
МАСС-СПЕКТРОМЕТРИЯ ДЛЯ БИОТЕХНОЛОГИИ

УДК 621.384.668.8

© Я. И. Лютвинский, Д. М. Петров, А. Н. Веренчиков,
Ю. И. Хасин, М. А. Гаврик

СИСТЕМА РЕГИСТРАЦИИ ДЛЯ ПАРАЛЛЕЛЬНОГО АНАЛИЗА В ВПМС-ТАНДЕМАХ

Описана система регистрации для времяпролетного масс-спектрометра на основе коммерческого быстродействующего АЦП с аппаратным накоплением сигнала AP100 фирмы Acqiris. Проведена оценка скоростных и временных параметров этой системы. Предложено понятие счетного режима использования АЦП при регистрации сигнала и исследованы особенности счетного режима подключения системы регистрации. Проведены испытания системы на действующем времяпролетном масс-спектрометре MX5303 и на макете многоотражательного времяпролетного масс-анализатора. Эксперименты проводились под управлением собственного программного обеспечения, включающего алгоритмы быстрого сжатия (ЖХ-МС)-данных. Продемонстрирована пригодность системы регистрации для времяпролетных тандемов с параллельным двухмерным (МС-МС)-анализом.

ВВЕДЕНИЕ. ПОСТАНОВКА ЗАДАЧИ

В нашей лаборатории разрабатывается тандемный времяпролетный масс-спектрометр (ВПМС), построенный на новых принципах двумерного (МС-МС)-анализа, описанного в [1]. На выходе из первого масс-анализатора ионы фрагментируются в столкновительной ячейке и группируются для анализа во втором масс-анализаторе. Для двумерного анализа время разделения в первом масс-анализаторе должно быть на 3 порядка больше времени анализа фрагментов во втором. Из ионов одного запуска первой ступени ВПМС-тандема получаются группы ионов для анализа во второй ступени ВПМС-тандема.

В качестве первой ступени масс-спектрометра используется многоотражательный времяпролетный анализатор (ВПА) [2], позволяющий расширить время пролета заряженных частиц до 10 миллисекунд, что необходимо для регистрации (МС-МС)-спектров в режиме вложенных времен. Даже при таком сильном расширении времени пролета разрешение обеих ступеней тандемного времяпролетного масс-спектрометра ограничено, что вынуждает прибегать к компромиссам. В качестве компромисса предполагается, что за время пролета ионов в ВПА1 (10 мс) мы успеем провести на ВПА2 до 300 анализов фрагментов фракций ионов, разделенных в ВПА1, что дает время пролета в ВПА2 около 30 мкс. В этом случае разрешение обеих ступеней будет эффективно использоваться, хотя и не достигнет предельных значений.

Заметим, что такой подход к анализу исходной смеси дает нам возможность получить информацию о фрагментации всех составляющих исходной

смеси и впоследствии судить о химической структуре всех компонент исходной смеси. Заметим также, что ни один из существующих приборов не дает подобной возможности.

На выходе второй ступени мы получаем сигнал, несущий непрерывную последовательность фрагментных масс-спектров. Таким образом, на детекторе 2 мы получаем неразделимый пакет сигналов продолжительностью 10 мс. В этом сигнале ионы, прошедшие анализ, представлены импульсами продолжительностью около 2 нс; для уверенной регистрации и оцифровки [3] таких импульсов требуется частота дискретизации не менее 1 ГГц.

Ионы из непрерывно работающего источника ионов накапливаются в ионном накопителе. Нет возможности прервать процесс накопления ионов и избежать переполнения ионного накопителя на время обработки сигналов. Соответственно для эффективной работы прибора система регистрации должна иметь минимальное время простоя.

Тандемные ВПМС часто используются в комплексе с жидкостным хроматографом (ЖХ). Жидкостной хроматограф используется для первоначального разделения сложной аналитической смеси. Именно в такой комбинации тандемные ВПМС используются для решения задач протеомики на экстремально сложных биологических пробах [4].

Для того чтобы эффективно использовать разделение смеси хроматографом, мы должны за время выхода хроматографического пика — 20÷40 с — получить не менее десяти наборов спектров от ВПМС-тандема, т. е. получение спектров должно происходить не реже чем раз в 3 секунды.

Создание системы регистрации для такого прибора налагает ряд требований:

- продолжительность непрерывной регистрации сигнала не менее 8 мс;
- частота дискретизации сигнала не менее 1 ГГц;
- полоса пропускания не менее 500 МГц;
- аппаратная обработка последовательности измерений — в течение 3 с;
- время передачи данных в память компьютера существенно меньше 3 с на пакет спектров.

Такой набор требований отличается от обычного набора требований, предъявляемых к системам скоростной регистрации сигналов. Рассмотрим наиболее типичные применения этих систем.

Цифровые осциллографы. Не предполагают сбора информации обо всем сигнале: при изучении работы электронных схем сигнал носит регулярный характер и достаточно фиксировать только малую часть полного сигнала.

Системы Lidar, Radar. В этих системах требуется фиксировать только незначительную область сигнала, содержащую информацию о волнах, отраженных от целевого объекта.

Специализированные системы регистрации для времяпролетной масс-спектрометрии. Для существующих масс-спектрометров не требуется такое длительное время непрерывной записи спектра. Как правило, время пролета не превышает 500 мкс и регистрация на более длительном промежутке времени не требуется.

Набор наших требований не типичен для большинства промышленных приложений, включая масс-спектрометрию, и выбор подходящих устройств на рынке ограничен, поскольку потребность в устройстве такого вида еще не создана.

После некоторых поисков среди производителей оборудования нами была выбрана фирма Acqiris, предлагающая карты осреднения сигнала, удовлетворяющие большинству указанных требований.

В качестве системы регистрации нами была выбрана плата программируемого цифрового осреднителя (в оригинале — *averager*) сигналов AP100 (www.acqiris.com).

КРАТКОЕ ОПИСАНИЕ AP100 В КАЧЕСТВЕ СИСТЕМЫ РЕГИСТРАЦИИ ВПМС-ТАНДЕМА

Плата швейцарской фирмы Acqiris представляет собой программируемый цифровой осреднитель сигналов, исполненный в виде полноформатной платы PCI, на основе восьмиразрядного АЦП с частотой дискретизации 1 ГГц, снабженный 32 Мбайт памяти для хранения данных.

При осреднении результат каждого последующего измерения складывается с суммой результатов всех предыдущих измерений. Точки измерений синхронизированы с запускающим триггером.

За счет этого в памяти устройства накапливается информация, характеризующая состояние измеряемого сигнала в среднем на протяжении некоторого количества измерений. Возможно, с точки зрения русского языка было бы правильнее называть такое устройство накопителем сигналов.

