
**МАСС-СПЕКТРОМЕТРИЯ ДЛЯ БИОТЕХНОЛОГИИ.
ИНТЕРПРЕТАЦИЯ ДАННЫХ, МЕТОДОЛОГИЯ, ПРИМЕНЕНИЕ**

УДК 621.384.668.8: 54–185

© Я. И. Лютвинский, В. В. Макаров, Б. Н. Козлов, Е. П. Подольская, А. С. Труфанов,
М. В. Апацкая, И. А. Краснов, Н. В. Краснов, А. Н. Веренчиков

**ОЦЕНКА ЕМКОСТИ МАСС-СПЕКТРОВ ВЫСОКОГО
РАЗРЕШЕНИЯ ПРИ АНАЛИЗЕ СЛОЖНЫХ СМЕСЕЙ**

Использование высокоразрешающих многоотражательных планарных времяпролетных масс-спектрометров (МОП ВПМС) привело к значительному увеличению емкости и информативности масс-спектров. В качестве модельных смесей обследовались триптические гидролизаты смеси двух белков как в режиме прямой инъекции пробы, так и в режиме ХЖ-МС. С помощью компьютерного моделирования показано преимущество МОП ВПМС перед традиционными ВПМС для анализа полного гидролизата бактерий по методу AMT Tag.

ВВЕДЕНИЕ

Быстрый прогресс аналитической химии и биоинформатики дает возможность использовать новые подходы для идентификации протеинов, для понимания их биологических функций, разработки методов ранней диагностики заболеваний, создания принципиально новых лекарств и методов лечения. Для реализации таких программ требуется экспресс-анализ сложных смесей (до миллиона компонент) в большом диапазоне концентраций (более 4 порядков величины) и при скоростном потоке таких анализов (сотни и тысячи анализов в день) для достижения статистически достоверных выводов. При разработке биохимических аналитических стратегий масс-спектрометрия играет ключевую роль как наиболее достоверный метод анализа.

Одна из наиболее производительных методик анализа смесей высокой сложности разработана группой Ричарда Смита (Smith R.D.) и носит название Accurate Mass and Time Tag (AMT Tag) [1]. Эта методика применима для анализа гидролизатов сложных смесей белков и подразумевает идентификацию пептидов по точной массе и времени элюирования из колонки жидкостного хроматографа. Методика предназначена для использования с масс-спектрометрами ионно-циклотронного резонанса (FT-ICR MS или FTMS) и времяпролетными масс-спектрометрами [2]. При скоростном анализе (1–5 мин) с применением жидкостной хроматографии ультравысокого давления (до 1000 атмосфер) времяпролетные масс-спектрометры предпочтительны из-за скоростной регистрации спектров без потери массовой точности [3].

Мы считаем перспективным использование многоотражательных планарных времяпролетных масс-спектрометров (МОП ВПМС) с разрешающей способностью 30 000–50 000, описанных

в статьях этого сборника, для проведения анализов по методике AMT Tag, поскольку МОП ВПМС обладает повышенной по сравнению с традиционными ВПМС точностью определения масс и сохраняет характерную для ВПМС скорость регистрации спектров. Однако для анализа по методике AMT Tag таких сложных смесей, как полный клеточный лизат, необходимо точно измерить массы значительного процента компонент смеси в большом динамическом диапазоне концентраций, поэтому предварительно мы должны оценить сложность смесей, анализ которых возможен при использовании МОП ВПМС и высокоэффективной жидкостной хроматографии (ВЭЖХ).

В данной работе показана разделяющая способность МОП ВПМС при анализе смесей пептидов, показана разделяющая способность хроматографического анализа при соединении с масс-спектрометром и сделана попытка при помощи численного эксперимента оценить сложность смесей, которые могут быть проанализированы посредством тандема ВЭЖХ—МОП ВПМС.

МАТЕРИАЛЫ И МЕТОДЫ

Масс-спектрометр детально описан в статье [4]. Планарный многоотражательный времяпролетный масс-анализатор согласован с источником ионов "электроспрей" с использованием линейной ионной ловушки с аксиальным выбросом ионов. Чувствительность прибора регулируется за счет изменения времени наполнения ионной ловушки. Минимальное время охлаждения ионов в ловушке составляет 2 мс. Однако время напуска ионов может быть уменьшено до 50 мкс. Характерная разрешающая способность варьируется от 30 000 до 100 000 в зависимости от режимов оперирования и массового диапазона.

Для данного исследования режим работы МОП ВПМС был выбран исходя из компромисса между высокой разрешающей способностью и чувствительностью. Разрешающая способность прибора составляет ~40 000 при времени наполнения ловушки 500 мкс. В этом режиме чувствительность МОП ВПМС достаточна для представленного количества вещества пробы.

Для сравнения был использован времяпролетный масс-спектрометр MX5303 с источником ионов "электроспрей" и ортогональным вводом, разработанный в нашей лаборатории [5]. Разрешающая способность этого прибора — около 8000.

