

УДК 543.51:539.1.09*54-128

СРАВНЕНИЕ ХАРАКТЕРИСТИК МАСС-СПЕКТРОВ ПРИ ПОМОЩИ МЕТОДОВ СТАТИСТИЧЕСКОГО АНАЛИЗА ДЛЯ СЛУЧАЯ ИОНИЗАЦИИ ОРГАНИЧЕСКИХ МОЛЕКУЛ ЭЛЕКТРОННЫМ УДАРОМ С РАЗЛИЧНОЙ ЭНЕРГИЕЙ ЭЛЕКТРОНОВ

© 2024 г. С. В. Силкин¹, А. В. Сахаров¹, С. И. Пеков^{2,3}, В. А. Елиферов¹, В. Г. Ткаченко¹,
Д. В. Колесник¹, Е. Н. Николаев^{2,*}, И. А. Попов^{1,3,**}

¹Московский физико-технический институт, Долгопрудный, Россия

²Сколковский институт науки и технологий, Москва, Россия

³Сибирский государственный медицинский университет, Томск, Россия

*E-mail: e.nikolaev@skoltech.ru

**E-mail: popov.ia@mipt.ru

Поступила в редакцию 05.04.2024 г.

После доработки 17.06.2024 г.

Принята к публикации 18.06.2024 г.

Повышение чувствительности и точности идентификации летучих органических соединений в случае их исследования при помощи масс-спектрометрических газоанализаторов с ионизацией на основе электронного удара может быть осуществлено путем увеличения числа образующихся в ионном источнике ионов за счет варьирования энергии ионизирующих электронов. В работе представлено сравнение данных, полученных ионизацией электронным ударом при значениях энергии электронов (ЭЭ) в диапазоне 25–105 эВ для ряда органических веществ, принадлежащих различным классам органических соединений. Для интерпретации динамики изменения интенсивностей пиков ионов-фрагментов был проведен анализ с использованием матриц подобия, основанных на различных метриках сходства и было показано влияние ЭЭ на вероятность образования основных фрагментарных частиц исследуемых веществ и, соответственно, на сходство регистрируемого масс-спектра с эталонным масс-спектром из базы данных.

Ключевые слова: масс-спектрометрия, электронный удар, ионизация, фрагментация, молекулярный ион

DOI: 10.31857/S0023119324060071 EDN: TNOJON

ВВЕДЕНИЕ

Одним из наиболее распространенных методов ионизации в масс-спектрометрии органических соединений является электронный удар [1]. Распространенность этого метода связана с его простотой, надежностью и универсальностью, что позволяет получать воспроизводимые масс-спектрометрические данные. Ионизация электронным ударом при количественном измерении масс-спектров обычно осуществляется при энергии электронов, определяемой в условиях конкретного эксперимента, и находится в интервале от 40 до 90 эВ, в большинстве методов энергию электронов конвенционально устанавливают равной 70 эВ [2, 3]. Эта величина выбрана масс-спектрометрическим сообществом исходя из близости интенсивностей пиков в масс-спектрах к максимуму при ионизации большинства нейтральных молекул и из-за отсутствия

при такой энергии ионизации двухзарядных ионов [3]. Основные базы данных масс-спектров (БД-МС) содержат спектры, измеренные при энергии электронного удара в 70 эВ. Самыми известными из них являются референсные базы NIST и Wiley. В настоящее время разрабатываются новые базы данных масс-спектров с ионизацией электронным ударом и новыми быстрыми алгоритмами их сопоставления [4].

В практических аналитических задачах при идентификации соединений по масс-спектрам электронного удара принято сравнивать измеренный спектр со спектральными данными из упомянутых выше БД-МС. Их прямое совпадение позволяет констатировать идентификацию конкретного вещества. Однако масс-спектры аналитов из-за особенностей масс-спектрометров и отличий условий измерения масс-спектров, даже

после разделения на хроматографической колонке, часто имеют лишь частичное совпадение с данными БД-МС, что приводит к неопределенности в интерпретации спектральных данных. Для решения этой проблемы используют различные методики обработки данных и метрики сходства спектров. Типичные шаги при этом включают удаление фонового спектра, программное подавление шумов, а также разделение перекрывающихся пиков. После проведения процедуры обработки данных становится возможным сравнивать спектры, применяя различные метрики сходства, такие как косинусная мера, коэффициент корреляции Пирсона, евклидово и манхэттенское расстояние [5–7].

Каждая из этих метрик имеет свои особенности, преимущества и недостатки. Косинусная мера позволяет сравнивать вектора в многомерном пространстве. Ее величина лежит в диапазоне от -1 до 1 [8, 9]. Большая величина косинусной меры говорит о высокой схожести векторов, а значит, схожести соответствующих им масс-спектров. Необходимо заметить, что так как вектора, соответствующие масс-спектрам, имеют только неотрицательные компоненты, косинусная мера может принимать лишь неотрицательные значения, т.е. в рассматриваемой задаче ее величина лежит в диапазоне от 0 до 1 . Такую метрику называют нормированной. Похожей метрикой является коэффициент корреляции Пирсона [10]. Она позволяет установить линейную зависимость между двумя многомерными векторами. К недостаткам этой метрики можно отнести чувствительность к выбросам и предположение о линейной зависимости данных. Величина коэффициента корреляции Пирсона лежит в диапазоне от -1 до 1 , то есть эта метрика не является нормированной, что представляется еще одним недостатком по сравнению с косинусной мерой. Евклидово и манхэттенское расстояния являются мерами дистанций между точками в многомерном пространстве, что позволяет учитывать изменение абсолютных величин интенсивностей пиков в масс-спектрах [11–13]. К существенным недостаткам этих метрик можно отнести их ненормированность, а также невозможность интерпретации абсолютных значений расстояний. Другими словами, уместно говорить лишь о сравнении расстояний, а не их значениях по отдельности, что нивелирует эффект влияния общей интенсивности масс-спектра на величину метрики.

Увеличение эффективности ионизации может не только повышать чувствительность масс-спектрометрического анализа, но и увеличивать точность идентификации соединений [14–16]. При ионизации электронным ударом поиск оптимальных условий эффективной ионизации может осуществляться путем варьирования величины энергии электронного удара. Действительно, органические молекулы, относящиеся к различным классам органических соединений, будут по-разному фрагментироваться под действием электронного удара. Варьирование величины ЭЭ, вероятно, может влиять на изменение фрагментарного состава органических молекул, относящихся к разным классам соединений. Качественный состав совокупности осколочных и перегруппировочных фрагментов при этом не должен меняться в силу относительной устойчивости химических связей, образующих ионизируемую молекулу [17].