На хранение суммы по каждой точке измерений отводится 4 байта; таким образом, 32 Мбайт памяти позволяют записывать усредненные спектры, в которых продолжительность регистрации составляет до 8 мс при частоте дискретизации 1 ГГц, что удовлетворяет требованиям к системе регистрации ВПМС-тандела.

AP100 способно передавать накопленные данные в память обрабатывающего компьютера по шине PCI со скоростью до 107 Мбайт/с, что составляет до 35 млн. точек/с. К сожалению, в AP100 не реализован режим работы с двойным буфером, из-за чего карта может либо проводить измерения, либо передавать данные, но не то и другое одновременно, что приводит к простоям прибора на время передачи данных между платой и компьютером. Это является наиболее существенным недостатком при использовании AP100 в качестве системы регистрации времяпролетных масс-спектрометров.

Собственное программное обеспечение, поставляемое с AP100, работает слишком медленно для его непосредственного использования (запись одного спектра около 1 минуты). Для программного доступа AP100 предоставляет развитый и хорошо документированный API, к которому можно обращаться из языков высокого уровня LabView, VisualBasic, Visual C++, Borland CBuilder. Наибольшую производительность показывают программы, написанные на диалектах языка C++ — Visual C++, Borland CBuilder. Наибольшую устойчивость — на Visual C++ 6.0. Для подробного практического исследования системы регистрации нами была реализована программа TOF+, снимки экранов которой использованы в этой работе.

Далее мы более подробно рассмотрим наиболее важные свойства системы регистрации, методы работы с AP100, как системой регистрации ВПМС, некоторые алгоритмы, необходимые для использования AP100 в качестве системы регистрации ВПМС.

ВРЕМЕННЫЕ И СКОРОСТНЫЕ ХАРАКТЕРИСТИКИ УСТРОЙСТВА

Временные характеристики устройства были исследованы в схеме минимизирующей простоту AP100. Впоследствии такая же схема применялась при подключении AP100 к ВПМС MX5303 (рис. 1). Задачей исследования было точно установить временные задержки при использовании системы регистрации AP100.

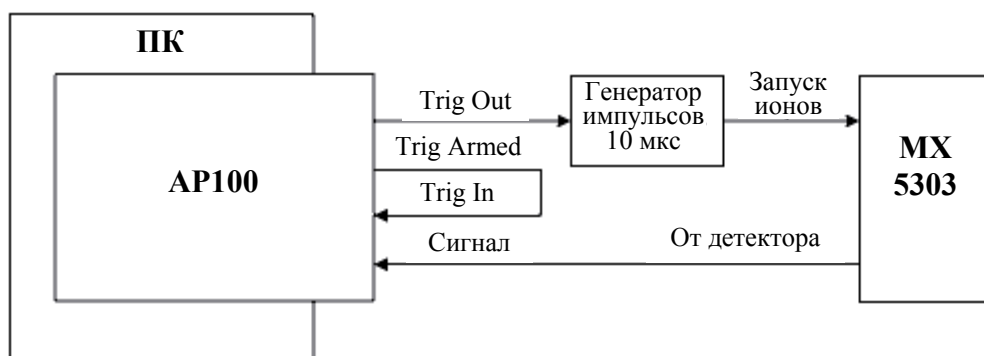


Рис. 1. Схема подключения системы регистрации на основе AP100 к ВПМС MX5303

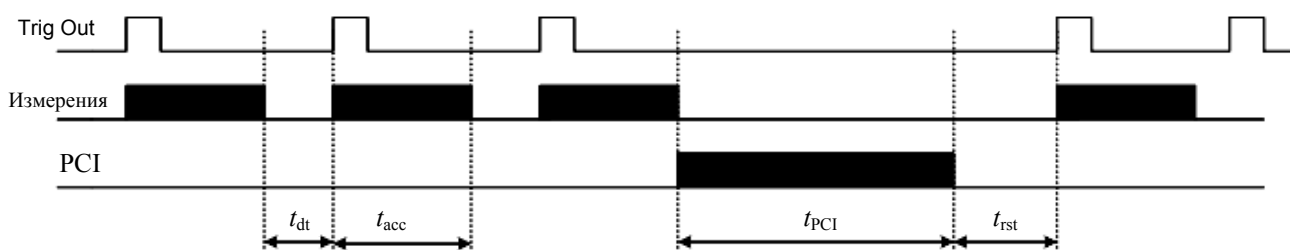


Рис. 2. Временная диаграмма работы AP100

Временные соотношения, наблюдающиеся при работе с AP100 показаны на рис. 2:

t_{acc} — время измерения, это время точно определено числом точек измерения и нам известно;

t_{dt} — пауза между измерениями, время, необходимое плате AP100, чтобы запустить цикл измерения заново, — как оказалось, $t_{dt} \sim 1$ мкс;

t_{rst} — время запуска нового измерения, составляет ~ 40 мкс;

t_{PCI} — время передачи информации по шине PCI, соответствует (и даже несколько превышает — 107 Мбайт/с) заявленной в документации скорости передачи по шине PCI (100 Мбайт/с).

При проведении компьютерного эксперимента мы можем измерить время $(t_{acc} + t_{dt}) \times n + t_{rst}$, где n — число просуммированных измерений, выполненных за весь сеанс измерений. Кроме этого, измерению доступно t_{PCI} . Эту информацию можно получить от встроенных в микропроцессор ПК высокоточных таймеров. Получение этого времени позволяет вычислить остальные временные характеристики AP100.

Как уже говорилось, во время t_{PCI} регистрации сигнала не происходит. Оценим потери времени работы системы регистрации по этой причине.

Понятно, что чем больше объем информации, передаваемой за единицу времени, тем больше потери времени простоя. Уменьшить процент времени простоя можно за счет большего времени накопления спектра или в экстремальных случаях при необходимости фиксировать очень большое число спектров за счет понижения частоты дискретизации.

Для проведения скоростных измерений с минимальными потерями нами была разработана программа AQServer, позволяющая сохранить до 800 спектров в 1 секунду. При помощи этой программы мы измерили число спектров в 1 секунду, которое можно получить на AP100 для разного накопления и для разной частоты дискретизации. Ось абсцисс на диаграммах рис. 3 показывает число циклов измерения для получения одного спектра. Спектры записаны для времени пролета ионов 100 мкс. В зависимости от частоты дискретизации длина спектров менялась: от 100 000 точек для частоты дискретизации 1 ГГц до 10 000 точек для частоты дискретизации в 100 МГц. Результат измерений, показанный на диаграмме рис. 3, а, позволяет вычислить эффективность использования входного потока ионов в разных режимах регистрации спектров (рис. 3, б).

