

УДК 004.42:543.51

## РАЗРАБОТКА ПРОГРАММНОГО МОДУЛЯ ДЛЯ ИДЕНТИФИКАЦИИ НИЗКОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ ПО ДАННЫМ МАЛДИ МАСС-СПЕКТРОМЕТРИИ

<sup>1</sup>Посылкина А.А., <sup>1</sup>Карякина О.Е., <sup>2</sup>Карякин А.А.

<sup>1</sup>ФГАОУ ВО «Северный (Арктический) федеральный университет имени М.В. Ломоносова», Архангельск, e-mail: alena\_posylkina@inbox.ru, novogil@mail.ru;

<sup>2</sup>ФГБОУ ВО «Северный государственный медицинский университет» Министерства здравоохранения РФ, Архангельск, e-mail: biophysica@gmail.com

В статье рассматриваются вопросы применения библиотек масс-спектральных данных для анализа химических соединений с помощью метода МАЛДИ масс-спектрометрии. Представлены результаты разработки программного модуля, позволяющего выполнять идентификацию низкомолекулярных соединений на основе анализа масс-спектрометрических данных. Программный модуль был создан с использованием среды визуального программирования «Lazarus» и технологией доступа к данным «ActiveX Data Objects». Хранение исходной информации реализовано с применением системы управления базой данных «MicrosoftAccess». Модуль предусматривает возможность автоматической обработки результатов МАЛДИ масс-спектрометрии анализируемого образца путем графического построения и масштабирования спектра. Идентификация низкомолекулярных соединений осуществляется в ходе процесса сопоставления массы пика экспериментального спектра со значением масс химических соединений в базе данных в выделенном пользователем участке спектра с учетом заданной величины погрешности. Для каждого химического соединения, представленного в базе данных, приведены название, молекулярная масса, брутто-формула и поясняющий комментарий, также при необходимости возможно обновление базы химических соединений информацией об интересующих компонентах. Результатом применения программы является повышение эффективности процесса идентификации неизвестного низкомолекулярного химического соединения с детализированным представлением информации о компонентах масс-спектра.

**Ключевые слова:** идентификация, МАЛДИ масс-спектрометрия, низкомолекулярные соединения

## THE DEVELOPMENT OF THE SOFTWARE MODULE FOR THE IDENTIFICATION OF LOW-MOLECULAR COMPOUNDS BY MALDI MASS SPECTROMETRY

<sup>1</sup>Posylkina A.A., <sup>1</sup>Karyakina O.E., <sup>2</sup>Karyakin A.A.

<sup>1</sup>Northern (Arctic) Federal University (NArFU), Arkhangelsk, e-mail: alena\_posylkina@inbox.ru, novogil@mail.ru, biophysica@gmail.com;

<sup>2</sup>Northern State Medical University, Arkhangelsk, e-mail: biophysica@gmail.com

The article discusses the use of mass spectral data libraries for the analysis of chemical compounds using the MALDI mass spectrometry. The results of the development of a software module that allows the identification of low-molecular compounds based on the analysis of MALDI mass spectrometric data are presented. Software module was developed in visual programming environment «Lazarus» with using data access technology «ActiveX Data Objects». Storage of initial information is implemented using the Microsoft Access database management system. Module makes it possible to automate the data processing of MALDI mass spectrometry results by graphical analysis and scaling of mass spectrum. Identification of low-molecular compounds is realized of compare experimental peak spectrum mass with spectrum of chemical compound mass in database with given quantity of data error. For each chemical compound replaced in database is present chemical name, molecular weight, gross-formula and comment. If necessary, it is possible to update the database of chemical compounds with information about the components of interest. The result of the application of the program is to increase the efficiency of the process of identifying an unknown low-molecular chemical compound with a detailed presentation of information about the components of the mass spectrum.

**Keywords:** identification, MALDI mass-spectrometry, low-molecular compounds

В настоящее время МАЛДИ масс-спектрометрия зарекомендовала себя как перспективный метод, используемый в современных лабораторных исследованиях в различных областях науки. В основе действия метода лежит облучение короткими лазерными импульсами исследуемого образца, который является твердым раствором анализируемого соединения в органической матрице [1].

