50\%$). При этом пик молекулярного иона является основным пиком в масс-спектрах 14% соединений. Пик молекулярного иона малоинтенсивен или вообще отсутствует ($I_{[M]} \leq 1\%$) в случае порядка 24% соединений.

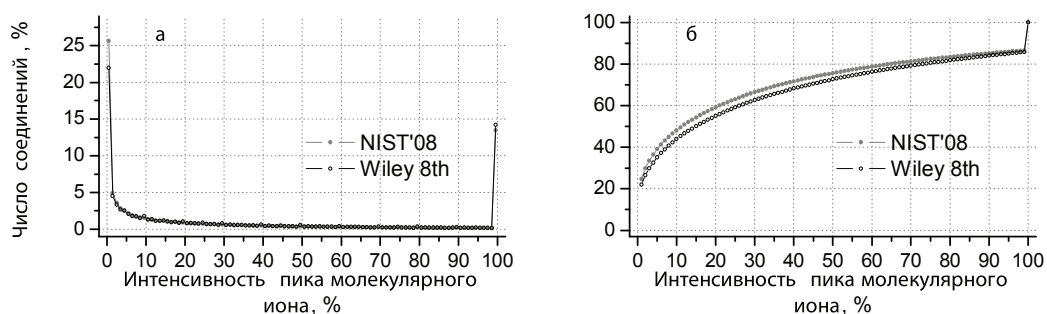


Рис. 2. Дифференциальное (а) и интегральное (б) распределения органических соединений в зависимости от интенсивности пика молекулярного иона (нормированной на основной пик).

Возможность достоверной идентификации неизвестных соединений, основанной на поиске по базам данных масс-спектров ЭИ.

Несмотря на широкое использование баз данных масс-спектров ЭИ и программ библиотечного поиска, на сегодняшний день не предложены граничные значения величины степени совпадения (ни для одного алгоритма библиотечного поиска), позволяющие говорить о достоверной идентификации неизвестного соединения. В связи с этим вещество считается идентифицированным, если значение величины степени совпадения для первого соединения в списке возможных кандидатов превышает установленную исследователем границу. Пороговое значение величины степени совпадения, используемое различными авторами, может изменяться в широких пределах, однако какое-либо обоснование выбора именно такого значения, как правило, отсутствует. Нами была поставлена цель – оценить надежность идентификации неизвестных соединений, основанной на проведении поиска по коммерческим базам данных масс-спектров ЭИ.

Из базы данных NIST'08 нами были выбраны 45 соединений, для каждого из которых в подбиблиотеке gerlib было представлено 5, 6 или 7 масс-спектров ЭИ. Общее число рассматриваемых масс-спектров составило 230. Каждый из масс-спектров был использован для проведения поиска по базе данных NIST'08 (подбиблиотека mainlib), поиск осуществляли с помощью программного обеспечения MS Search. В каждом случае во внимание принимали позицию заданного соединения в списке возможных кандидатов, значение величины Match, рассчитанное для заданного соединения, а также значение величины Match, рассчитанное для соединения, занимающего вторую (или первую в случае неправильной

идентификации) строчку в списке возможных кандидатов. Исходя из полученных данных, нами были определены величины ошибок первого и второго рода⁸ в зависимости от значения величины степени совпадения (Match), выбранного в качестве порогового (Табл. 2). Высокая величина ошибки первого рода для низких пороговых значений величины Match (ниже 850) объясняется тем, что заданное соединение не возглавляло список возможных кандидатов в 18,1% случаев. Как видно из Табл. 2, в общем случае (когда заранее неизвестно, присутствует ли масс-спектр исследуемого соединения в базе данных) вне зависимости от значения величины Match, выбранного в качестве порогового, вероятность неверной идентификации (ложноположительной или ложноотрицательной) превышает 40%.