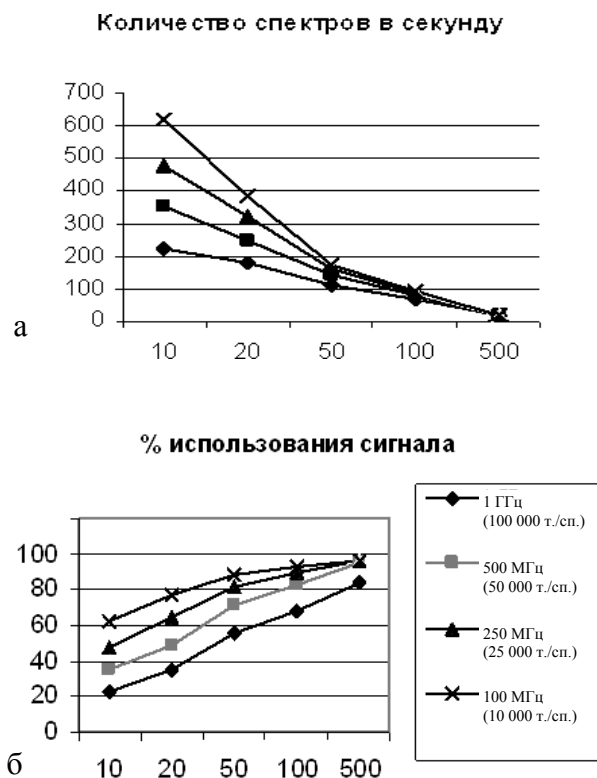


Рис. 3. Число спектров в секунду (а) и процент использования сигнала (б) в зависимости от числа запусков ВПА, накопленных в одном спектре, и частоты дискретизации сигнала

ПРЕДВАРИТЕЛЬНЫЕ ИСПЫТАНИЯ AP100 КАК СИСТЕМЫ РЕГИСТРАЦИИ ВПМС

Некоторые свойства AP100 как системы регистрации ВПМС можно установить только в процессе практических испытаний AP100 в комплексе с действующим ВПМС, поскольку свойства сигнала ВПМС плохо воспроизводимы любыми другими испытательными установками.

Наши испытания проводились на времяпролетном масс-спектрометре MX5303 с ортогональным вводом и ионизацией электрораспылением. Этот ВПМС — разработка нашей лаборатории и используется для актуальных биохимических исследований, результат одного из которых описан в статье [5]. Интенсивность и насыщенность сигнала детектора этого прибора при работе с биохимическими пробами характерны для времяпролетных приборов умеренного разрешения и интенсивности. В частности, прибор, находившийся в нашем распоряжении, обеспечивал разрешающую способ-

ность 8000–10 000 и максимальную интенсивность потока ионов около 0.4 иона на пик на выстрел. При подключении к этому прибору использовалась схема, изображенная на рис. 1.

Процесс измерений начинается по готовности триггера. Сигнал готовности триггера поступает на входной триггер AP100. Запуск генератора импульсов MX5303 происходит по синхронизованному с внутренними часами AP100 сигналу Trig Out. Такая схема минимизирует время простоя AP100 и рассинхронизацию между ВПМС и системой регистрации.

Далее приводятся некоторые результаты исследований, проведенных на комплексе из ВПМС MX5303 и системы регистрации AP100.

Способ регистрации сигнала ВПМС

Первой проблемой, с которой приходится сталкиваться при регистрации сигнала детектора, является шум канала регистрации, возникающий в процессе доставки сигнала от детектора к АЦП и его последующей оцифровки. Такой шум, как показало исследование AP100, занимает не менее 1/2000 от полной шкалы измерений. Так, например, на рис. 4 показан сигнал с AP100 при осреднении 1000. Полный диапазон измерений такого сигнала составляет 256 000. Как видно на этом рисунке, ширина шумовой дорожки для этого сигнала — около 150, т. е. 1/1700 от диапазона измерений.

Для того чтобы избавиться от этого вида шума, мы выносим этот шум за пределы измерений. Для этого нуль шкалы измерений смещается к середине шумовой дорожки и устанавливается порог чувствительности карты в значение, превышающее уровень шума канала регистрации. В этом случае карта будет регистрировать только статистические выбросы шума, превышающие уровень чувствительности. При таком способе регистрации сигнала важно выбрать правильный уровень усиления сигнала детектора. Если усиление сигнала

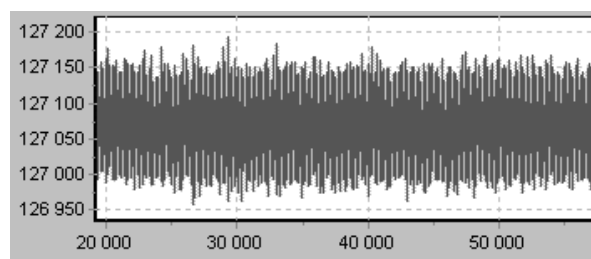


Рис. 4. Шум канала регистрации

детектора будет слишком слабым и сигнал детектора на входе АЦП будет составлять 1–2 бита, то такой сигнал будет потерян за пределами уровня чувствительности. В принципе при малом осред-

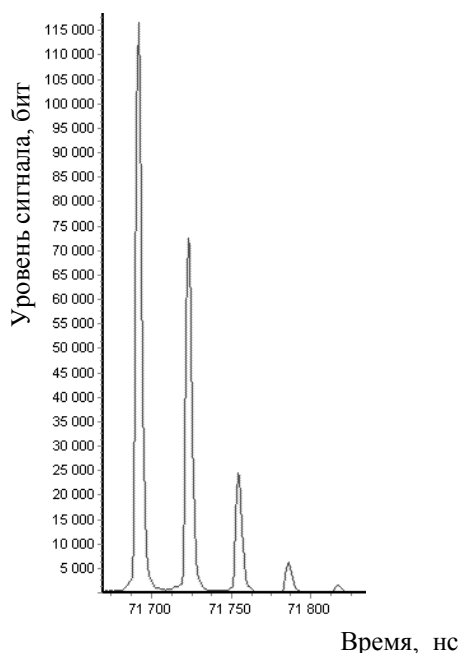


Рис. 5. Спектр *граммицина-С*, записанный на приборе MX5303 с помощью системы регистрации на основе AP100