Анализируемый образец наносится на мишень МАЛДИ, которая затем помещает-

ся в масс-спектрометр, и выполняется анализ исследуемого вещества. Молекулами матрицы поглощается испускаемый короткий световой импульс лазера и происходит разрушение ее кристаллической решетки. При этом часть молекул вырывается с поверхности мишени и образует высокотемпературный суперплотный газ (зона шлейфа), где присутствуют молекулы анализируемого вещества, сохранившие исходную структуру. Нейтральные молекулы и частицы плазмы откачиваются насосами, а образо-

вавшиеся ионы вытягиваются и ускоряются высоким потенциалом по направлению к анализатору. Процесс ионизации и регистрации ионов повторяется после того, как детектора достигнет наиболее тяжелый ион за счет испускания нового лазерного импульса. Спектры всех лазерных импульсов суммируются до получения информативных данных о молекулярной массе анализируемого соединения [2, 3].

Метод МАЛДИ характеризуется рядом преимуществ в сравнении с классическими методами лабораторных исследований, к которым относятся:

- существенное сокращение финансовых затрат на проведение лабораторного анализа за счет низкой стоимости расходных материалов;

- возможность проведения экспресс-тестов (пробоподготовка занимает приблизительно 5–10 мин, а процедура регистрации спектра – 1–2 мин);

- возможность проведения анализа большого количества образцов за одно исследование;

- использование специализированного программного обеспечения, автоматизирующего процесс анализа и самостоятельно интерпретирующего получаемые молекулярные спектры [4, 5].

Проведение исследований с использованием метода МАЛДИ получило широкое распространение в ведущих медицинских и научных центрах, как в России, так и за рубежом с большой вариацией аппаратных и программных решений и комплексными базами эталонных данных. Так метод МАЛДИ масс-спектрометрии активно используется как при традиционных исследованиях биополимеров, микроорганизмов и вирусов, так и при разработке новых лекарственных препаратов, в области клинической лабораторной диагностики различных заболеваний [6].

Широкое применение метод МАЛДИ нашел в анализе высокомолекулярных органических структур, однако представляет интерес и в изучении низкомолекулярных соединений массой менее 1000 Да. Существенным ограничением при идентификации низкомолекулярных МАЛДИ масс-спектров является отсутствие узкоспециализированных баз данных.

В международной практике широкое распространение получили библиотеки масс-спектральных данных, содержащие информацию о структуре анализируемых соединений. Среди наиболее известных – библиотека масс-спектральных данных «WILEY», поддерживаемая компанией «VCH Data Group» (Германия) и содержащая

около 450 тыс. эталонных масс-спектров. Следует выделить также библиотеку «NIST», разработанную и поддерживаемую Национальным институтом стандартизации и технологии США, которая включает более 300 тыс. эталонных масс-спектров [7]. Значительный объем данных в вышеупомянутых библиотеках обеспечивается за счет присутствия информации о синтетических химических соединениях, в то время как поиск сведений о низкомолекулярных химических соединениях сопряжен для исследователя со значительными временными затратами и информационными трудностями. Кроме того, первоначально основное направление МАЛДИ масс-спектрометрии было связано с анализом высокомолекулярных органических соединений. В настоящее время актуальным является изучение возможностей данного метода в области низких масс, в связи с чем требуется создание специализированных программных средств для обеспечения эффективности лабораторных исследований.

Цель исследования: разработка программного модуля для повышения эффективности процесса идентификации низкомолекулярных соединений по данным МАЛДИ масс-спектрометрии.

#### Материалы и методы исследования

При разработке программного модуля была использована классическая каскадная модель программирования, включающая пять основных этапов: проектирование, дизайн, кодирование, тестирование и поддержка. На первом этапе была разработана структура программного модуля, определены функциональные требования. Далее был сформирован удобный в применении и интуитивно понятный для пользователя интерфейс. Программный модуль для идентификации низкомолекулярных органических соединений был разработан с использованием среды визуального программирования «Lazarus» и технологией доступа к данным «ActiveX Data Objects» (ADO). Хранение исходной информации реализовано с применением системы управления базами данных «Microsoft Access».

#### Результаты исследования и их обсуждение

На первом этапе исследования был осуществлен поиск прототипов разрабатываемого программного модуля, в качестве которых был выбран интерпретатор масс-спектров NIST (США) и масс-спектральный идентификатор «АИПСИН АнтиНаркотики» (Беларусь).