Табл. 2. Величины ошибок первого и второго рода, возникающих при идентификации неизвестных соединений с использованием поиска по базам данных масс-спектров ЭИ, в зависимости от значения величины степени совпадения (Match), выбранного исследователем в качестве граничного.

Пороговое значение Match	Ошибка I рода	Ошибка II рода*	Ошибка II рода**
950	94,9	1,7	1,7
900	43,5	8,4	18,6
850	22,4	14,8	46,8
800	19,8	17,3	67,1
750	19,4	17,7	78,1
700	18,6	18,1	89,9

* величина ошибки II рода была рассчитана для случая, когда заданное соединение присутствует в базе данных масс-спектров;

** величина ошибки II рода была рассчитана для случая, когда заданное соединение отсутствует в базе данных масс-спектров.

Надежное сопоставление масс-спектров (зарегистрированных при одних и тех же условиях), основанное на совместном использовании метода главных компонент и t-критерия Стьюдента.

На основании представленных выше данных можно заключить, что идентификация, основанная на использовании баз данных масс-спектров ЭИ, в общем случае носит предположительный характер. В ряде случаев учет библиотечных (или расчетных) индексов удерживания может увеличить надежность идентификации. Однако в случае метода ГХ/МС наивысшая надежность идентификации может быть достигнута только при проведении анализа неизвестного и предполагаемого соединений в одних и тех же условиях и сопоставлении экспериментально полученных времен удерживания и масс-спектров ЭИ. Времена удерживания являются одномерными аналитическими сигналами, поэтому их сравнение не вызывает трудностей. В свою очередь, масс-спектры ЭИ – многомерны,

⁸ Ошибка первого рода имеет место в случае ложноотрицательной идентификации, ошибка второго рода – в случае ложноположительной идентификации.

и их сравнение является более сложной задачей. В практике химического анализа для сопоставления масс-спектров (зарегистрированных при одних и тех же условиях) обычно используют всего два подхода, однако оба подхода имеют весьма серьезные ограничения.

- Визуальное сравнение масс-спектров необъективно и не позволяет оценить «похожесть» масс-спектров количественно.
- Сравнение интенсивностей отдельных масс-спектральных пиков используется при проведении идентификации предполагаемых соединений⁹. Однако во внимание принимаются только несколько значений m/z (обычно три).

Нами был предложен подход к надежному сопоставлению масс-спектров (зарегистрированных при одних и тех же условиях), основанный на совместном использовании метода главных компонент и t -критерия Стьюдента. Краткое описание подхода представлено ниже. Несколько параллельных масс-спектров, отвечающих предположительно двум различным соединениям, должны быть зарегистрированы при одних и тех же экспериментальных условиях. Каждый масс-спектр (нормированный на основной пик) должен быть представлен в виде вектора-строки. Матрица, образованная путем объединения соответствующих векторов, должна быть центрирована и проанализирована с помощью метода главных компонент¹⁰.

Если масс-спектры отвечают разным соединениям, то первая главная компонента должна разделить объекты (то есть масс-спектры) на две группы. Если масс-спектры отвечают одному и тому же соединению, то различия между объектами обусловлены случайными факторами, скрытая структура данных отсутствует, а все главные компоненты (включая первую) содержат только шум. В результате при сопоставлении масс-спектров можно использовать значения только первой главной компоненты. Предполагая, что интенсивности масс-спектральных пиков подчиняются нормальному распределению, можно показать, что первая главная компонента

⁹ Данный способ сравнения спектров описан в ряде нормативных документов:

- Commission Decision 2002/657/EC of 12 August 2002, Official journal of the European Communities 2002, L221, 8–36.
- Guidance for industry: mass spectrometry for confirmation of the identity of animal drug residues: final guidance. FDA, Washington, 2003.
- WADA Technical Document – TD2010IDCR, 2011.