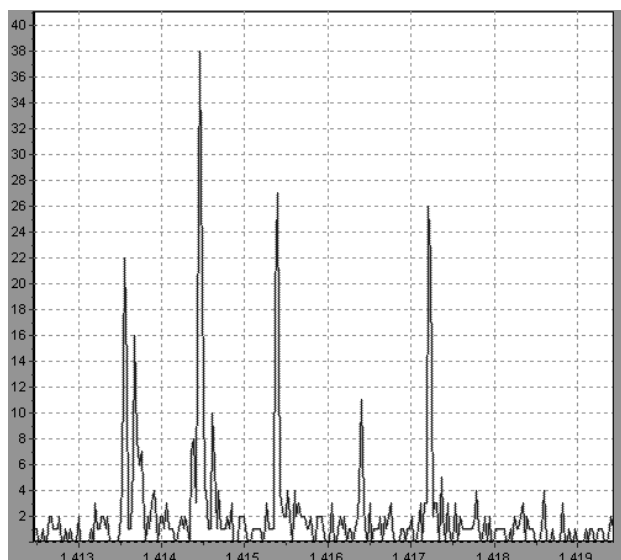


Рис. 6. Слабый аналитический сигнал, зафиксированный AP100 при включении в счетном режиме

нении мы могли бы улавливать и такие слабые сигналы. Однако при значительном осреднении такие сигналы будут потеряны в шуме канала регистрации. Поэтому выбирается такой уровень предварительного усиления сигнала, чтобы срабатывание детектора было представлено всплеском на уровне около 16 бит, или 30 мВ при полной шкале AP100 в 500 мВ; уровень чувствительности AP100 при этом устанавливается в 6–10 бит. В этом случае все сигналы детектора будут уверенно зафиксированы в результатах измерений, а шум канала измерений будет практически исключен.

Такой режим измерений при помощи АЦП мы называем *счетным* режимом измерений (в противоположность традиционному *аналоговому* режиму, подразумевающему полную фиксацию шумовой дорожки).

Конечно, таким способом уменьшается динамический диапазон имеющегося АЦП с 256 до 16, однако одновременная регистрация в течение одной–двух наносекунд 16 ионов — это крайне редкое событие. Современные источники, как правило, не в состоянии обеспечить подобную плотность потока ионов даже для самых сильных компонент спектра [6]. Наш подход позволяет фиксировать как основные компоненты исходной смеси (рис. 5), так и компоненты, представленные в сигнале буквально несколькими ионами. Спектр на рис. 6 составлен из 11–12 зарегистрированных ионов, тем не менее этот спектр остается информативным.

Шумы и шумоподобные сигналы

После перевода AP100 в счетный режим влияние шума канала регистрации сводится к минимуму. Шум канала наблюдается только в виде статистических выбросов, превышающих порог регистрации. Однако в сигнале остаются другие виды шумов.

Ионный шум. Несистематические срабатывания детектора из-за темнового тока и не отличимые от них статистические выбросы шума канала регистрации образуют так называемый ионный шум.

Кроме того, в сигнале масс-спектрометра могут присутствовать слабые аналитические сигналы, не отличимые по уровню от ионного шума. При работе с последовательностями спектров ЖХ-МС из такого разреженного сигнала можно извлечь информацию о минорных компонентах аналитической смеси. Работа по автоматическому вычленению таких слабых сигналов сейчас ведется.

Химический шум. В хорошо накопленных спектрах на каждой целой массе наблюдается сигнал от неких слабых примесей из анализируемого раствора или из атмосферы (рис. 7). Нехарактерная форма пиков с несколькими апексами позволяет предположить, что каждый из этих пиков

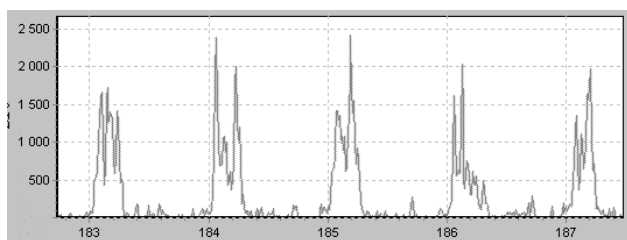


Рис. 7. Химический шум

образован несколькими веществами с одной целочисленной молекулярной массой, но разными брутто-формулами. Такие сигналы препятствуют распознаванию минорных составляющих аналитической смеси, особенно в диапазоне масс до 250 а. е. м. Для отбора истинных сигналов на малых массах можно использовать фильтрацию на основе разрешения пика. Пики, образованные химическим шумом, существенно шире истинных сигналов, ширина пиков которых определяется разрешением прибора. Тем не менее для счетного режима системы регистрации наличие естественного химического шума является одним из основных факторов, ограничивающих чувствительность прибора.

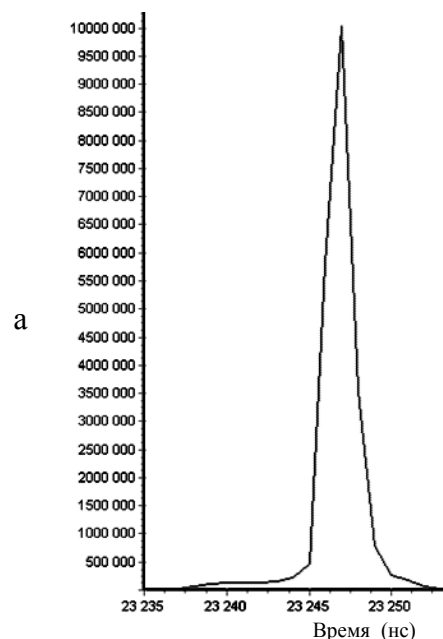
Динамический диапазон и точность определения масс

Важнейшими аналитическими характеристиками масс-спектрометра являются динамический диапазон и точность определения масс. Для определения предельного динамического диапазона нами был проведен эксперимент по длительному (в течение 10 минут) накоплению стабильного сигнала коронного разряда на ВПМС MX5303. Этот эксперимент уверенно показал динамический диапазон сигнала, равный 100 000. Фрагменты спектра с таким диапазоном показаны на рис. 8.

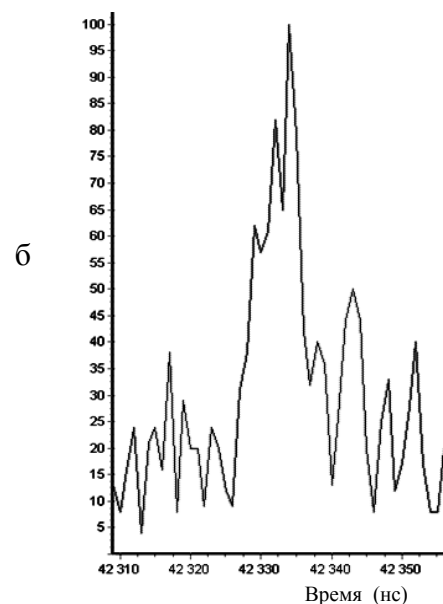
Дальнейшему увеличению динамического диапазона препятствует накопление ионного шума, который образует подложку под слабыми сигналами, как это видно на рис. 8, б.

Для коротких времен накопления сигнала наблюдаемый динамический диапазон составляет около 3000 для спектра, накопленного в течение 1 секунды, что соответствует возможностям данного прибора. Используемый у нас в настоящее время источник не позволяет развивать плотность потока ионов более 0.4 за выстрел на пик.