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Эксперименты проводились на высоколетучих веществах, принадлежащих к различным классам органических соединений – *n*-гексан, 2,2,4-триметилпентан, циклогексан, нитрометан, хлороформ, изопропанол, диэтиловый эфир, тетрагидрофуран, этилформиат, этилацетат, изопропилацетат, бутилацетат, толуол, анилин и пиридин. Все вещества соответствовали маркам ЧДА или ОСЧ. Предварительно чистоту выбранных органических соединений проверяли методом газовой хроматографии, совмещенной с масс-детектором с помощью ГХ-МС Альфа Маэстро (Интерлаб, Россия).

Спектры измеряли с помощью газового масс-анализатора Residual Gas Analyzer (RGA-300) с прямым интерфейсом ввода пробы. Данный анализатор представляет собой масс-спектрометр квадрупольного типа и позволяет изменять ЭЭ от 25 до 105 эВ, поэтому были выбраны следующие энергии электронов: 25 , 45 , 70 , 90 и 105 эВ.

В эксперименте добивались устойчивого сигнала от молекул атмосферного воздуха при энергии электронного удара 70 эВ и записывали фоновые масс-спектры при скорости сканирования 1491 а.е.м/с в диапазоне 1 – 150 а.е.м. Затем в анализатор вводилась проба исследуемого вещества из виалы, и измерялся масс-спектр без изменения параметров анализатора. Эксперимент проводился с трехкратным повторением. Аналогично получались масс-спектры при остальных значениях ЭЭ. Перед началом ввода пробы очередного вещества входной интерфейс

промывали смесью растворителей с последующей дегазацией, в результате чего удалось добиться отсутствия сигналов от них.

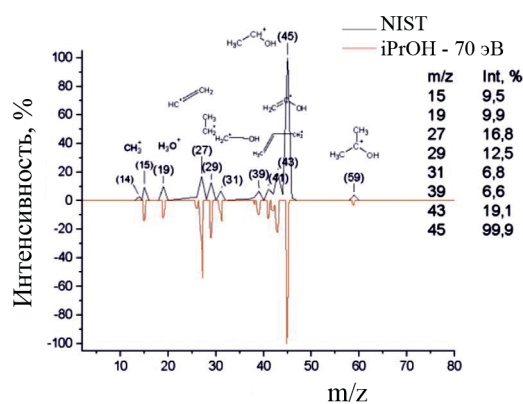
Для получения корректных масс-спектров органических веществ из спектров аналитов вычитались соответствующие фоновые спектры при каждой энергии ионизации в автоматическом режиме. Фоновые спектры содержали пики с m/z равными 2, 14, 16, 18, 28, 32, 40 и 44 Да, соответствующих основному составу атмосферного воздуха.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

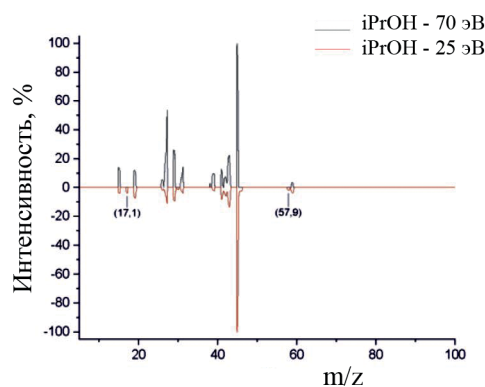
Среди проанализированных веществ присутствуют алифатические органические соединения без функциональных групп, алкилгалогениды, нитрозамещенные алканы, одноатомные спирты, простые и сложные эфиры, а также ароматические соединения, содержащие гетероатом в цикле и, соответственно, не содержащие его. Представленные в данной работе вещества не

покрывают все классы органических соединений, но позволяют наблюдать общие закономерности изменения масс-спектров при изменении ЭЭ. Гораздо больший интерес вызывает сравнение гомологов, изомеров или соединений с единственным различием, например, функциональной группы.

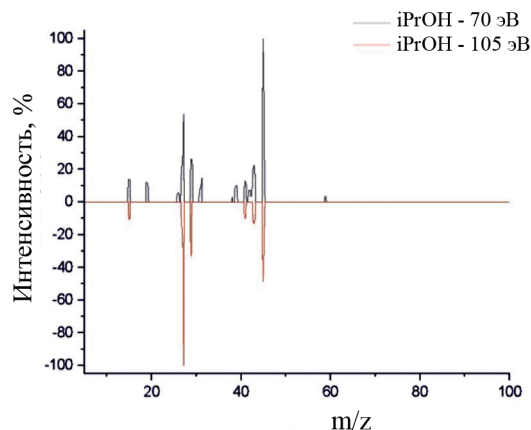
После нормирования интенсивностей пиков на максимальную величину ионного тока, полученные масс-спектры органических веществ при ионизации электронным ударом в 70 эВ сравнивали с таковыми из базы данных NIST 2020 (рис. 1а). Для вычисления метрик сходства масс-спектры представлялись в виде векторов одинаковой размерности, для которых можно использовать классические метрики: косинусную меру, коэффициент корреляции Пирсона, евклидово и манхэттенское расстояния. Для каждого вещества были составлены матрицы сходства из приведенных выше метрик [6]. Со-



(a)



(б)



(в)

Рис. 1. Масс-спектры изопропанола, полученные ионизацией электронным ударом. (а) при 70 эВ в сопоставлении со спектром из БД NIST; (б) при 25 эВ в сопоставлении со спектром, полученным при 70 эВ; (в) при 105 эВ в сопоставлении со спектром, полученным при 70 эВ.