¹⁰ Матрица размером $K \times N$ может быть представлена в виде K точек в пространстве размерностью N . Основная идея метода главных компонент заключается в переходе к новой системе координат, оси которой получили название главных компонент. Вначале, как правило, проводят центрирование исходной матрицы, графическая интерпретация которого заключается в смещении начала координат N -мерного пространства в центр облака данных. Затем определяют направление наибольшего изменения данных. Это направление и является осью первой главной компоненты. Каждая следующая главная компонента выбирается с тем расчетом, чтобы она была ортогональна уже выбранным главным компонентам и также определяла направление наибольшего изменения данных.

распределена нормально. Таким образом, для сравнения значений первой главной компоненты может быть использован *t*-критерий Стьюдента.

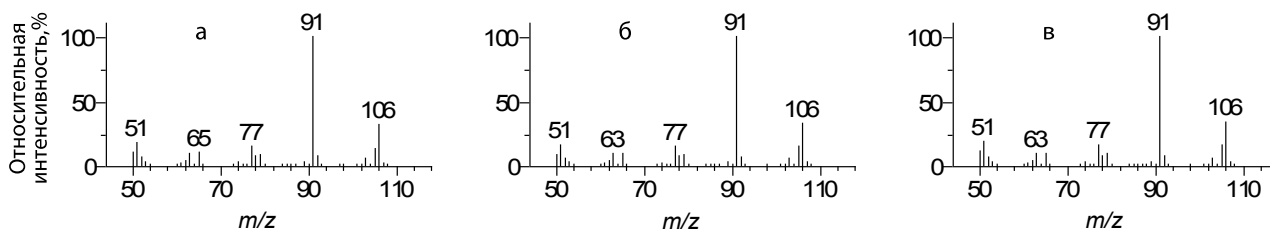


Рис. 3. Масс-спектры ЭИ *o*-диметилбензола (а), *m*-диметилбензола (б) и *n*-диметилбензола (в), зарегистрированные в идентичных условиях с помощью времяпролетного масс-спектрометра. В хроматографическую колонку было введено около 1 нг аналитов.

Возможность использования предложенного подхода для нахождения различий между практически идентичными масс-спектрами была показана на примере нескольких групп органических соединений:

- *o*-, *m*-, *n*-диметилбензол и этилбензол;
- 1-октанол, 1-нонанол, 1-деканол, 1-ундеканол, 1-додеканол и 1-тетрадеканол;
- бенз[б]флуорантен, бенз[к]флуорантен и бенз[а]пирен.

Для каждой пары соединений, принадлежащих одной и той же группе, было найдено статистически значимое различие между соответствующими масс-спектрами (даже для доверительной вероятности $P = 99,9\%$). В качестве примера на Рис. 3 представлены масс-спектры *o*-, *m*-, *n*-диметилбензола, зарегистрированные при одних и тех же условиях с использованием времяпролетного масс-спектрометра. Как видно из Рис. 3, спектры содержат один и тот же набор пиков и визуально не отличаются друг от друга. Результаты сопоставления масс-спектров с использованием метода главных компонент и *t*-критерия Стьюдента обобщены в Табл. 3.

Табл. 3. Рассчитанные для каждой пары соединений значения коэффициента Стьюдента ($n_1 = n_2 = 5$). Табличное значение коэффициента Стьюдента: $t_{\text{крит.}}(P = 99,9\%; f = 8) = 5,0$.

	<i>o</i> -диметилбензол	<i>m</i> -диметилбензол	<i>n</i> -диметилбензол
<i>o</i> -диметилбензол	–	16	12
<i>m</i> -диметилбензол	16	–	10
<i>n</i> -диметилбензол	12	10	–

Возможность достоверной идентификации следовых количеств неизвестных органических соединений с помощью метода ГХ/МС(ЭИ). Предложенный нами подход к сопоставлению масс-спектров может быть использован для повышения достоверности идентификации следовых количеств неизвестных соединений, когда заключительным этапом является анализ неизвестного и предполагаемого соединений

методом ГХ/МС(ЭИ) в идентичных экспериментальных условиях. Раствор 2,4-диметилфенола (концентрацией 0,1 нг мкл⁻¹) был рассмотрен нами в качестве «неизвестного»¹¹ образца, а хроматографический пик 2,4-диметилфенола – в качестве пика «неизвестного» соединения, идентификацию которого нужно было провести.