Для определения возможностей по точному определению масс использовались спектры *полиэтилэмина*, полученные при помощи системы



а



б

Рис. 8. Максимальный (а) и минимальный (б) пики в спектре коронного разряда (масштаб ординат на а и б одинаков)

регистрации на основе AP100 при подключении к MX5303. Такие спектры содержат множество известных пиков в широком диапазоне масс, что позволяет набрать статистику точности определения масс в пределах одного спектра. Низкая интенсивность сигнала требовала значительного времени накопления спектра: нами использова-

лись спектры, накопленные в промежутке от 1 до 140 с. Разрешение прибора составляло около 10 000.

Для калибровки использовались только два известных пика с массами 276.2876 и 663.6674. На рис. 9 представлен спектр с временем накопления 140 с. Этот спектр показывает, что система из ВПМС и AP100 потенциально способна работать с точностью определения масс около 1 ppm.

Измерения, проведенные на спектрах с меньшим временем накопления, показывают, что увеличение времени накопления спектра повышает точность определения масс. Это происходит за счет большей статистической достоверности пиков в спектре. Плотность потока ионов в проведенном эксперименте составляет 0.05 иона на пик за выстрел.

Длительность времени накопления спектра критично зависит от эффективности источника ионов. В экспериментах измерения производились на потоке ионов намного меньше 1 иона на пик за выстрел. Динамический диапазон AP100 при включении в счетном режиме — ~ 16 ионов на пик за выстрел. Если использовать высокоэффективные источник и детектор ионов, позволяющие получать и регистрировать плотности потоков ионов до 16 ионов на пик за выстрел, можно получить: точность определения масс — около 1 ppm для основных компонент спектра в режиме ЖХ-МС при записи до 10 спектров в 1 секунду и динамический диапазон в 100 000 при времени накопления около 20 с. Заметим, что уменьшение времени накопления потенциально способно уменьшить влияние ионного шума, ограничивающего динамический

диапазон прибора при длительном накоплении сигнала.

AP100 предоставляет метод, который позволяет увеличить динамический диапазон еще примерно в 8 раз при увеличении времени накопления спектра вдвое. Это может быть достигнуто за счет пропорционального суммирования спектров, накопленных на шкалах 500 и 50 мВ. В этом случае сигналы, превышающие шкалу 50 мВ, могут быть оценены на шкале 500 мВ. Такой режим реализован в нашем программном обеспечении, однако практические эксперименты в таком режиме не проводились из-за отсутствия источника сигнала с таким динамическим диапазоном.

Добавочное время простоя AP100 для переключения диапазонов усиления не превышает 2 мс, что составляет малую часть типичного времени накопления спектра в режиме ЖХ-МС.

ПОДКЛЮЧЕНИЕ AP100 К МАКЕТУ МНОГООТРАЖАТЕЛЬНОГО ВПМС

Сложность исследования AP100 заключается в том, что сложно найти устройство, при помощи которого можно проверить работу AP100 при большом времени накопления и на сигналах с большим динамическим диапазоном. Подобная возможность появилась с разработкой в нашей лаборатории макета планарного многоотражательного ВПМС, описанного в работе [6].

Этот макет является прототипом ВПА1 для ВПМС-тандема. На этом макете время пролета ионов через анализатор близко к времени регистрации последовательности (МС-МС)-спектров в тандемном ВПМС. Испытания системы регистрации на этом макете позволили судить об устойчивости работы AP100 при длительном времени регистрации сигнала высокой интенсивности. Потенциально такая высокая интенсивность сигнала может быть достигнута и при работе с биохимическими пробами за счет введения более эффективных источников ионов.

При подключении AP100 к макету источником синхронизации для запуска ВПА и начала измерения AP100 служил внешний генератор импульсов. Время простоя платы определяется частотой этого генератора. Такая схема подключения увеличивает джиттер до половины периода частоты дискретизации.

Был получен набор спектров сигнала цезиевой пушки, описанной в статье [7], с временем пролета ионов около 10 мс. Информативные участки одного из таких спектров показаны на рис. 10. Набор сигналов в спектре ограничен изотопами калия, рубидия и цезия в материале цезиевой пушки и динамический диапазон соответствует их относительной концентрации. Соотношение инте-

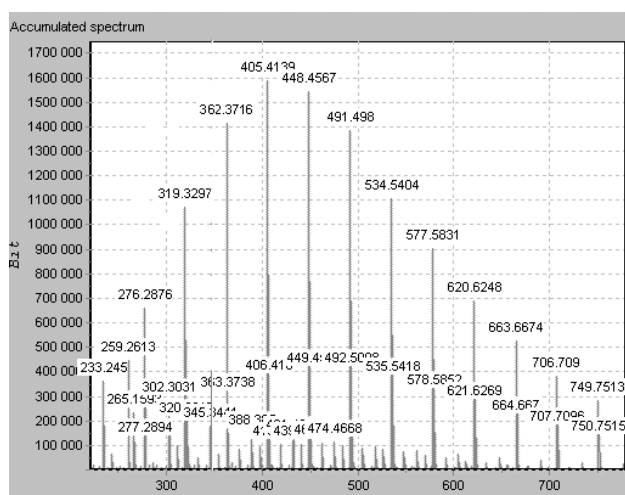


Рис. 9. Спектр полиэтилмина, накопленный в течение 140 секунд на MX5303 с системой регистрации AP100

гальной интенсивности зарегистрированных сигналов Cs^{133} и Ca^{41} дает нам динамический диапазон в этом спектре $\sim 250\,000$. Этот спектр накоплен всего за $\sim 100\,000$ запусков ВПА. Это было достигнуто за счет введения эффективного источника ионов в виде цезиевой пушки и отсутствия химического шума в сигнале.

Нам не удалось обнаружить какие-либо дополнительные недостатки AP100, связанные с большим временем пролета. Правда, следует отметить, что разрешение макета около 4000 дает нам продолжительность пиков порядка 100 нс и максимальную интенсивность пиков около 20 ионов на пик за выстрел. При такой ширине и интенсивности пиков малые отклонения в работе AP100 могли остаться незамеченными.

СОХРАНЕНИЕ ДАННЫХ

Известно, что в процессе экспериментов ЖХ-МС-МС, проводимых на современных приборах, обеспечивающих снятие 1–3 (МС-МС)-спектров в 1 секунду, поток данных для сохранения стано-

вится очень велик — порядка 1 Мбайт в 1 секунду. Для разрабатываемого в нашей лаборатории ВПМС-тандема проблема сохранения очень большого объема данных еще более обостряется, т. к. прибор в состоянии генерировать заметно больший объем данных, чем существующие системы.