Интерпретатор масс-спектров NIST представляет собой утилиту, для которой характерно наличие функций построения и обработки спектральной информации, встроенного изотопного калькулятора, используемого для получения информации

о теоретически рассчитанном изотопном распределении конкретного иона [8]. Интерпретатор может быть использован для генерации различных брутто-формул при анализе фрагментов масс-спектра, однако его библиотека данных не содержит сведений о низкомолекулярных органических соединениях.

Масс-спектральный идентификатор «АИПСИН АнтиНаркотики» представляет собой специализированное информационное приложение для поиска химических соединений по масс-спектрам. Основной задачей данного приложения является автоматизация процесса обработки, расшифровки и интерпретации аналитических данных, получаемых с помощью хроматографического оборудования, оснащенного масс-детектором (хромато-масс-спектрометра). Идентификация анализируемых веществ осуществляется путем сравнения экспериментально полученной масс-спектральной информации с эталонными спектрами, содержащимися в библиотеке системы АИПСИН. Основным недостатком рассматриваемого масс-спектрального идентификатора является ориентирование на установление структуры и статуса контроля наркотических, психотропных веществ, их производных, аналогов, метаболитов и прекурсоров [9].

На рис. 1 представлена структура разработанного программного модуля, демонстрирующая основные функциональные блоки.

Работа с разработанным программным модулем для идентификации низкомоле-

кулярных соединений по данным МАЛДИ масс-спектрометрии включает в себя следующие этапы: установка модуля на ПК; ввод данных масс-спектра; автоматическая обработка введенных данных; идентификация масс-спектральных пиков и получение информации об идентифицированном соединении.

Установка модуля на персональный компьютер осуществляется автоматически при переносе пакета программных файлов. Масс-спектральные данные, представленные в формате таблицы числовых значений масс, загружаются пользователем в главном окне программного модуля с расширением «\*.txt».

На следующем этапе выполняется автоматическая обработка введенных МАЛДИ масс-спектральных данных, в результате которой в левой части диалогового окна программного модуля отображаются результаты графического построения исходного спектра анализируемого образца, в правой части – его вариант с масштабированием для обеспечения удобства обработки данных пользователем (рис. 2). Предусмотрена также возможность построения спектра в заданном пользователем диапазоне масс.

Для удобства работы со спектрами в программном модуле предусмотрена функция дополнительного масштабирования графиков с использованием опции «Значение интенсивности в точках», которая позволяет оператору с высокой степенью детализации оценить значение интенсивности сигнала в выделенном участке спектра (рис. 3).



Рис. 1. Структура программного модуля

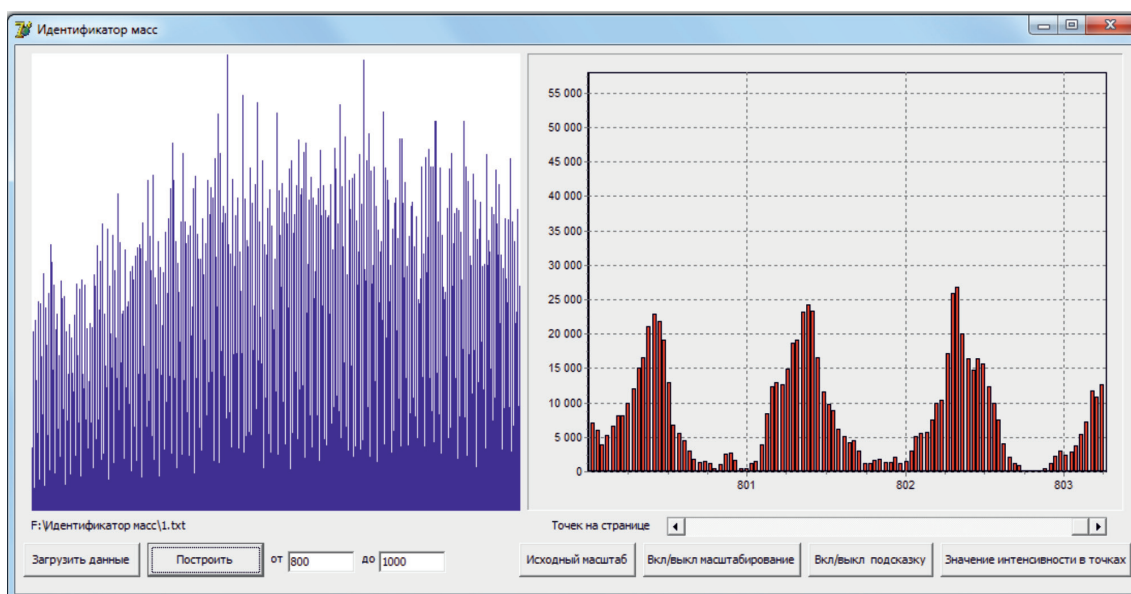


Рис. 2. Построение графического масс-спектра анализируемого соединения в разработанном программном модуле