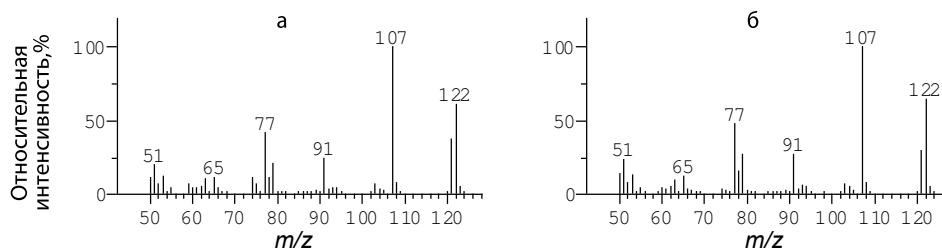


Рис. 4. Масс-спектры ЭИ «неизвестного» соединения (а) и 2,5-диметилфенола (б), зарегистрированные в идентичных условиях при использовании времяпролетного масс-спектрометра.

«Неизвестное» соединение было охарактеризовано масс-спектром ЭИ (Рис. 4-а) и индексом удерживания ($RI_{\text{«неизвестное» соед.}} \sim 1155$), которые были использованы для поиска по соответствующим базам данных. По результатам поиска наиболее вероятным кандидатом являлся 2,5-диметилфенол ($RI_{2,5\text{-диметилфенол}} \sim 1152$ ¹²). Раствор 2,5-диметилфенола (концентрацией 0,1 нг мкл⁻¹) и «неизвестный» образец были проанализированы при одних и тех же экспериментальных условиях. Разница между временами удерживания «неизвестного» и предполагаемого соединения составила всего 0,15 с (или 0,06%). Визуальное различие между зарегистрированными масс-спектрами ЭИ отсутствовало (Рис. 4). Исходя из полученных данных, можно было бы сделать обоснованный вывод, что «неизвестное» соединение является 2,5-диметилфенолом, однако это была бы ошибочная идентификация.

Чтобы увеличить надежность идентификации, три параллельных масс-спектра «неизвестного» соединения и 2,5-диметилфенола были зарегистрированы в идентичных условиях и сопоставлены с помощью метода главных компонент и *t*-критерия Стьюдента. График счетов для первой и второй главных компонент представлен на Рис. 5-а. Объекты, отвечающие масс-спектрам «неизвестного» соединения и возможного кандидата, разделены первой главной компонентой. Рассчитанное значение коэффициента Стьюдента было равно 36 и значительно превышало табличное значение $t_{\text{крит.}}(P = 99,9\%; f = 4) = 8,6$. Наличие статистически значимого различия между масс-спектрами «неизвестного» соединения и возможного кандидата позволило избежать проведения ошибочной идентификации.

¹¹ В данном случае термин «неизвестный» написан в кавычках, поскольку в исследовании использовали раствор, известного нам состава.

¹² Представленный индекс удерживания был экспериментально получен в следующей работе: *Mjos S.A., Meier S., Boitsov S. Alkylphenol retention indices, // J. Chromatogr. A. 2006. Vol. 1123. N 1. 98–105.*