Оценим поток данных, который может генерировать тандемный ВПМС. Оцифровка сигнала набора (МС-МС)-спектров от детектора происходит в течение 8 мс с предельной частотой дискретизации 1 ГГц — это дает 8 000 000 точек. Последовательность таких сигналов суммируется в аппаратном накопителе AP100.

При совмещении тандемного ВПМС с жидкостным хроматографом нужно получить несколько наборов спектров за время выхода хроматографического пика. Для этого мы должны снимать накопленный спектр не реже чем 1 раз в 3 с. То есть каждые три секунды мы получаем 8 000 000 точек, по 4 байта на каждую точку — 32 Мбайт. Таким образом, система регистрации создает поток данных около 10 Мбайт в секунду. В течение одного (ЖХ-МС-МС)-эксперимента продолжительностью 1 ч мы получим 36 Гбайт несжатых данных.

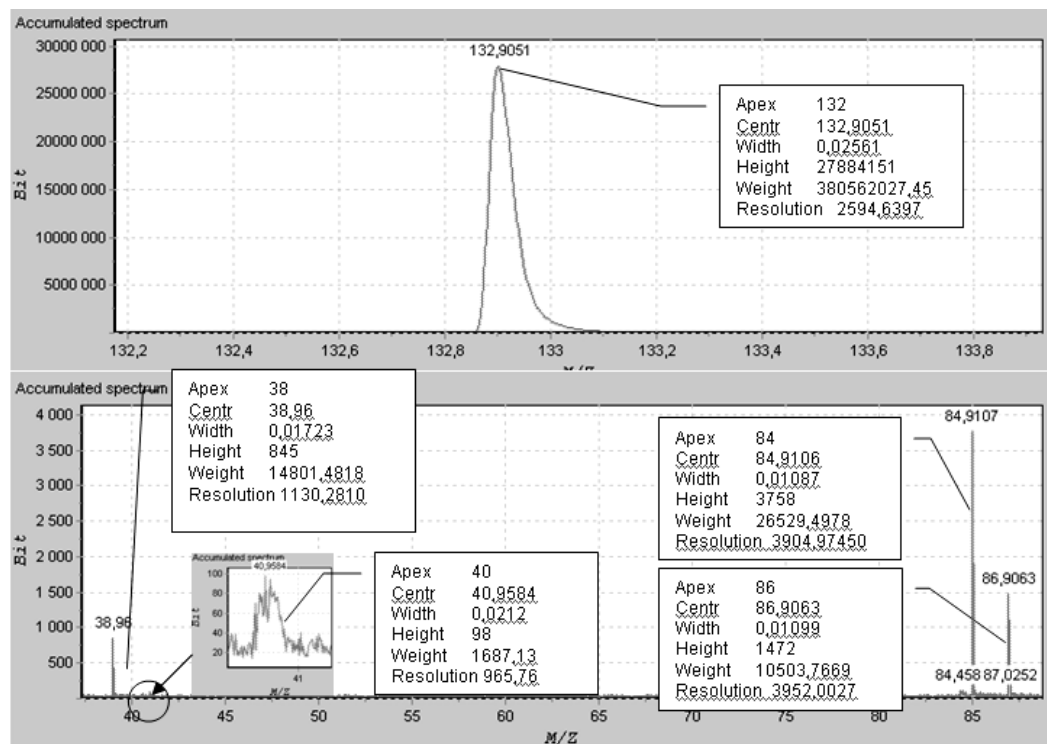


Рис. 10. Спектр, полученный при помощи AP100 при подключении к макету многоотражательного ВПМС

Такой объем данных слишком велик для хранения или передачи, при последующей обработке потребуется слишком много времени для чтения такого объема информации. Для сохранения дискового пространства, для ускорения процесса последующего анализа информации мы разработали процедуру компрессии информации.

Компрессия информации происходит в два этапа:

- 1) фильтрация по порогу чувствительности и удаление нулей;
- 2) сжатие с использованием кодов Хаффмана.

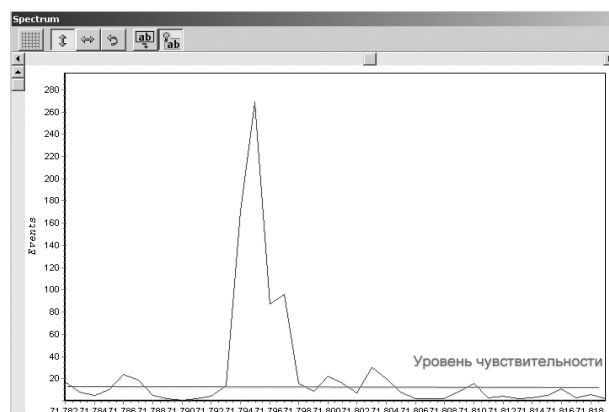
Принцип фильтрации очень прост. Нужно заменить в измеренной последовательности все значения, меньшие определенного, нулями и затем заменить последовательности нулей числом нулей и маркером замены. В качестве маркера можно выбрать как раз число ноль, после замены в последовательности таких самостоятельных значений не остается. Такая замена показана на рис. 11, б, точками на этом рисунке отмечены числа, которые только и нужно сохранить.

Некоторую сложность может представлять выбор правильного уровня чувствительности: в идеале такой уровень должен быть выше уровня шума канала, но ниже уровня ионного шума. Тогда никакая полезная часть сигнала не будет потеряна.

Код Хаффмана — это классический код для сжатия информации, подробно описанный, например, в [9]. Его использование для сигнала ВПМС основано на том, что в этом сигнале значения неравномерно распределены по числу эффективно используемых бит. На рис. 12 показано распределение для сигнала, уже пропущенного через процедуру подавления нулей и фильтрацию по порогу чувствительности. Для необработанного сигнала неравномерность распределения еще более выражена — в нем сохранено большое количество малых чисел. Возникновение такого распределения естественно: в сигнале ВПМС, как правило, фиксируется только несколько основных компонент и намного больше менее интенсивных сигналов. Если же не проводить фильтрацию, большую часть данных будут составлять малые значения шума.

При таком неравномерном распределении можно оптимизировать хранение информации, организовав данные в виде битового массива. Для этого представим каждое число из спектра как пару <длина—данные>. В качестве данных будем хранить только используемые биты значения, кроме первого (первый используемый бит всегда равен единице). Префикс в виде числа бит кодируется по Хаффману и служит маркером начала следующего числа.

Эффективность такой обработки зависит от исходной информационной наполненности сигнала.