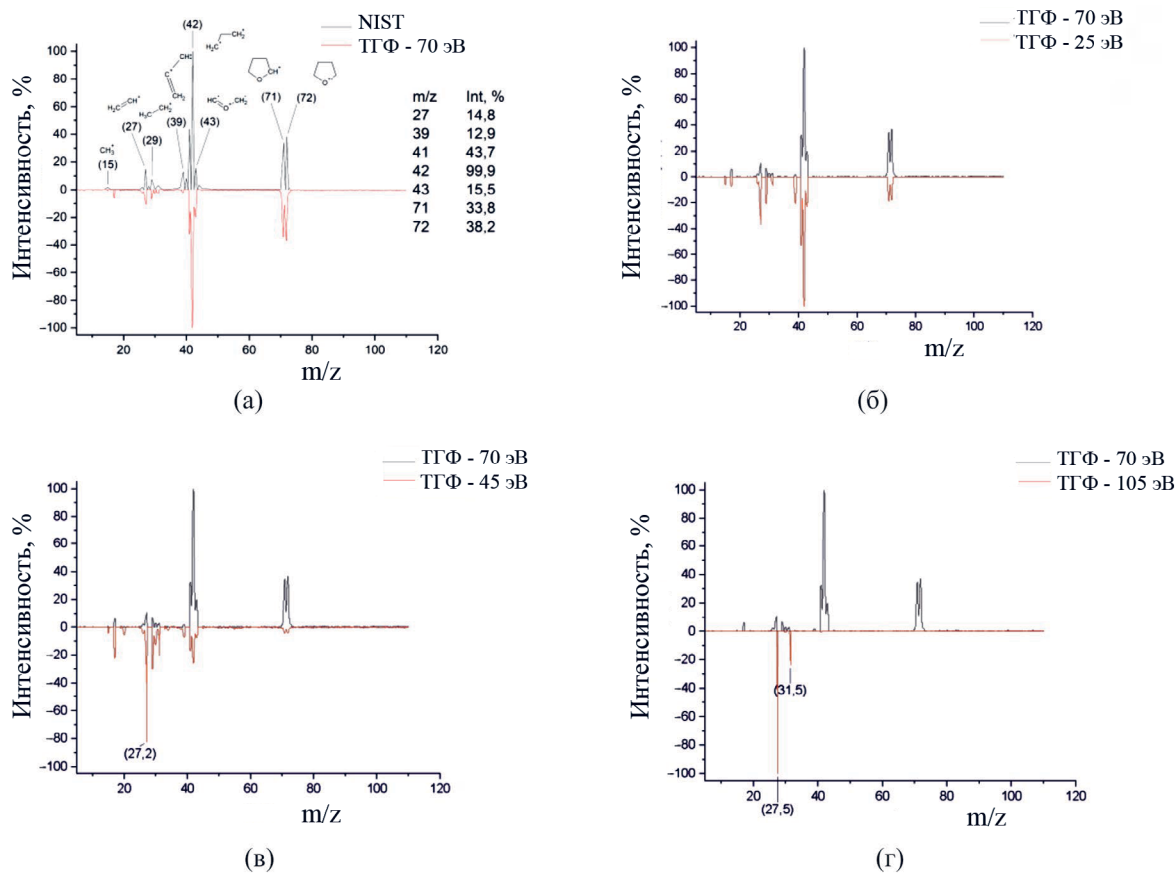


Рис. 2. Масс-спектры ТГФ, полученные ионизацией электронным ударом (а) при 70 эВ в сопоставлении со спектром из БД NIST; (б) при 25 эВ в сопоставлении со спектром, полученным при 70 эВ; (в) при 45 эВ в сопоставлении со спектром, полученным при 70 эВ; (г) при 105 эВ в сопоставлении со спектром, полученным при 70 эВ.

ставление матриц осуществлялось с помощью написанной нами на языке Python программы, включающей загрузку спектральных данных, их предобработку, вычисление метрик и визуализацию результатов.

При детальном анализе полученных при различных ЭЭ спектров, нами учитывались только ионные фрагменты, соответствующие значениям m/z из базы данных NIST. При этом не принимались во внимание пики, интенсивность которых оказывалась менее 0.9% от максимального. Каждое массовое число в спектре из БД-МС NIST было соотнесено с ионным фрагментом молекулы и сравнивалось с полученным спектром при ЭЭ в 70 эВ, затем этот спектр сравнивался со спектрами, полученными при других энергиях электронного удара (рис. 1). Как видно из рис. 1, в масс-спектре изопропанола, измеренного при ЭЭ = 25 эВ, наблюдается пик, соответствующий молекулярному иону, а в масс-спектре, регистрируемом при ЭЭ = 105 эВ, он исчезает.

Максимальные интенсивности пиков в масс-спектрах, представленных в данной работе веществ, соответствовали наиболее устойчивым ионам. Так, например, в спектрах изопропанола пик с максимальной интенсивностью соответствует этанольному карбкатиону, полученному при отщеплении одной метильной группы. При этом интенсивность изопропанольного карбкатиона оказывается достаточно низкой (3.8%) несмотря на то, что такой карбкатион более стабилен. Вместо этого наблюдается максимальный пик с массовым числом 45, соответствующий иону $C_2H_5O^+$, что объясняется правилом выброса максимального алкила [2]. Таким образом, интенсивность пика зависит не только от устойчивости иона, но и от геометрического строения молекулы. Это позволяет различать изомеры при регистрации масс-спектров по соотношению интенсивностей пиков [8, 9, 18]. Также максимальной или, по крайней мере, весьма высокой интенсивностью обладают пики,

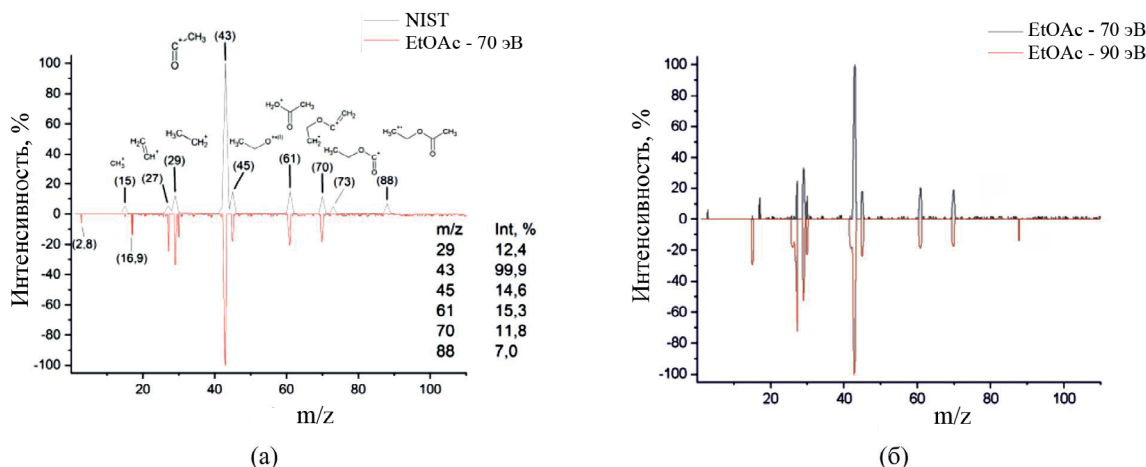


Рис. 3. Масс-спектры этилацетата, полученные ионизацией электронным ударом (а) при 70 эВ в сопоставлении со спектром из БД NIST; (б) при 90 эВ в сопоставлении со спектром, полученным при 70 эВ.