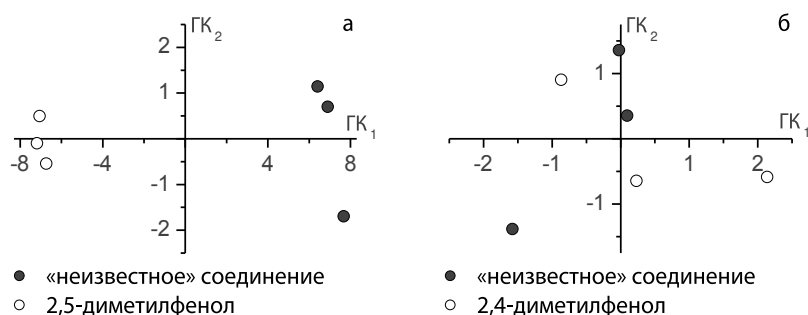


Рис. 5. Графики счетов для первой и второй главных компонент, полученные в процессе сравнения масс-спектров «неизвестного» соединения с масс-спектрами 2,5-диметилфенола (а) и 2,4-диметилфенола (б). Каждая точка на графике представляет собой масс-спектр. $ГК_1$ и $ГК_2$ – оси первой и второй главной компоненты. Высоты хроматографических пиков «неизвестного» соединения и предполагаемого кандидата были приблизительно равны.

После исключения 2,5-диметилфенола из рассмотрения наиболее вероятным кандидатом стал 2,4-диметилфенол ($RI_{2,4\text{-диметилфенол}} \sim 1150$). Масс-спектры «неизвестного» соединения и 2,4-диметилфенола (зарегистрированные при одних и тех же условиях) были сопоставлены с использованием описанного выше подхода. График счетов для первой и второй главных компонент представлен на Рис. 5-б. В данном случае табличное значение коэффициента Стьюдента $t_{\text{крит.}}(P = 95\%; f = 4) = 2,8$ превышало рассчитанное ($\xi = 1,0$), исходя из чего был сделан вывод о том, что «неизвестное» соединение является 2,4-диметилфенолом.

В процессе проведения данного исследования было установлено, что масс-спектры ЭИ, зарегистрированные для разных количеств аналита, могут статистически значимо отличаться друг от друга. В связи с этим при использовании данного подхода к сопоставлению масс-спектров в реальной практике необходимо подбирать концентрации растворов таким образом, чтобы высоты хроматографических пиков неизвестного и предполагаемого соединений были приблизительно одинаковыми.

На заключительном этапе исследования нами была рассмотрена возможность увеличения надежности идентификации неизвестного соединения при анализе реального объекта. В качестве объекта исследования был выбран фармацевтический препарат Ксеникал. Один из компонентов, зарегистрированных в процессе анализа, был идентифицирован с использованием базы данных масс-спектров NIST'08 как 1-додеканол. Для увеличения надежности идентификации образец фармацевтического препарата и модельный раствор 1-додеканола были проанализированы при одних тех же экспериментальных условиях, а зарегистрированные масс-спектры были сопоставлены с помощью метода главных компонент и t -критерия Стьюдента. Статистически значимое различие между масс-спектрами обнаружено не было ($t_{\text{крит.}}(P = 95\%; f = 4) = 2,8 > \xi = 1,9$).

Предложенный нами подход к сопоставлению масс-спектров (основанный на совместном использовании метода главных компонент и t -критерия Стьюдента) позволяет значительно увеличить надежность идентификации следовых количеств аналитов в случае проведения анализа неизвестного соединения и предполагаемого кандидата (выбранного, например, на основании библиотечного поиска) методом ГХ/МС(ЭИ) в идентичных условиях.

Использование математических подходов для выделения «чистых» масс-спектров. Возможности метода ГХ/МС позволяют за один анализ проводить разделение и определение более 100 соединений. Однако при работе со сложными многокомпонентными смесями возможно наложение пиков исследуемых веществ друг на друга или на мешающие компоненты матрицы, что может затруднить процедуру идентификации неизвестных соединений. В связи с этим одной из актуальных задач является разработка математических подходов, позволяющих проводить выделение «чистых» масс-спектров соединений, хроматографические профили которых разделены неполностью.