Источник ионов оборудован пневматическим электрораспылителем, обеспечивающим большой диапазон расходов анализируемой жидкости. Для инъекции индивидуальных растворов использовался объемный расход 5 мкл/мин. При стыковке масс-спектрометра с хроматографом объемный расход жидкости увеличивался до 100 мкл/мин.

В работе был использован жидкостный хроматограф "Милихром А-02" (ЗАО "ЭкоНова", Новосибирск, Россия). Колонка 2×75 мм с Nucleosil 100-5 C18 (Macherey-Nagel, Duren, Germany). Температура колонки 35 °С. Подвижные фазы: А — раствор 0.5 % HCOOH в H₂O (pH 2), В — ацетонитрил. Градиентное элюирование: от 5 до 70 % В за 1500 мкл. Скорость потока 100 мкл/мин.

При проведении эксперимента были использованы следующие образцы:

1) смесь пептидов, полученных в процессе триптического гидролиза бычьего сывороточного альбумина (BSA) (производство компании Sigma, чистота 96 %, концентрация 10^{-5} М);

2) смесь пептидов, полученных в процессе триптического гидролиза бычьего бета-лактоглобулина (bLG) (производство компании Sigma, чистота 98 %, концентрация 10^{-5} М).

Смесь пептидов была приготовлена по следующей методике: белок растворяли в 8 М мочеvine (1 мг/мл), инкубировали в течение 2 ч при комнатной температуре. Отбирали аликвоту 50 мкл. Затем к раствору добавляли 50 мкл 25 мМ раствора DTT в 50 мМ NH₄HCO₃ и инкубировали при 57 °С в течение 45 минут. После чего к смеси добавляли 50 мкл 40 мМ IAcNH₂ в 50 мМ NH₄HCO₃ и выдерживали в темноте 20 мин. Трипсинолиз проводили при 37 °С. Трипсин добавляли в соотношении 1/100 моль, инкубировали 3 ч, затем добавляли трипсин еще раз в том же соотношении и выдерживали еще 3 ч. Трипсинолиз останавливали добавлением 10 мкл ледяной уксусной кислоты. Контроль полноты прохождения трипсинолиза выполняли за счет осаждения высокомолекулярных остатков белка 100 % ацетонитрилом, так чтобы его конечная концентрация составляла 70%. При отсутствии замутнения результат трипсинолиза считали удовлетворительным. Обессоливание

пробы проводили с помощью обращеннофазной хроматографии на сорбенте POROS (C18). Затем пробу высушивали и растворяли в смеси CH₃CN и H₂O в соотношении 1:1 до достижения концентрации белка 10^{-5} М.

ЕМКОСТЬ МАСС-СПЕКТРА МОП ВПМС ПРИ ПРЯМОЙ ИНЖЕКЦИИ ПРОБЫ

Раствор, содержащий смесь триптических гидролизатов двух белков (BSA и bLG) в отношении 1:1 распылялся в источнике ионов "электроспрей" в течение 1 мин. Результирующий масс-спектр представлен на рис. 1, а. Область спектра в диапазоне масс 462–464 Да представлена на рис. 1, б. В спектре выявлено 1011 пиков. С использованием разработанного в лаборатории алгоритма IPeX [6] пики автоматически сгруппированы в изотопные мультиплеты, соответствующие 296 индивидуальным компонентам смеси. При такой комплексности смеси в спектре наблюдаются множественные перекрытия изотопных мультиплетов. Пример такого перекрытия наглядно представлен на рис. 1, б. На этом участке спектра наблюдается перекрытие трех изотопных мультиплетов:

- изотопный мультиплет однозарядных ионов, соответствующий моноизотопной массе $[M+H]^+ = 462.1498$ Да;
- изотопный мультиплет однозарядных ионов, соответствующий моноизотопной массе $[M+H]^+ = 462.2228$ Да;
- изотопный мультиплет двузарядных ионов, соответствующий моноизотопной массе $[M+H]^+ = 921.4855$ Да — на рисунке представлены второй и третий сигналы этого мультиплета.

Стоит отметить что, разделение индивидуальных изобарных пиков соединений, соответствующих одной и той же номинальной массе, но имеющих различные массовые дефекты из-за различий в элементном составе, происходит благодаря высокой разрешающей способности прибора (~40 000).

Для сравнения на рис. 1, в приведен выделенный сегмент масс-спектра в том же массовом диапазоне, полученный на приборе MX5303 с разрешающей способностью ~ 8000. Из рисунка видно, что разрешающей способности порядка 8000 недостаточно для разделения изобарных пиков. Как следствие, при обработке данных этого масс-спектра зарегистрировано меньшее количество пиков (382 против 1011 для МОП ВПМС), а значения найденных масс искажены наложением неразделенных пиков.