соответствующие реакции перегруппировки по Мак-Лафферти. Такое поведение демонстрируют, например, сложные эфиры. Вместе с тем при электронном ударе могут образовываться малоустойчивые катионы и катион-радикалы, которые тем не менее могут быть детектированы, но с низкой интенсивностью. Например, образование крайне неустойчивого арильного карбокатиона при регистрации масс-спектра толуола маловероятно, в то время как устойчивый аллильный карбокатион дает сигнал в масс-спектре изопропанола.

Понижение ЭЭ ниже 70 эВ ожидаемо приводит к меньшей фрагментации и, как следствие, повышению интенсивности молекулярного иона (рис. 1) [14]. Однако в случае, если молекулярный ион не наблюдается в спектре при 70 эВ, то при понижении ЭЭ его появление не ожидается по причине неустойчивости конкретного молекулярного иона. Фрагментация исходной молекулы происходит даже при значительном понижении ЭЭ. В случае же устойчивого молекулярного иона малые ионные фрагменты молекулы могут не наблюдаться при снижении ЭЭ либо их интенсивность оказывается существенно ниже референсной интенсивности, представленной в БД. Следует отметить, что интенсивности сигналов от ионных фрагментов в абсолютных величинах падают относительно регистрации спектра при 70 эВ.

Повышение ЭЭ относительно 70 эВ приводит к росту фрагментации молекул. При этом качественного изменения ионных форм не наблюдается, происходит лишь более глубокая фрагментация молекул (рис. 2). Так, например, масс-спектры

ТГФ, полученные при ионизации электронным ударом (рис. 2), демонстрируют наличие молекулярного иона при 25 эВ, уменьшение его интенсивности при 45 эВ и отсутствие при 105 эВ. При этом в спектре, полученном при 105 эВ, растут интенсивности пиков, соответствующих малым ионным фрагментам молекулы ТГФ (рис. 2г).

Однако с увеличением энергии электронного удара интенсивности сигналов от ионных фрагментов растут в абсолютных величинах. В ряде случаев это приводит к появлению наблюдаемого молекулярного иона и крупным ионным фрагментам молекул, нерегистрируемых при ЭЭ ниже 70 эВ (рис. 3). Как правило, это наблюдается для небольших органических соединений, при электронной бомбардировке которых образуются стабильные молекулярные ионы. В масс-спектрах этилацетата, представленных на рис. 3, при ЭЭ 90 эВ наблюдается молекулярный ион, в то время как при ЭЭ 70 эВ он отсутствует. Следует отметить, что в эталонном спектре этилацетата (БД-МС NIST) присутствует пик с массовым числом 88 и интенсивностью 7%, соответствующий молекулярному иону, что показывает его принципиально возможную устойчивость. Молекулярный ион, однако, может не наблюдаться в конкретных условиях (давление, конструкция ионной оптики), реализованных в используемом в настоящей работе масс-спектрометре. Повышение энергии электронов приводит к увеличению абсолютного количества образующихся ионов, что позволяет детектировать их.

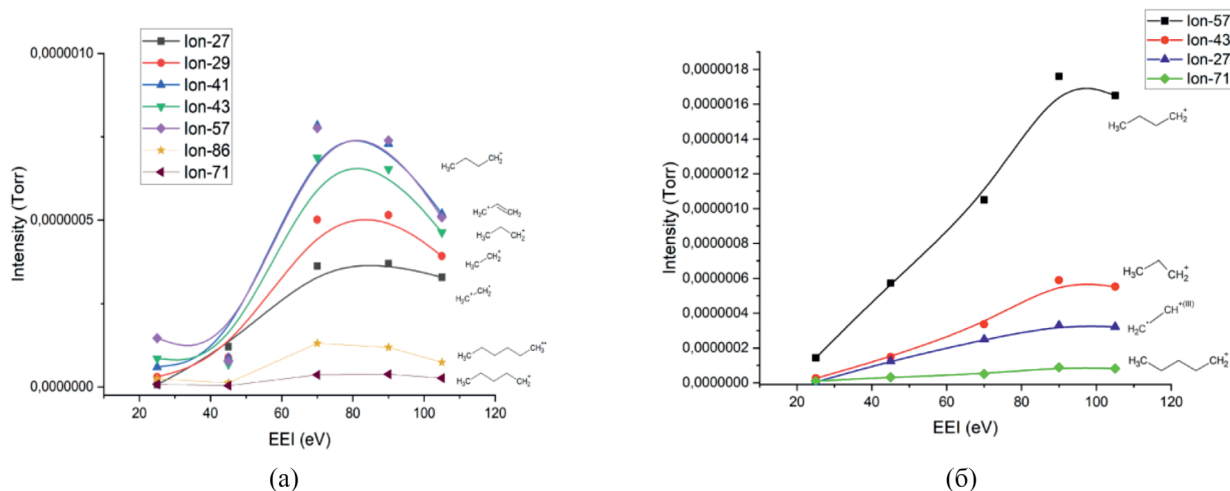


Рис. 4. Кривые эффективности ионизации для основных фрагментарных ионов (а) н-гексана, (б) 2,2,4-триметилпентана.

Таким образом, нельзя однозначно считать, что повышение ЭЭ приводит к исчезновению в масс-спектрах крупных ионных фрагментов молекул или молекулярных ионов, равно как и понижение энергии электронного удара не всегда приводит к регистрации молекулярных ионов и отсутствию мелких ионных фрагментов молекул. Данная неоднозначность продиктована спецификой строения каждой из органических молекул и электронными эффектами в них.

Кривые эффективности ионизации для ряда катион-радикальных ионов н-гексана (рис. 4) имеют характерный вид с максимумом при 70 эВ и дальнейшим монотонным спадом. Похожее зависимость наблюдалась для заряженных частиц, полученных при ионизации большинства соединений, участвовавших в данном исследовании

(циклогексан, изопропанол, нитрометан, бутилацетат, толуол, анилин и пиридин). Для вещества из того же класса органических соединений — 2,2,4-триметилпентана — наблюдается аналогичный вид кривой, хотя максимум приходится на 90 эВ. Такой вид семейства кривых наблюдался для большинства заряженных частиц этилацетата, изопропилацетата и тетрагидрофурана.