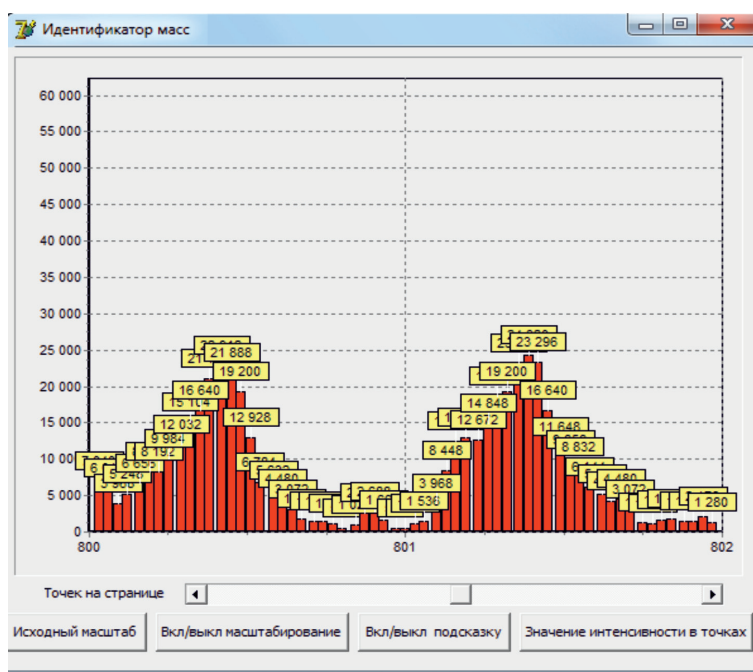


Рис. 3. Диалоговое окно программного модуля для оценки спектра анализируемого образца с использованием опции «Значение интенсивности в точках»

Идентификация спектральных пиков происходит при нажатии на интересующий участок спектра. Программный модуль сравнивает массу пика экспериментального спектра со значением масс химических соединений в базе данных и выдает все совпадающие результаты в соответствии со значением погрешности, введенным оператором (рис. 4). Для

каждого соединения в базе данных представлены название, молекулярная масса, брутто-формула и поясняющий комментарий. Поскольку при масс-спектральном анализе идентифицируемое соединение может содержать ионы H, Li, Na или H<sub>2</sub>O, для более точного определения предусмотрен пересчет массы с учетом массы указанных ионов.