При наложении хроматографических пиков соединений, масс-спектры которых не содержат общих значений m/z , выделение «чистых» и неискаженных масс-спектров обычно не представляет сложностей. Гораздо сложнее провести очистку масс-спектров, содержащих большое число общих значений m/z . При использовании алгоритмов библиотечного поиска небольшое искажение интенсивностей масс-спектральных пиков не приведет к существенному изменению величины степени совпадения. Однако в случае сравнения масс-спектров с использованием метода главных компонент даже незначительные ошибки на стадии выделения «чистых» масс-спектров могут привести к неправильной идентификации.

В связи с этим нами была рассмотрена возможность выделения «чистых» и неискаженных масс-спектров коэлюируемых соединений с помощью метода «вычитания фона»¹³, при использовании прокрустового анализа¹⁴, а также с помощью коммерческих компьютерных программ AMDIS и ChromaTOF, предназначенных для

¹³ Данный метод заключается в вычитании из масс-спектра, зарегистрированного в максимуме хроматографического пика, усредненного масс-спектра близлежащей области хроматограммы. Обычно этот подход используют для исключения из экспериментального масс-спектра значений m/z , отвечающих фону прибора, однако в ряде случаев он может быть успешно применен для очистки масс-спектров коэлюируемых соединений.

¹⁴ Прокрустов анализ основан на использовании метода главных компонент, позволяющего провести разложение матрицы, содержащей хромато-масс-спектральные данные, на матрицу счетов и матрицу нагрузок. Линейное преобразование (масштабирование и вращение) матрицы нагрузок позволяет перейти к матрице «чистых» масс-спектров (однако для этого нужно обладать информацией о направлении осей чистых соединений в пространстве главных компонент). В настоящей работе элементы матрицы перехода рассчитывали, используя априорную информацию о характеристичных значениях m/z коэлюируемых соединений.

автоматической обработки хромато-масс-спектральных данных. Использование смоделированных хромато-масс-спектральных данных позволило варьировать хроматографическое разрешение компонентов, отношение концентраций компонентов, величину шума, а также число общих значений m/z , присутствующих в масс-спектрах коэлюируемых соединений.

При наличии общих значений m/z в масс-спектрах коэлюируемых соединений только прокрустов анализ позволил во всех случаях выделить совершенно неискаженные масс-спектры, которые не имели статистически значимого различия с исходными спектрами, используемыми при моделировании данных. Метод «вычитания фона» позволил получить приемлемые результаты только при отсутствии общих значений m/z в масс-спектрах коэлюируемых соединений. Масс-спектры, очищенные с помощью программного обеспечения AMDIS, во многих случаях были сильно искажены (масс-спектральные пики, отвечающие общим значения m/z , либо вовсе не были извлечены, либо их интенсивности были извлечены некорректно). В случае программного обеспечения ChromaTOF (при условии, что отношение концентраций компонентов не превышало 10) очищенные масс-спектры визуально не отличались от исходных спектров, используемых при моделировании хромато-масс-спектральных данных (однако использование метода главных компонент в ряде случаев показало наличие статистически значимых различий между очищенными и исходными спектрами).

Нами показано, что прокрустов анализ может быть использован для выделения «чистых» масс-спектров неизвестного соединения и предполагаемого кандидата при проведении их сопоставления с использованием метода главных компонент и t -критерия Стьюдента. Однако для применения данного подхода необходима априорная информация о числе химических компонентов (присутствующих на рассматриваемом фрагменте хроматограммы) и характеристичных значениях m/z коэлюируемых соединений. В реальной практике данная информация может быть получена, например, с помощью программного обеспечения ChromaTOF.