Однако кривые эффективности ионизации сложных эфиров, за исключением бутилацетата, имеют монотонно восходящий вид (рис. 5). Вероятно, точка перегиба находится при ЭЭ выше 105 эВ. Это может быть связано с высокой устойчивостью образующихся заряженных частиц в условиях образования довольно небольших фрагментов молекул. Рост массового числа фрагментов от более тяжелых сложных эфиров приводит, по-видимому,

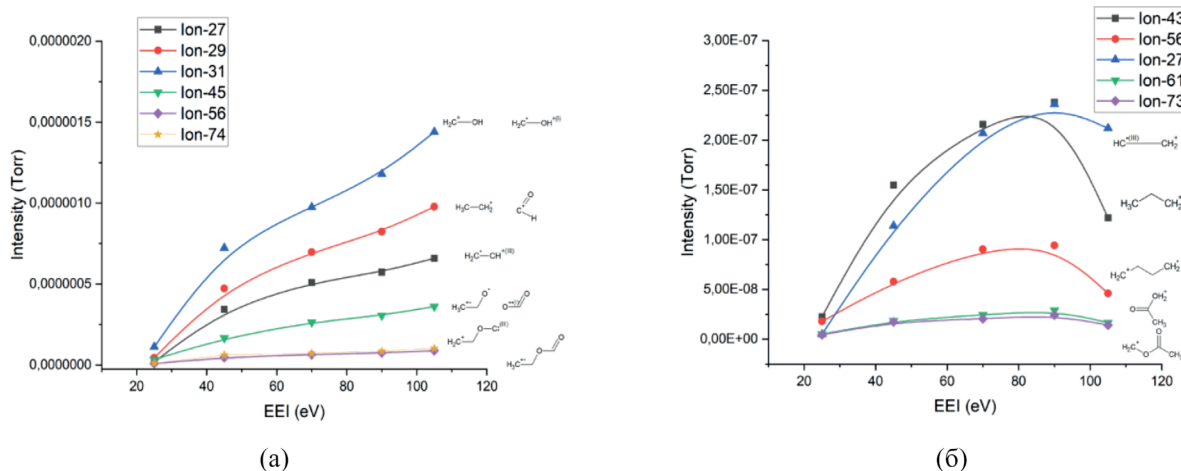
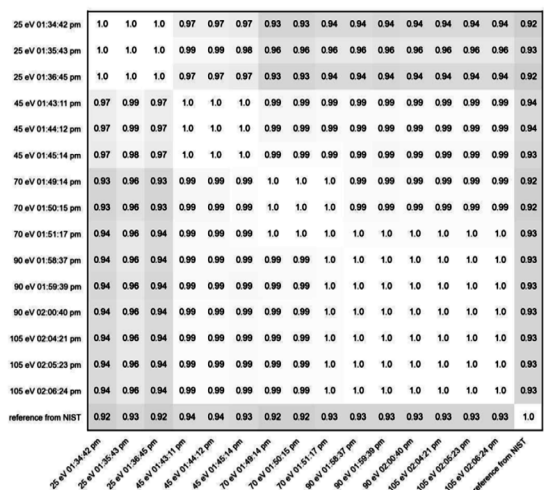
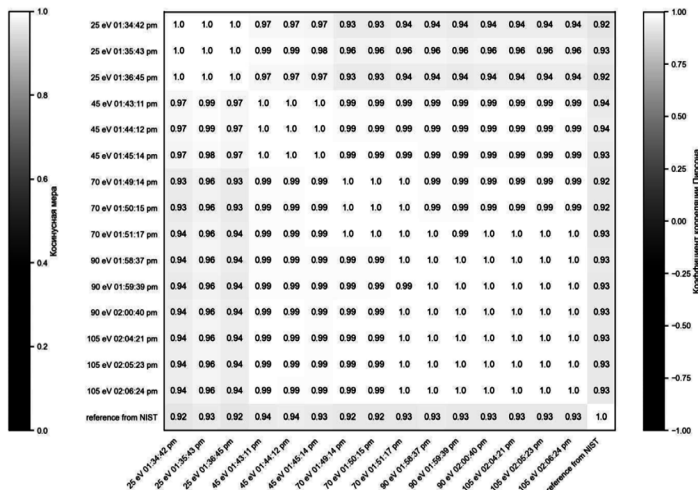


Рис. 5. Кривые эффективности ионизации для основных фрагментарных ионов (а) этилформиата, (б) бутилацетата.



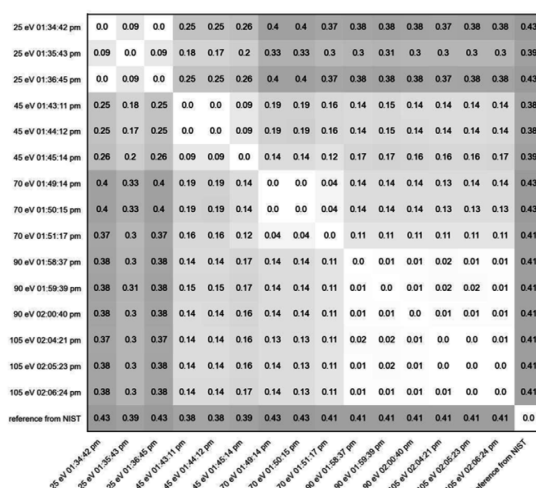
(а)

Рис. 6а. Матрица косинусных мер для 2,2,4-триметилпентана.



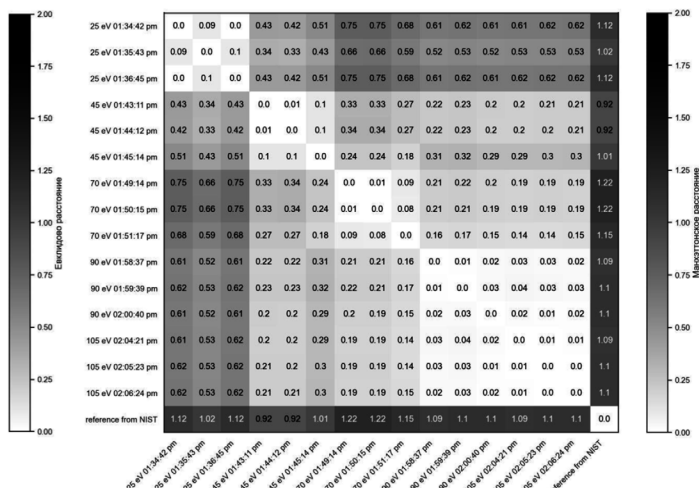
(б)

Рис. 6б. Корреляционная матрица (Пирсона) для 2,2,4-триметилпентана.