Для анализа значимости дополнительной информации, доступной в экспериментах с использованием МОП ВПМС, стоит сравнить результаты, полученные при интерпретации спектров

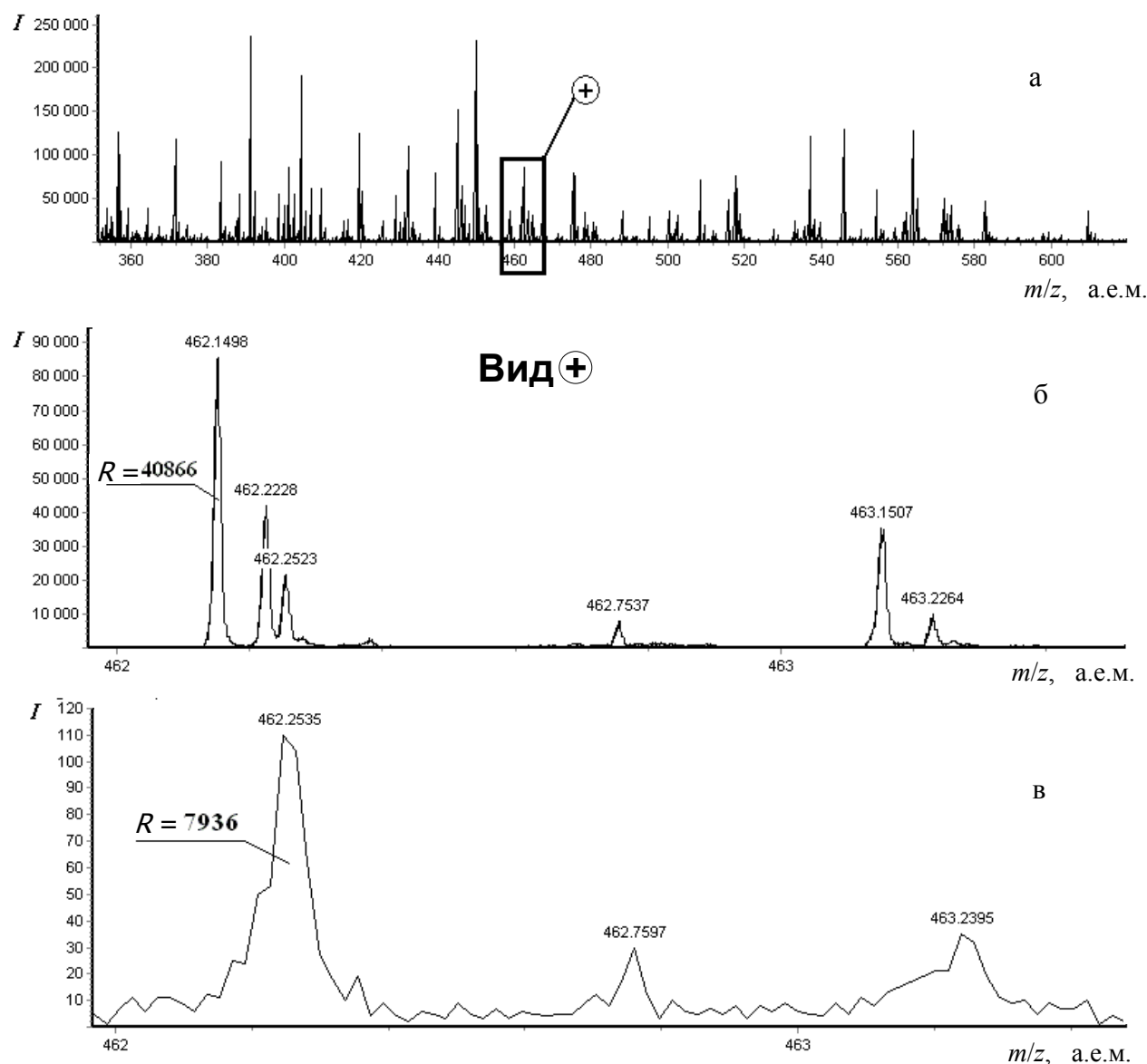


Рис. 1. Масс-спектры гидролизата смеси белков BSA и bLG, полученные на МОП ВПМС (а, б) и на ВПМС MX5303 (в)

с высоким и низким разрешением. Интерпретация была проведена с помощью пакета программного обеспечения Mascot для идентификации белков методом PMF [7].

В отчете программы Mascot [8] в качестве наиболее вероятного результата идентификации показана смесь из двух белков, включающая в себя как BSA, так и bLG. Отчет Mascot показан на рис. 2. Рейтинг 315, присвоенный этому результату идентификации, позволяет считать идентификацию практически однозначной. Этот рейтинг вычисля-

ется как $-10 \cdot \log_{10}(P)$, где P — вероятность образования спектра из случайных событий. При этом в спектре обнаружено 69 сигналов соответствующих пептидам, покрывающим значительную часть аминокислотной последовательности исходных белков.