ВЫВОДЫ

- Проведено сопоставление масс-спектров ЭИ, зарегистрированных в неидентичных условиях. Показано, что различия между масс-спектрами проявляются в основном не в наборе пиков фрагментных ионов, а в значениях их интенсивностей.
- Впервые определено распределение органических соединений (представленных в базах данных масс-спектров ЭИ) в зависимости от интенсивности пика молекулярного иона. Установлено, что пик молекулярного иона отсутствует в масс-спектрах ЭИ порядка четверти органических соединений (что снижает достоверность результатов библиотечного поиска).
- Экспериментально показано, что результаты поиска по базам данных масс-спектров ЭИ следует рассматривать как ориентировочные, поскольку нет однозначной корреляции между величиной степени совпадения и достоверностью идентификации.
- Предложен подход, позволяющий значительно увеличить достоверность идентификации следовых количеств аналитов при использовании метода ГХ/МС(ЭИ). Подход основан на выборе наиболее вероятного кандидата с помощью библиотечного поиска и последующем сопоставлении экспериментальных масс-спектров неизвестного и предполагаемого соединений (зарегистрированных при одних и тех же условиях) с помощью метода главных компонент и t -критерия Стьюдента.
- Проведено сопоставление ряда математических подходов для выделения «чистых» и неискаженных масс-спектров коэлюируемых соединений. Показано, что при наличии общих значений m/z в масс-спектрах коэлюируемых соединений и высоком содержании мешающего компонента только прокрустов анализ позволяет выделить неискаженные масс-спектры.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Самохин А.С., Ревельский И.А. Применение метода главных компонент для выделения «чистых» масс-спектров в газохроматографическом – масс-спектральном анализе. // Масс-спектрометрия. 2010. Т. 7. № 2. С. 132–138.
2. Самохин А.С., Ревельский А.И., Чепелянский Д.А., Ревельский И.А. Возможность достоверной идентификации неизвестных соединений при использовании программы MS Search и коммерческой базы данных масс-спектров электронной ионизации. // Масс-спектрометрия. 2011. Т. 8. № 1. С. 65–67.
3. Самохин А.С., Ревельский А.И., Ревельский И.А. Сопоставление масс-спектров электронной ионизации ряда соединений ($MM < 200$ Да), зарегистрированных с помощью различных типов масс-спектрометров. // Масс-спектрометрия. 2011. Т. 8. № 2. С. 106–110.
4. Samokhin A., Revelsky I. Distinguishing by principal component analysis o-xylene, m-xylene, p-xylene and ethylbenzene using electron ionization mass spectrometry. // Eur. J. Mass Spectrom. 2011. Vol. 17. N 5. P. 477–480.
5. Самохин А.С., Ревельский И.А. Интенсивность пика молекулярного иона в масс-спектрах электронной ионизации. // Масс-спектрометрия. 2012. Т. 9. № 1. С. 58–60.
6. Самохин А.С., Ревельский И.А. Надежное сопоставление масс-спектров, зарегистрированных в идентичных условиях, с помощью метода главных компонент. // Аналитика и контроль. 2012. Т. 16. № 3. С. 269–274.
7. Samokhin A., Revelsky I. Reliable comparison of electron ionization mass spectra recorded under identical experimental conditions. // Spectrosc. Lett. 2013. Vol. 46. N 2. P. 117–121.
8. Самохин А.С., Ревельский А.И., Ревельский И.А. Применение метода главных компонент для выделения «чистых» масс-спектров в ГХ-МС анализе. Тезисы докладов III Всероссийской конференции с международным участием «Масс-спектрометрия и ее прикладные проблемы», 18 – 22 мая, Москва, 2009. С. 35.
9. Samokhin A.S., Revelsky I.A. Application of principal component analysis for extraction of pure mass spectra from gas chromatography/mass spectrometry data. Book of Abstracts. 7th Winter Symposium on Chemometrics «Modern Methods of Data Analysis». February 15 – 19, 2010, Saint Petersburg, Russia. P. 37–38.
10. Самохин А.С., Ревельский И.А., Ревельский А.И. Увеличение достоверности идентификации неизвестных соединений при использовании баз данных масс-спектров электронной ионизации. Тезисы докладов Четверной Всероссийской конференции «Фундаментальные вопросы масс-спектрометрии и ее аналитические применения», 10 – 14 октября, Звенигород, 2010. С. 58–59.