(в)

Рис. 6в. Матрица евклидовых расстояний для 2,2,4-триметилпентана.



(г)

Рис. 6г. Матрица манхэттенских расстояний для 2,2,4-триметилпентана.

к снижению эффективности ионизации при увеличении ЭЭ, в результате чего кривая приобретает классический вид (рис. 5).

На рис. 6 приведены матрицы сходства спектров 2,2,4-триметилпентана, полученные на энергиях 25, 45, 70, 90 и 105 эВ, а также со спектром этого соединения из базы данных NIST. Эти матрицы позволяют сделать вывод о хорошем сходстве экспериментальных спектров 2,2,4-триметилпентана с эталонным спектром, взятым из БД NIST. Спектры на одинаковых уровнях энергий совпадают с хорошей точностью, о чем свидетельствуют значения косинусных мер и коэффициента корреляции Пирсона близкие

к 1 и сравнительно небольшие значения расстояний, из чего следует сходимость эксперимента. При сравнении спектров, полученных при разных ЭЭУ, метрики показывают меньшее сходство.

Поскольку косинусная мера обладает рядом существенных преимуществ, по сравнению с другими метриками, далее будем использовать только ее для определения сходства спектров. Она нормирована (в отличие от коэффициента корреляции Пирсона) и обладает информативностью величины (в отличие от евклидового и манхэттенского расстояний, где смыслом обладают лишь отношения расстояний друг к другу).

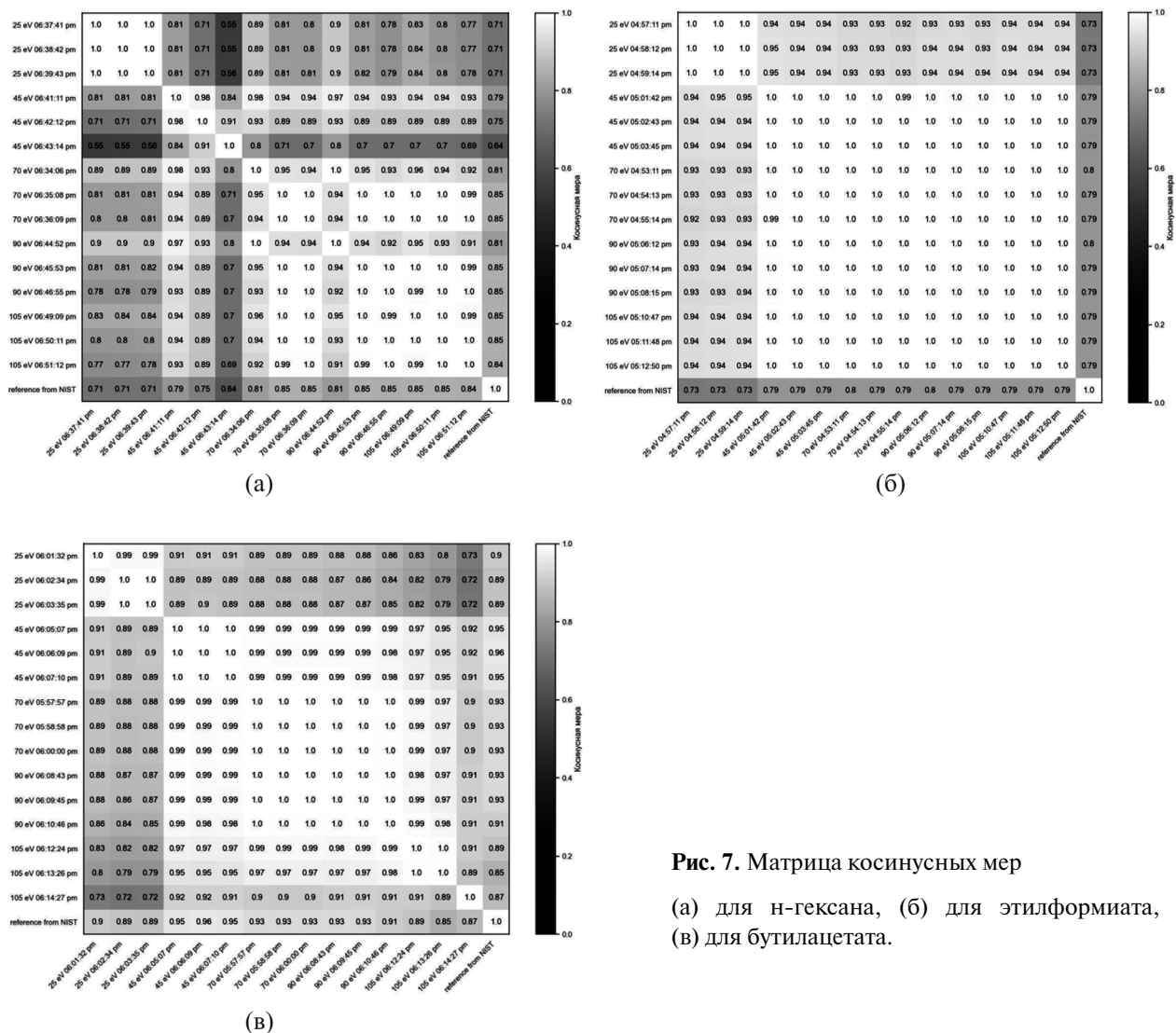


Рис. 7. Матрица косинусных мер

(а) для н-гексана, (б) для этилформиата, (в) для бутилацетата.

Матрица косинусных мер для серии масс-спектров н-гексана показывает наибольшее сходство с базой данных NIST при 70, 90 и 105 эВ (рис. 7а). При сравнении спектральных данных между собой на разных энергиях ионизации видно высокое сходство спектров, полученных при 70 эВ, со спектрами, полученными при других энергиях ионизации. В тоже время заметна слабая корреляция между образцами, измеренными при 25 и 45 эВ. Это связано прежде всего с различиями интенсивностей пиков или вовсе их отсутствием для малых осколочных катион-радикальных фрагментов молекулы. Матрицы косинусных мер для масс-спектров этилацетата и бутилацетата (рис. 7б, в) показывают хорошую повторяемость эксперимента, сходимость спектральных данных с БД-МС NIST, а также свидетельствуют о высоком подобии спектров при изменении ЭЭ.