Интерпретация данных масс-спектра той же пробы, полученного на приборе с разрешающей способностью ~ 8000 не дает надежной идентификации ни одного из белков пробы. Реально присутствующий в смеси BSA находится только

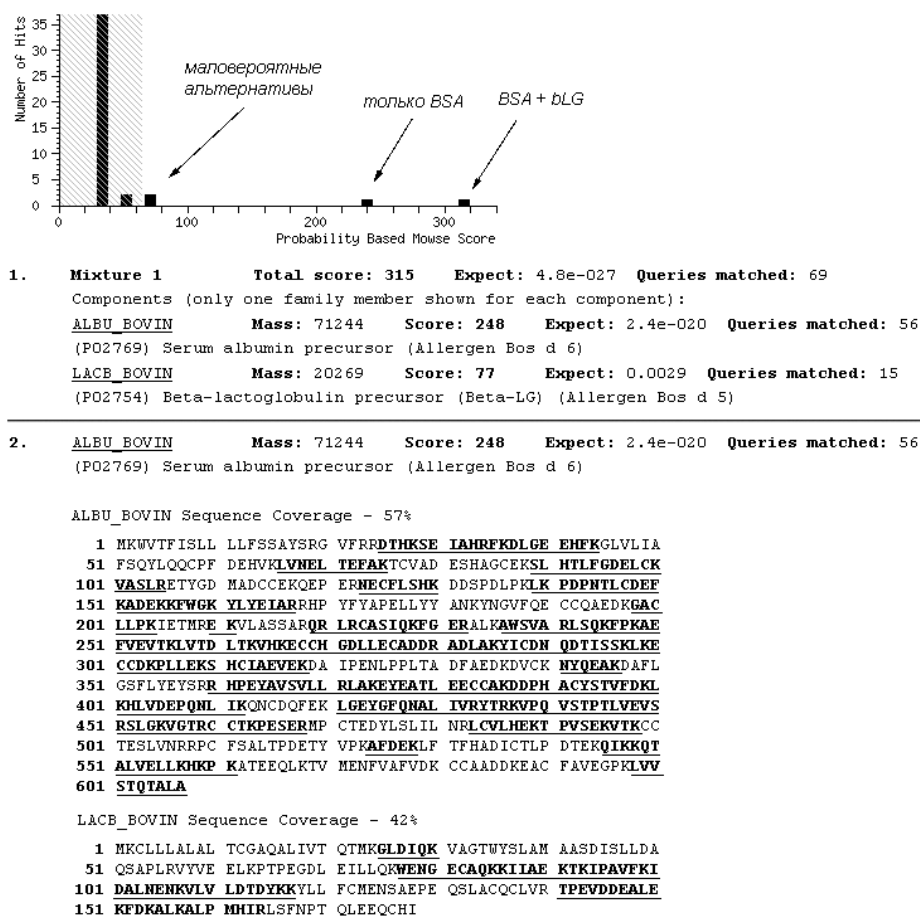


Рис. 2. Результат интерпретации спектра МОП ВПМС программой Mascot

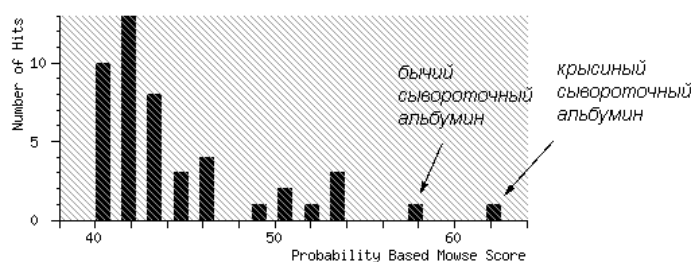


Рис. 3. Результат интерпретации спектра ВПМС MX5303 программой Mascot

на втором месте по рейтингу вероятной идентификации, в то время как на первом месте по рейтингу стоит гомолог BSA — сывороточный альбумин крысы (рис. 3).

Таким образом, различие в биохимическом результате оказалось качественным — это различие между верным и неверным ответом. Стоит отметить, что автоматическая надежная идентификация смеси двух белков методом PMF — это редкий результат в практике протеомных работ. Используемые для проведения идентификации белков этим методом стандартные масс-спектрометры ESI-TOF и MALDI-TOF не обладают достаточной разрешающей способностью для разделения компонент смеси, показанной на МОП ВПМС.

Как правило, идентификация состава смеси белков умеренной сложности, требует проведения (ЖХ-МС-МС)-эксперимента продолжительностью 20–40 мин, т. е. требует существенно большего времени для идентификации белков, чем метод PMF, подразумевающий прямой ввод неразделенной пробы в масс-спектрометр.

ПРЯМАЯ СТЫКОВКА МОП ВПМС С ЖИДКОСТНЫМ ХРОМАТОГРАФОМ

Дополнительно к эксперименту с прямым вводом гидролизата белков в масс-спектрометр была выполнена стыковка МОП ВПМС с жидкостным хроматографом "Милихром А-02", после чего проведен (ЖХ-МС)-анализ той же смеси двух белков. На рис. 4 представлен результат проведенного эксперимента в виде двумерной растровой диаграммы в осях "Время элюирования—Отношение массы к заряду" [9]. На карте видны хорошо разделенные друг от друга сигналы изотопных мультиплетов пептидов. Такое представление наглядно показывает, что масс-спектры смеси пептидов BSA и bLG в (ЖХ-МС)-эксперименте существенно упрощаются. Хроматографическое разделение позволяет уменьшить влияние химического шума и разнести по времени информативные сигналы пептидов. На показанных данных эксперимента ВЭЖХ-МС отдельные спектры содержат не более 20 компонент.