а



б

Рис. 11. Фильтрация по уровню чувствительности и подавление нулей

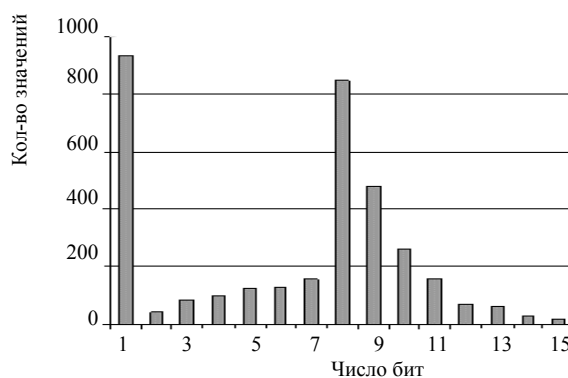


Рис. 12. Статистика распределения чисел в спектре по количеству эффективно используемых бит

Для спектров с малым временем накопления размер файла после сжатия явно соотносится с общим ионным током — 0,8–1,2 байта на ион. Это позволяет увязать максимальный размер файлов с общим ионным током; например, для MX5303 поток данных после компрессии не превышает 20 Кбайт/с и слабо зависит от количества записанных за 1 секунду отдельных спектров.

Для спектров с большим временем, в которых представлены статистически хорошо накопленные пики, этот показатель, разумеется, падает. Например, спектр *полиэтиламина*, показанный рис. 9, с временем накопления 1 с сжимается до размера 10 Кбайт. Исходный размер данных при этом — 480 Кбайт. Заметим, что такой насыщенный многокомпонентный спектр, как показанный на рис. 9, — редкость в (ЖХ-МС)-экспериментах. Обычно спектры ЖХ-МС намного беднее и содержат 2–3 основных компонента и 10–20 минорных. Такие спектры при сжатии редко превышают предел в 5 Кбайт при сжатии спектра, записанного в течение 1 с.

Оценим, каков может быть поток информации после сжатия от тандемного ВПМС. Сейчас можно предположить, что при работе со сложными биохимическими пробами фрагментные спектры будут иметь примерно ту же наполненность, что и спектры ЖХ-МС. Типичный коэффициент сжатия для таких спектров — 50–100. В таком случае исходный информационный поток в 10 Мбайт/с может быть сжат до 100–200 Кбайт/с. Для эксперимента ЖХ-МС-МС продолжительностью 1 ч

общий объем данных составит 300–600 Мбайт, т. е. объем одного CD-диска.

Время работы алгоритма сжатия для одного спектра размером 500 Кбайт составляет не более 5 мс. Соответственно для обработки потока данных в 10 Мбайт/с требуется каждую секунду выделять для сжатия около 100 мс ресурса процессора. Таким образом, задача сжатия на лету данных (ЖХ-МС-МС)-эксперимента выполняется при использовании не более чем 10 % ресурса процессора.

Алгоритм является удачным компромиссом между эффективностью сжатия и скоростью работы алгоритма. Можно было бы использовать более эффективные методики сжатия, основанные на вейвлет- или Фурье-преобразованиях [10], но при этом сильно возросла бы нагрузка на процессор. Мы ограничились своим поиском простыми и эффективными алгоритмами, которые не только позволяют выполнить сжатие информации за время ее получения, но и оставляют большую часть ресурса процессора свободной для выполнения других задач.

СХЕМА ПОДКЛЮЧЕНИЯ СИСТЕМЫ РЕГИСТРАЦИИ К ВПМС-ТАНДЕМУ

Целью выполняемой работы является создание высокопроизводительного ВПМС-тандема, и система регистрации на основе AP100 предназначена именно для использования в составе этого тандема. Результаты предварительных испытаний показывают, что возможности AP100 как системы регистрации выглядят адекватными задаче регистрации данных в биохимических экспериментах

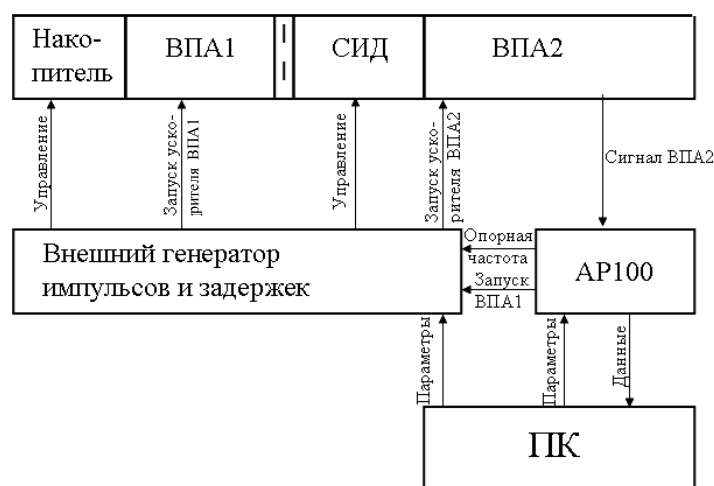


Рис. 13. Схема подключения системы регистрации к ВПМС-тандему

на ВПМС-тандеме. Схема подключения AP100 к ВПМС-тандему показана на рис. 13. Поскольку именно в системе регистрации требуется особая точность соблюдения временных параметров, при подключении к ВПМС-тандему именно система регистрации становится центром синхронизации. Цикл измерений системы регистрации завязан на работу платы AP100. Эта плата служит источником опорной частоты для работы всей системы. Эта же плата запускает цикл ВПА1. Сигнал для запуска ВПА2 генерируется отдельным блоком на основе тактовой частоты, также предоставленной AP100. Только в этом случае возможно обеспечить устойчивость сигналов фрагментных спектров во временной шкале AP100.

Как в памяти AP100 образуется пакет спектров, можно наблюдать на рис. 14. На этом рисунке представлена иерархия промежутков времени в полном цикле накопления сигнала. Эта схема показывает, как за 3 с времени накопления получить набор из ~250 фрагментных спектров, соответствующих диапазонам масс в спектре ВПА1.

Все временные параметры на этом рисунке могут варьироваться. Число фрагментных спектров, на которое разбивается исходный сигнал ВПА1, зависит от решаемой задачи. Время пролета в ВПА2 может изменяться для разного числа фрагментных спектров. Эта схема только качественно показывает способ использования системы регистрации в работе ВПМС-тандема.

Отметим, что сдвиг времени регистрации спектра относительно цикла работы ВПА2 позволит увеличить время пролета ионов в ВПА2. При этом самые тяжелые ионы спектра будут регистрироваться уже после запуска следующей порции ионов, в то время когда эта следующая порция еще не достигла детектора. За счет этого память AP100 используется наиболее эффективно, без пустых областей.

Есть некая потеря времени на передачу накопленных данных по шине PCI. Можно предложить во время передачи сигнала по шине PCI использовать альтернативную более низкочастотную систему регистрации для снятия спектра с ВПА1.