ВЫВОДЫ

Сравнение данных масс-спектрометрических экспериментов с помощью матриц подобию, основанных на метриках сходства, наглядно подтверждают, что понижение ЭЭ приводит к снижению эффективности ионизации и, как следствие, уменьшению интенсивности сигналов от ионных фрагментов в абсолютных величинах. Ионизация молекул электронным ударом с энергией ниже 70 эВ может приводить к менее жесткой ионизации и образованию молекулярных ионов за счет снижения эффективности фрагментации генерируемых ионов. При этом не следует ожидать образования молекулярных ионов, если они не наблюдаются в спектре ЭЭ 70 эВ, поскольку это связано с их низкой стабильностью, а следовательно, понижение ЭЭ не может служить методом получения молеку-

лярных ионов. Малые ионные фрагменты молекулы при ЭЭ с энергией электронов ниже 70 эВ могут отсутствовать в масс-спектре, или их интенсивности окажутся ниже эталонных интенсивностей, представленных в БД.

Напротив, увеличение ЭЭ приводит к повышению эффективности ионизации и росту интенсивностей сигналов от ионных фрагментов в абсолютных величинах. Повышение ЭЭ может приводить к появлению наблюдаемого молекулярного иона и крупным ионным фрагментам молекул, не регистрируемых при более низких ЭЭ, что, однако, связано не с изменением механизма ионизации, а с ростом абсолютных интенсивностей всех сигналов. Повышение ЭЭ приводит к увеличению интенсивностей малых фрагментов молекул, при этом новых ионных форм не образуется, а лишь увеличиваются интенсивности пиков, соответствующих небольшим ионным фрагментам.

Зависимости интенсивности сигнала от энергии электронного удара для многих заряженных частиц проходят через максимум в районе 70–90 эВ, но возможно смещение максимума в область больших ЭЭ. Поиск таких областей является одной из фундаментальных задач масс-спектрометрии и может быть осуществлен путем анализа полученных на спектрометрах с широким диапазоном варьирования ЭЭ-данных с применением матриц подобия. Среди матриц подобия наиболее информативными и удобными в использовании являются матрицы, основанные на косинусных мерах. Таким образом, создание программ по обработке и сравнению масс-спектрометрических данных и приборов с широким диапазоном варьирования ЭЭ также является актуальной задачей для будущих исследований.

ИСТОЧНИК ФИНАНСИРОВАНИЯ

Работа выполнена в рамках госзадания Министерства науки и высшего образования Российской Федерации (соглашение № 075-03-2023-106).

СПИСОК ЛИТЕРАТУРА

1. *McLafferty F.W.* A century of progress in molecular mass spectrometry // *Annual review of analytical chemistry*, 2011. V. 4. P. 1-22.
2. *Вульфсон Н.С., Заикин В.Г., Микая А.И.* Масс-спектрометрия органических соединений // *Н.С. Вульфсон и др. М.: Химия*, 1986. 312 с.
3. *Field F. H., Franklin J. L.* Electron impact phenomena: and the properties of gaseous ions // *Academic Press*, 2013. V. 1. 564 p.
4. *Yang Q., Ji H., Xu Z. et al.* Ultra-fast and accurate electron ionization mass spectrum matching for compound identification with million-scale in-silico library // *Nat Commun*, 2023, V. 14. P. 372. <https://doi.org/10.1038/s41467-023-39279-7>
5. *Moorthy A.S., Kearsley A.J.* (2021). Pattern Similarity Measures Applied to Mass Spectra. In: *Cruz M., Parés C., Quintela P.* (eds.) *Progress in Industrial Mathematics: Success Stories. SEMA SIMAI Springer Series*, V. 5. Springer, Cham., 2020. https://doi.org/10.1007/978-3-030-61844-5_4
6. *Zhvansky, E.S., Pekov, S.I., Sorokin, A.A. et al.* Metrics for evaluating the stability and reproducibility of mass spectra // *Sci Rep*, 2019. V. 9, P. 914. <https://doi.org/10.1038/s41598-018-37560-0>
7. Numpy and Scipy Documentation [Электронный ресурс] / Distance computations (scipy.spatial.distance). URL: <https://docs.scipy.org/doc/scipy/reference/spatial.distance.html>
8. Learn Data Science [Электронный ресурс] / Cosine Similarity — LearnDataSci. URL: <https://www.learndatasci.com/glossary/cosine-similarity/>
9. *Jiawei Han, Micheline Kamber and Jian Pei*, *Data Mining: Concepts and Techniques* (Third Edition), Elsevier, 2012. <https://doi.org/10.1016/C2009-0-61819-5>
10. *Debbie L. Hahs-Vaughn.* Foundational methods: descriptive statistics: bivariate and multivariate data (correlations, associations), Editor(s): *Robert J Tierney, Fazal Rizvi, Kadriye Ercikan*, *International Encyclopedia of Education* (Fourth Edition), Elsevier, 2023. <https://doi.org/10.1016/B978-0-12-818630-5.10084-3>.
11. *M. Greenacre, Correspondence Analysis.* Editor(s): *Penelope Peterson, Eva Baker, Barry McGaw.* *International Encyclopedia of Education* (Third Edition), Elsevier, 2010. <https://doi.org/10.1016/B978-0-08-044894-7.01317-8>
12. NIST [Электронный ресурс] / Manhattan distance. URL: <https://xlinux.nist.gov/dads/HTML/manhattanDistance.html>
13. *Xin-She Yang.* Data mining techniques, Editor(s): *Xin-She Yang*, *Introduction to Algorithms for Data Mining and Machine Learning*, Academic Press, 2019. <https://doi.org/10.1016/B978-0-12-817216-2.00013-2>
14. *Glineur A., Beccaria M., Purcaro G.* Exploring 20 eV electron impact ionization in gas chromatography-tandem mass spectrometry for the determination of estrogenic compounds // *Journal of Chromatography A*. 2021. V. 1652. P. 462359.