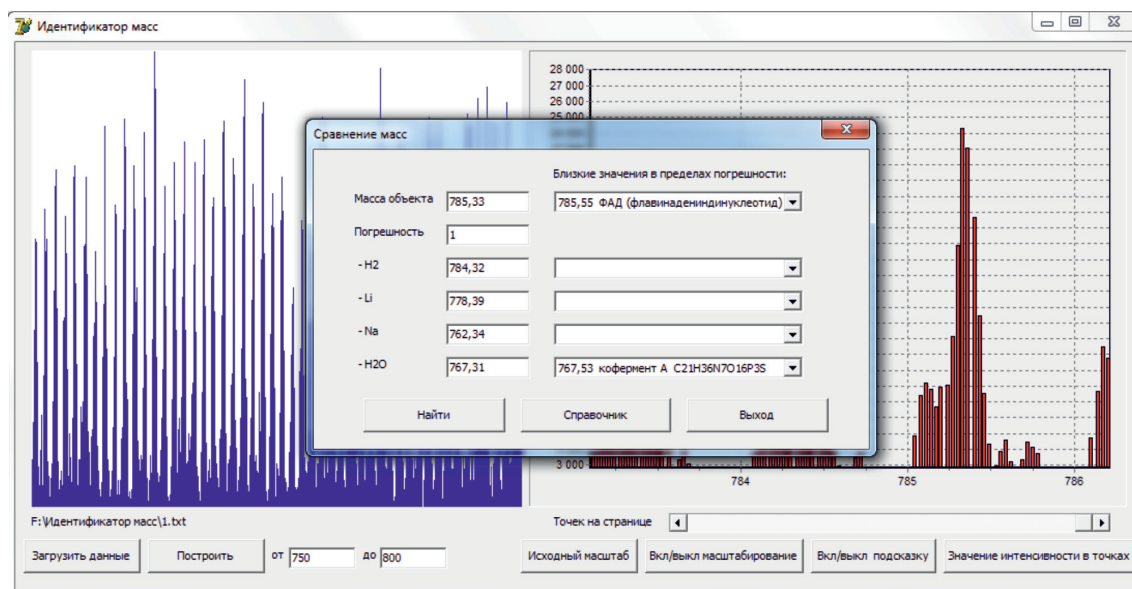


Рис. 4. Диалоговое окно идентификации масс-спектральных пиков низкомолекулярных соединений

В настоящее время в базе данных разработанного программного модуля аккумулированы сведения о более чем 500 химических соединениях. При необходимости оператор может самостоятельно вносить информацию в базу данных программного модуля, воспользовавшись опцией «Перейти к базе данных». В электронной форме «Справочник» необходимо внести данные химического соединения (массу, название, брутто-формулу и комментарий), выбрав опцию «Добавить». Обновление библиотек данных предусмотрено локально, то есть путем внесения новых данных пользователем. При необходимости использования программного модуля пользователями на нескольких рабочих станциях возможна настройка базы данных по локальной сети организации.

### Заключение

Таким образом, разработанный программный модуль позволяет:

- выполнять графическое построение масс-спектра анализируемого образца и его масштабирование;
- выводить на экран детализированную информацию о компонентах масс-спектра;
- осуществлять идентификацию низкомолекулярных соединений с учетом введенной пользователем погрешности и представлением результатов;
- импортировать в базу данных новые сведения о значениях молекулярных масс.

Использование программного модуля позволит оператору повысить эффек-

тивность идентификации неизвестного низкомолекулярного соединения, однако корректность результата идентификации, предложенного программным модулем, должна быть проверена специалистом спектрометрической лаборатории. Кроме того, модуль может быть использован при для идентификации химических соединений вне зависимости от их массы при условии обновления оператором базы данных.

### Список литературы

1. Спицын А.Н., Уткин Д.В., Куклев В.Е. Применение MALDI масс-спектрометрии в диагностике особо опасных инфекционных болезней: современное состояние и перспективы // Проблемы особо опасных инфекций. 2014. № 3. С. 77–82.
2. Лебедев А.Т. Масс-спектрометрия для анализа объектов окружающей среды. М.: Техносфера, 2013. 632 с.
3. Ходорковский М.А. МАЛДИ спектроскопия сложных соединений: методическое пособие. М.: Наука, 2014. 20 с.
4. Баранцевич Е.П., Баранцевич Н.Е. Применение MALDI-TOF масс-спектрометрии в клинической микробиологии // Трансляционная медицина. 2014. № 3. С. 23–28.
5. Беризовская Е.И., Ихалайнен А.А., Антохин А.М. Методы обработки масс-спектрометрических данных при идентификации пептидов и белков // Вестник Московского университета. Серия 2. Химия. 2015. № 5. С. 266–278.
6. Мильман Б.Л., Ильясов Ю.Ю., Луговкина Н.В. «Низкомолекулярный» подход к идентификации микроорганизмов методом масс-спектрометрии МАЛДИ // Масс-спектрометрия. 2016. № 3. С. 157–166.
7. Мильман Б.Л., Журкович И.К. Избранная статистика использования библиотек масс-спектров // Масс-спектрометрия. 2014. № 11 (2). С. 123–125.
8. NIST Chemistry WebBook [Электронный ресурс]. URL: <https://webbook.nist.gov/chemistry/name-ser/> (дата обращения: 20.10.2018).
9. Мисник В.А., Павловец Ю.С., Шевчук Т.А., Юрченко Р.А. Информационно-поисковая система «АИПСИН АнтиНаркотики» и возможности ее использования в правоохранительных органах // Вопросы оборонной техники. 2013. № 11–12. С. 22–27.