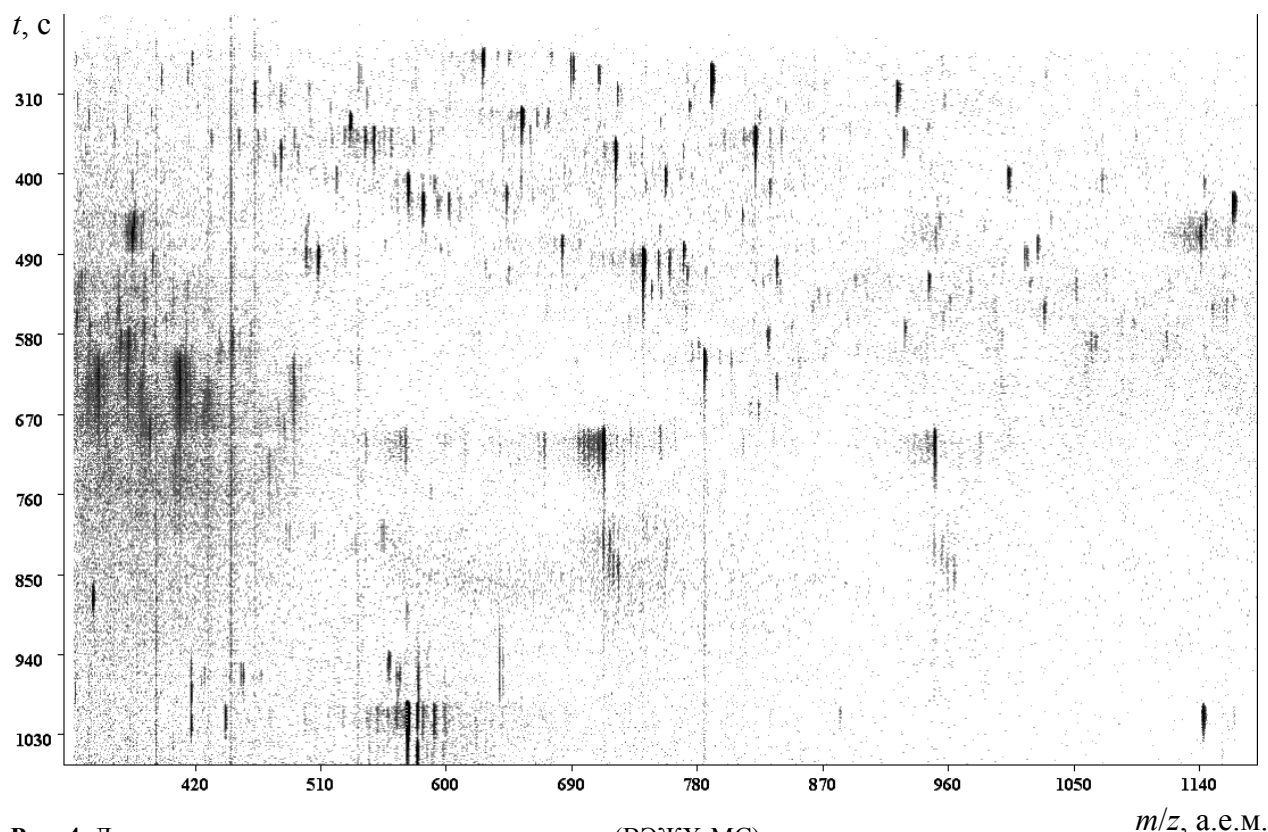


Рис. 4. Двумерная растровая диаграмма результатов (ВЭЖХ-МС)-эксперимента

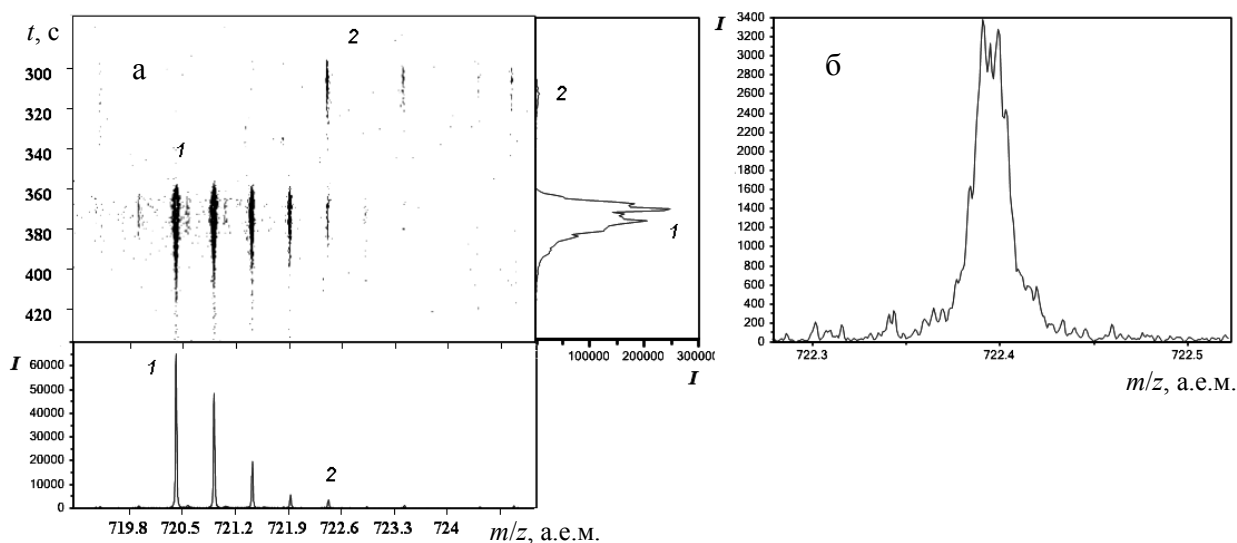


Рис. 5. Фрагмент двумерной растровой диаграммы и проекции диаграммы на оси масс-спектра и ионного тока в интервале масс (а) и неразрешенный пик, включающий сигналы 1 и 2, (б)

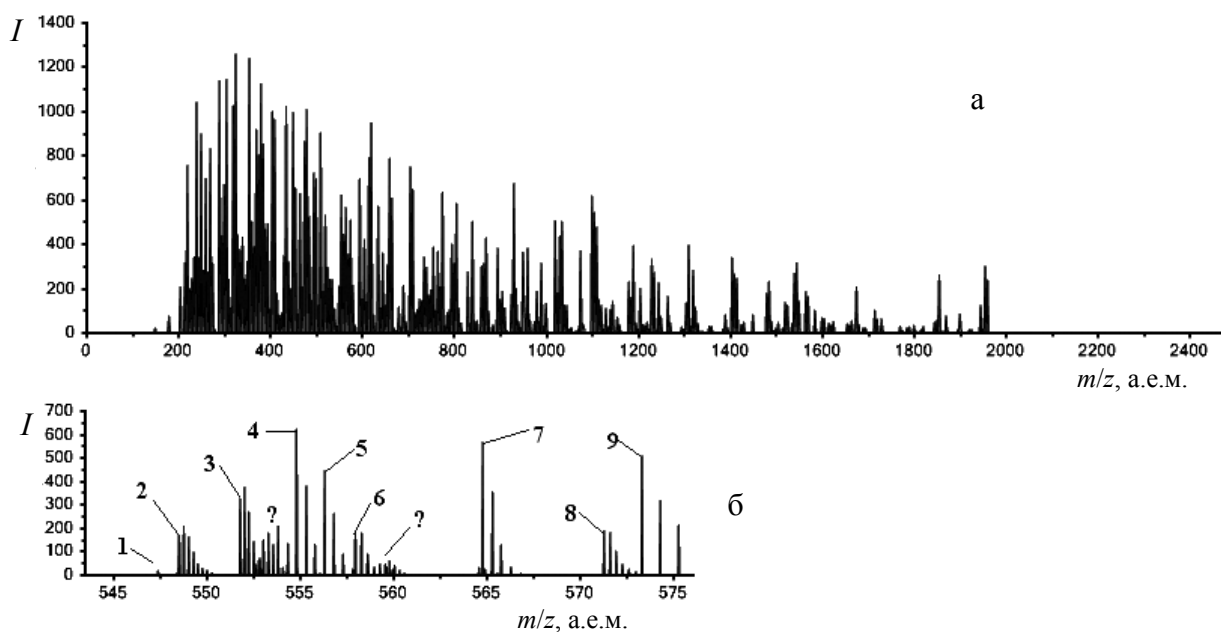


Рис. 6. Моделированный масс-спектр — число компонент 500, разрешение 100 000.
а — общий вид, б — фрагмент

Использование хроматографа "Милихром А-02" позволило упростить состав пробы на входе масс-спектрометра в среднем примерно в 20 раз. Хотя для данного состава пробы хроматографическое разделение не является необходимой стадией эксперимента, тем не менее в ходе (ВЭЖХ-МС)-анализа удастся дополнительно выявить отдельные компоненты, неразличимые в прямом МС-анализе.

На рис. 5 продемонстрировано разделение двух сигналов. Фрагмент двумерной растровой диа-

граммы, представленный на рис. 5, а, содержит сигналы двух изотопных мультиплетов: 1 — от двузарядного иона с $m/z = 720.7$ Да и 2 — от однозарядного иона с $m/z = 722.3$ Да. Эти изотопные мультиплеты перекрыты таким образом, что разрешения 40 000 недостаточно для выделения моноизотопного пика второго сигнала, что показано на проекциях ионного тока на оси масс-спектра и времени элюирования. Рис. 5, б показывает неразрешенный пик, соответствующий перекрытию моноизотопного

пика сигнала 2 с $m/z = 722.391$ Да и пятого пика изотопного мультиплета сигнала 1 с $m/z = 722.398$ Да. Только благодаря хроматографическому разделению удается выделить оба сигнала.