15. Metz T. O. et al. High-resolution separations and improved ion production and transmission in metabolomics // *TrAC Trends in Analytical Chemistry*. 2008. V. 27 (3). P. 205-214.
16. Fialkov A. B. et al. Sensitivity and noise in GC–MS: Achieving low limits of detection for difficult analytes // *International journal of mass spectrometry*. 2007. V. 260 (1). P. 31-48.
17. Ключев Н. А., Бродский Е. С. Современные методы масс-спектрометрического анализа органических соединений // *Российский химический журнал*. 2002. Т. 46. №. 4. С. 57-63.
18. Honkanen O. et al. Mass Spectra of Seven Isomeric Hexen-1-ols // *Acta Chemica Scandinavica*. 1963. V. 17. P. 2051-2054.

COMPARISON OF MASS SPECTRA CHARACTERISTICS USING STATISTICAL ANALYSIS METHODS FOR THE CASE OF IONISATION OF ORGANIC MOLECULES BY ELECTRON IMPACT WITH DIFFERENT ELECTRON ENERGIES

S. V. Silkin^a, A. V. Sakharov^a, S. I. Pekov^{b, c}, V. A. Eliferov^a, V. G. Tkachenko^a, D. V. Kolesnik^a,
E. N. Nikolaev^{b, *}, I. A. Popov^{a, c, **}

^a *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*

^b *Skolkovo Institute of Science and Technology,
Skolkovo, Russia*

^c *Siberian State Medical University,
Tomsk, Russia*

*E-mail: e.nikolaev@skoltech.ru

**E-mail: popov.ia@mipt.ru

The sensitivity and accuracy of volatile organic compound (VOC) identification can be enhanced through the manipulation of ionisation energy in electron impact ionisation mass spectrometric gas analysers. This is achieved by modulating the number of ions formed in the ion source. This paper presents a comparison of data obtained by electron impact ionisation at electron energy (EE) values in the range of 25–105 eV for a number of organic substances belonging to different classes of organic compounds. In order to interpret the dynamics of changes in the peak intensities of fragment ions, an analysis was conducted using similarity matrices based on different similarity metrics. This analysis demonstrated the influence of electron energy (EE) on the probability of formation of the main fragment particles of the studied substances, and consequently, on the similarity of the recorded mass spectrum with the reference mass spectrum from the database.

Keywords: mass spectrometry, electron impact, ionisation, fragmentation, molecular ion

REFERENCES

1. McLafferty F.W. A century of progress in molecular mass spectrometry // *Annual review of analytical chemistry*, 2011. V. 4, P. 1-22.
2. Wulfson N. S., Zaikin V.G., Mikaya A.I. Mass-spectrometry of organic compounds // N.S. Wulfson et al. Moscow: Khimiya, 1986. 312 p.
3. Field F. H., Franklin J. L. Electron impact phenomena: and the properties of gaseous ions // *Academic Press*, 2013. V. 1. 564 p.
4. Yang, Q., Ji, H., Xu, Z. et al. Ultra-fast and accurate electron ionization mass spectrum matching for compound identification with million-scale in-silico library // *Nat Commun*, 2023, V. 14. P. 372. <https://doi.org/10.1038/s41467-023-39279-7>
5. Moorthy A.S., Kearsley A.J. (2021). Pattern Similarity Measures Applied to Mass Spectra. In: Cruz M., Parés C., Quintela P. (eds.) *Progress in Industrial Mathematics: Success Stories*. SEMA SIMAI Springer Series, V. 5. Springer, Cham., 2020. https://doi.org/10.1007/978-3-030-61844-5_4
6. Zhvansky E.S., Pekov S.I., Sorokin A.A. et al. Metrics for evaluating the stability and reproducibility of mass spectra // *Sci Rep*, 2019. V. 9, P. 914. <https://doi.org/10.1038/s41598-018-37560-0>
7. Numpy and Scipy Documentation [Электронный ресурс] / Distance computations (scipy.spatial.distance). URL: <https://docs.scipy.org/doc/scipy/reference/spatial.distance.html>

8. Learn Data Science [Электронный ресурс] / Cosine Similarity — LearnDataSci. URL: <https://www.learndatasci.com/glossary/cosine-similarity/>
9. *Jiawei Han, Micheline Kamber and Jian Pei*. Data Mining: Concepts and Techniques (Third Edition), Elsevier, 2012. <https://doi.org/10.1016/C2009-0-61819-5>
10. *Debbie L. Hahs-Vaughn*. Foundational methods: descriptive statistics: bivariate and multivariate data (correlations, associations), Editor(s): *Robert J Tierney, Fazal Rizvi, Kadriye Ercikan*, International Encyclopedia of Education (Fourth Edition), Elsevier, 2023. <https://doi.org/10.1016/B978-0-12-818630-5.10084-3>.
11. *M. Greenacre*. Correspondence Analysis, Editor(s): *Penelope Peterson, Eva Baker, Barry McGaw*, International Encyclopedia of Education (Third Edition), Elsevier, 2010. <https://doi.org/10.1016/B978-0-08-044894-7.01317-8>
12. NIST [Electronic resource] / Manhattan distance. URL: <https://xlinux.nist.gov/dads/HTML/manhattanDistance.html>
13. *Xin-She Yang*. Data mining techniques, Editor(s): *Xin-She Yang*, Introduction to Algorithms for Data Mining and Machine Learning, Academic Press, 2019. <https://doi.org/10.1016/B978-0-12-817216-2.00013-2>
14. *Glineur A., Beccaria M., Purcaro G.* Exploring 20 eV electron impact ionization in gas chromatography-tandem mass spectrometry for the determination of estrogenic compounds // *Journal of Chromatography A*. 2021. V. 1652. P. 462359.
15. *Metz T. O. et al.* High-resolution separations and improved ion production and transmission in metabolomics // *TrAC Trends in Analytical Chemistry*. 2008. V. 27 (3). P. 205-214.
16. *Fialkov A. B. et al.* Sensitivity and noise in GC–MS: Achieving low limits of detection for difficult analytes // *International journal of mass spectrometry*. 2007. V. 260 (1). P. 31-48.
17. *Klyuev N. A., Brodsky E. S.* Modern methods of mass spectrometric analysis of organic compounds // *Russian Chemical Journal*. 2002. V. 46. №. 4. P. 57-63.
18. *Honkanen O. et al.* Mass Spectra of Seven Isomeric Hexen-1-ols // *Acta Chemica Scandinavica*. 1963. V. 17. P. 2051-2054.