Представленная схема — это базовый, типичный способ включения тандема. За счет использования информации из спектра ВПА1 и более сложного способа управления СИД и ВПА2 можно получать фрагментные спектры не для диапазонов масс, как в представленной схеме, а для компонент исходного спектра ВПА1.

ЗАКЛЮЧЕНИЕ

Работа по испытанию системы регистрации на основе скоростного осреднителя сигналов AP100 носит в сущности промежуточный характер. Эта

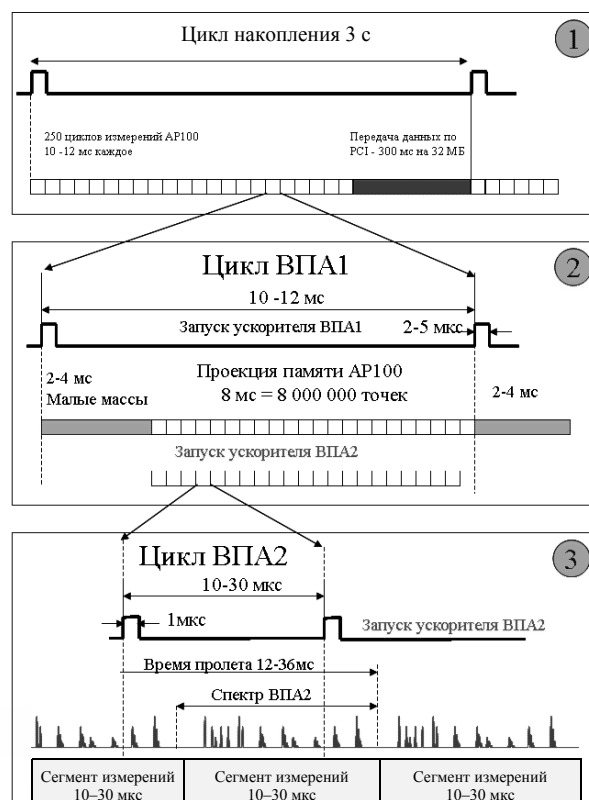


Рис. 14. Иерархия времен при получении пакета фрагментных спектров на ВПМС-тандеме при помощи системы регистрации на основе AP100

работа значима только в контексте всей работы лаборатории по созданию высокопроизводительного ВПМС-тандема, и на пути к его созданию необходимо решить множество проблем, связанных в первую очередь с прикладной обработкой результатов измерений. Однако уже сейчас нарабатываемые способы использования системы регистрации позволяют судить, что возможности AP100 как системы регистрации выглядят адекватными задаче регистрации данных в биохимических экспериментах на ВПМС-тандеме.

Счетный режим использования АЦП AP100 дает возможность фиксирования сигнала ВПМС в большом динамическом диапазоне и с хорошей точностью масс.

Временные параметры AP100 по предварительным оценкам устойчивы на промежутке времени, необходимом для регистрации сигнала ВПМС-тандема.

Разработанные алгоритмы сжатия и сохранения информации позволяют сохранять всю полезную информацию за время эксперимента при разумном объеме хранимых данных.

Задача по реализации на основе AP100 системы регистрации для ВПМС-тандела остается актуальной. Предложенная схема подключения к ВПМС-танделу носит предварительный характер и, наверняка, будет нуждаться в уточнениях по ходу работы. Система регистрации для ВПМС-тандела на основе AP100 позволяет фиксировать фрагментные спектры для всех компонент исходной смеси, что дает принципиально новый уровень масс-спектрометрического анализа сложных смесей.

Количественно система регистрации для ВПМС-тандела позволяет уверенно фиксировать до 100 фрагментных спектров в 1 секунду, что дает в 30 раз больше информации для последующего анализа средствами биоинформатики, чем существующие системы. Скоростная обработка такого объема информации для получения конечного биологического ответа — предмет наших будущих исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Веренчиков А.Н. Параллельный (МС-МС)-анализ во времяпролетном танделе. Постановка задачи, метод и схемы приборов // *Здесь*. С. 24–37.
2. Явор М.И., Веренчиков А.Н. Планарный многоотражательный времяпролетный масс-анализатор, работающий без ограничения диапазона масс // *Здесь*. С. 38–45.
3. Gedcke D.A. ORTEC Application Note AN61: How Counting Statistics and the ADC Sampling Interval Control Mass Accuracy in Time-of-Flight Mass Spectrometry. 2002. 27 p. www.ORTEC-inline.com.
4. Веренчиков А.Н. Тандемные масс-спектрометры в биохимии // *Здесь*. С. 4–23.
5. Подольская Е.П., Новиков А.В., Миргородская О.А. Масс-спектрометрическое определение концентрационных соотношений галогенмодифицированных 2'-дезоксинуридинов // *Здесь*. С. 116–120.
6. Flora J.W., Null A.P., Muddiman D.C. Dual-Micro-ESI Source for Precise Mass Determination on a Quadrupole Time-of-Flight Mass Spectrometer for Genomics and Proteomics Applications // *Analytical and Bioanalytical Chemistry*. 2002. N 7. P. 538–546.
7. Хасин Ю.И., Веренчиков А.Н., Гаврик М.А., Явор М.И. Первые экспериментальные исследования планарного многоотражательного ВПМС // *Здесь*. С. 59–71.
8. Хасин Ю.И., Гаврик М.А., Явор М.И., Демидов В.Н. Импульсная цезиевая пушка для исследования свойств многоотражательного времяпролетного масс-анализатора // *Здесь*. С. 72–79.
9. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Диалог-МИФИ, 2002. 384 с.
10. Короновский А., Храмов А. Непрерывный вейвлетный анализ и его приложения. М.: Физматлит, 2003. 176 с.

Институт аналитического приборостроения РАН, Санкт-Петербург

Материал поступил в редакцию 20.04.2004.

DATA ACQUISITION SYSTEM FOR TOF-MS AND TOF-MS TANDEM

Y. I. Lutvinskiy, D. M. Petrov, A. N. Verenchikov, Y. I. Khasin, M. A. Gavrik

Institute for Analytical Instrumentation RAS, Saint-Petersburg

This paper describes data acquisition system for Time-of-Flight Mass Spectrometer based on a commercial high-speed signal averager AP100 (AcqirisTM). Counting Mode for AP100 ADC has been characterized. Feasibility of acquisition system has been tested with two different instruments. First instrument is a home-built Time-of-Flight Mass Spectrometer used in our lab for biochemical studies. Second instrument is a prototype of Multi Reflection Time-of-Flight mass Spectrometer with long (about 10 milliseconds) time of flight. All experiments were controlled by our data acquisition software, which includes fast compression algorithms and allows storing large arrays of LC-MC experimental data without any losses. Our studies show that AP100 based data acquisition system is satisfying requirements of TOF tandems with 2D parallel MS-MS analysis.