Предварительное хроматографическое разделение становится особенно эффективным при анализе смесей, число компонент в которых существенно превышает максимальную вместимость масс-спектра.

ВЫЧИСЛИТЕЛЬНАЯ ОЦЕНКА ЕМКОСТИ ВЫСОКОРАЗРЕШАЮЩИХ МАСС-СПЕКТРОВ

Для проведения теоретической оценки числа детектируемых компонент для МС и (ЖХ-МС)-экспериментов с использованием высокоразрешающего масс-спектрометра был проведен следующий численный эксперимент.

Расчетным путем был сгенерирован набор модельных масс-спектров для различных значений разрешающей способности прибора и различной сложности состава исходной пробы [6]. Каждый спектр содержит теоретически рассчитанные изотопные и зарядовые распределения для каждого из предварительно отобранного набора пептидов. В среднем каждому пептиду соответствует 2–3 зарядовых состояния, каждое из которых состоит из 3–5 изотопных пиков. Таким образом, в среднем каждому пептиду в спектре соответствует около 10 сигналов. Интенсивность сигналов пептидов была выбрана случайным образом в пределах двух порядков динамического диапазона.

Каждый модельный масс-спектр был обработан при помощи алгоритма зарядовой и изотопной декомпозиции IPeX. Набор молекулярных масс, извлеченный при помощи IPeX, был подвергнут сравнению с начальным набором масс пептидов. Если ошибка в точности определения масс не превышала 1 ppm, считалось, что масса компоненты корректно выделена.

Пример модельного масс-спектра из 500 компонент с моделируемой разрешающей способностью 100 000 представлен на рис. 6. Из рис. 6, б видно, что множественные пересечения изотопных мультиплетов не поддаются ручной обработке. Тем не менее алгоритм IPeX уверенно выявляет более 95 % пептидов в этом спектре.

В первой серии расчетных экспериментов спектры были построены с характерной для действующего МОП ВПМС разрешающей способностью 40 000. Рис. 7 показывает число корректно распознанных компонент в зависимости от сложности исходной смеси. На рисунке видно, что насыщение спектра по числу выявленных компонент наступает на уровне 700 компонент. Однако эффекты перекрытия изотопных мультиплетов становятся заметны уже для смесей из нескольких со-

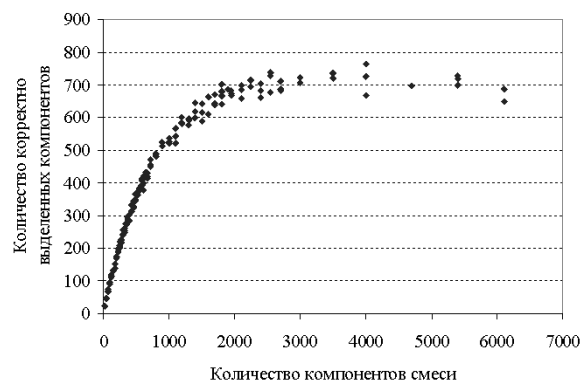


Рис. 7. Характеристика выделения компонент из масс-спектров смесей различной сложности с разрешающей способностью $R = 40\,000$

тен компонент. Так, например, в смеси действительно состоящей из 700 компонент мы можем различить около 70 % сигналов.

Во второй серии численного эксперимента была проведена попытка установить, какая разрешающая способность необходима для масс-спектрометрического анализа клеточного лизата в сочетании с жидкостной хроматографией. Для этого был промоделирован процесс триптического гидролиза для всех белков, относящихся к бактерии *E. Coli*, доступных в базе данных Swiss-prot. *E. Coli* является традиционным полигоном биологических экспериментов и благодаря этому протеом *E. Coli* хорошо известен и практически полностью представлен в базе данных.

На основе полученного набора аминокислотных последовательностей пептидов были сделаны выборки, соответствующие степени разделения исходной смеси пептидов при помощи различных инструментов жидкостной хроматографии:

- 7500 пептидов — разделение пептидной смеси в 20 раз, соответствующее использованию штатного хроматографа "Милихром А-02";
- 2000 пептидов — разделение пептидной смеси в 75 раз, соответствующее использованию капиллярной хроматографии — стандарта де факто в практике современных протеомных исследований [10];
- 500 пептидов — разделение пептидной смеси в 300 раз, соответствующее использованию хроматографии ультравысокого давления [11].