11. Самохин А.С., Рыбальченко И.В., Ревельский И.А. Различие между масс-спектрами электронной ионизации, зарегистрированными в неидентичных условиях. Тезисы докладов IV Всероссийской конференции с международным участием «Масс-спектрометрия и ее прикладные проблемы», 5 – 9 сентября, Москва, 2011. С. 28.
12. Самохин А.С., Ревельский И.А. Статистическое исследование интенсивности пика молекулярного иона в масс-спектрах электронной ионизации. Тезисы докладов IV Всероссийской конференции с международным участием «Масс-спектрометрия и ее прикладные проблемы», 5 – 9 сентября, Москва, 2011. С. 47.
13. Ревельский И.А., Самохин А.С., Леднева А.В., Чепелянский Д.А., Бурмыкин Д.А., Ревельский А.И. Достоверное установление молекулярной массы – необходимое условие надежной идентификации следов компонентов смесей. Тезисы докладов IV Всероссийской конференции с международным участием «Масс-спектрометрия и ее прикладные проблемы», 5 – 9 сентября, Москва, 2011. С. 58.
14. Samokhin A.S., Revelsky I.A. Differentiation of electron ionization mass spectra of o-, m-, p-xylene and ethylbenzene using principal component analysis. Book of Abstracts. Conferentia Chemometrica 2011. September 18 – 21, 2011, Sumeg, Hungary. P. P35.
15. Samokhin A.S., Revelsky I.A. Identification of unknown organic compounds based on comparison of electron ionization mass spectra. Book of Abstracts. 8th Winter Symposium on Chemometrics «Modern Methods of Data Analysis». February 27 – March 2, 2012, Drakino, Russia. P. 24–25.
16. Samokhin A.S., Revelsky I.A. Reliable comparison of electron ionization mass spectra of unknown compound and possible candidate registered under the same experimental conditions. Book of Abstracts. XIII Chemometrics in Analytical Chemistry. June 25 – 29, 2012, Budapest, Hungary. P. 90.
17. Самохин А.С., Ревельский И.А. Возможность использования масс-спектров электронной ионизации для проведения достоверной идентификации неизвестных органических соединений. Тезисы докладов Всероссийской конференции по аналитической спектроскопии с международным участием, 23 – 29 сентября, Краснодар, 2012. С. 205.
18. Samokhin A.S., Revelsky A.I., Revelsky I.A. Absence of spectral skew, high-speed acquisition, automatic peak finding and mass spectra purification as key aspects of non-target GC/MS analysis. Book of Abstracts. 13th European Meeting on Environmental Chemistry. December 5 – 8, 2012, Moscow, Russia. P. 24.
19. Samokhin A.S., Revelsky I.A. Extraction of unbiased mass spectra in case of unresolved chromatographic peaks. Book of Abstracts. 13th European Meeting on Environmental Chemistry. December 5–8, 2012, Moscow, Russia. P. 45.

Автор работы выражает искреннюю благодарность:

- своему научному руководителю *И.А. Ревельскому*;
- компании *Leco* за предоставление времяпролетных масс-спектрометров *TruTOF HT* и *Pegasus HT* для проведения исследований;
- *А.В. Гармашу, Э.Д. Вирюсу и А.И. Ревельскому* за обсуждение результатов, представленных в работе;
- *О.В. Поляковой и А.Т. Лебедеву* за предоставление смеси среднелетучих органических соединений, используемой в работе;
- *И.В. Рыбальченко* за предоставление базы данных *OSAD*;
- *А. Шнайдеру* за предоставление образца фармацевтического препарата *Ксеникал*, исследуемого в работе;
- *А. Фиалкову и А. Амираву* за предоставление фрагмента кандидатской диссертации *С. Дагана*.