Далее для полученных выборок были составлены масс-спектры с различной разрешающей способностью. На рис. 8 показана зависимость доли корректно выявленных компонент от разрешения масс-спектра. Из рисунка видно, что при повышении разрешающей способности прибора с 20 000 до 200 000 наблюдается наиболее существенный

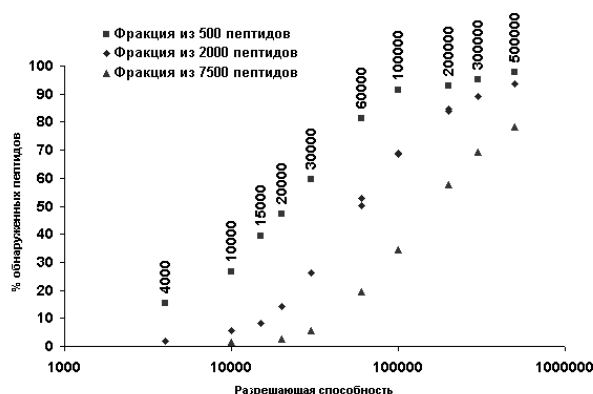


Рис. 8. Процент обнаруженных пептидов в модельных масс-спектрах триптического гидролизата *E. Coli* в зависимости от разрешающей способности масс-спектрометра

прирост числа выделяемых компонент. Таким образом, переход от разрешающей способности современных приборов (порядка 15 000) к МОП ВПМС с разрешающей способностью 40 000 позволит существенно повысить сложность анализируемых смесей.

Следует отметить, что при проведении численных экспериментов невозможно учесть множество факторов, существенных при регистрации экспериментального масс-спектра, таких как наличие химического шума, реальный динамический диапазон сигналов масс-спектра и многое другое. Поэтому представленный численный эксперимент следует рассматривать как верхнюю оценку потенциальных возможностей высокоразрешающих масс-спектрометров в совокупности с имеющимися методами обработки масс-спектров.

ВЫВОДЫ

Результаты вычислительных экспериментов, представленные на рис. 7 и 8, позволяют сделать следующие выводы. Применение высокоэффективной хроматографии, и особенно хроматографии ультравысокого давления, при анализе смесей пептидов сложности, сопоставимой с полным лизатом бактерий, позволяет определять десятки процентов компонент, исходно представленных в смеси. Этого принципиально достаточно для анализов с использованием метода AMT Tag, поскольку для идентификации белков в методе AMT Tag достаточно идентифицировать 5–8 пептидов для каждого белка.

В случае недоступности в лаборатории хроматографии ультравысокого давления или хотя бы капиллярной хроматографии для применения ме-

тодик типа AMT в экспериментах, проводимых на МОП ВПМС, требуется снизить сложность исходной смеси пептидов до такого уровня, при котором в каждом спектре регистрация 500–700 компонент в каждом спектре будет достаточной для идентификации. Это может быть достигнуто несколькими способами.

- Увеличением числа ступеней предварительного разделения пептидов, например введением двухступенчатой (ЖХ-ЖХ) и(или) параллельной хроматографии.

- Снижением сложности смеси пептидов, например, за счет выделения для анализа только цистеин-содержащих пептидов [12].

- Снижением сложности исходной пробы за счет предварительного фракционирования белков.

При анализе многокомпонентных смесей пептидов возможности МОП ВПМС, равно как и приборов ионно-циклотронного резонанса FTMS, не гарантируют извлечения 100 % компонентов смеси. Однако на данном этапе исследования возможности МОП ВПМС представляются адекватными методике AMT Tag, что дает основания для продолжения исследований в этом направлении, и в первую очередь проверки расчетных предположений в практической работе с МОП ВПМС.

СПИСОК ЛИТЕРАТУРЫ

1. Pasa-Tolic L., Masselon C., Barry R.C., Shen Y., Smith R.D. Proteomic analyses using an accurate mass and time tag strategy // *Biotechniques*. 2004. V. 37, N 4. P. 621–624, 626–633.
2. Strittmatter E.F., Ferguson P.L., Tang K., Smith R.D. Proteome analyses using accurate mass and elution time peptide tags with capillary LC time-of-flight mass spectrometry // *J. Am. Soc. Mass Spectrom.* 2003. V. 14, N 9. P. 980–991.
3. Shen Y. et al. Making broad proteome protein measurements in 1–5 min using high-speed RPLC separations and high-accuracy mass measurements // *Anal. Chem.* 2005. V. 77, N 23. P. 7763–7773.
4. Козлов Б.Н. и др. Многоотражательный время-пролетный масс-спектрометр с ионной ловушкой на входе // *Этот выпуск*. С. 40–48.
5. Новиков А.В. и др. Контроль производства генно-инженерного инсулина с помощью тандема "микроколоночный жидкостный хроматограф—масс-спектрометр" в режиме прямого ввода образца // *Научное приборостроение*. 2004. Т. 14, № 2. С. 109–115.
6. Макаров В.В. и др. Алгоритм извлечения Аналитически значимой информации из масс-спектрометрических данных экспериментов протеомики // *Научное приборостроение*. 2006. Т. 16, № 2. С. 92–100.

7. Pappin D.J., Hojrup P., Bleasby A.J. Rapid identification of proteins by peptide-mass fingerprinting // *Curr. Biol.* 1993. V. 3, N 6. P. 327–332.
8. Perkins D.N., Pappin D.J., Creasy D.M., Cottrell J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data // *Electrophoresis*. 1999. V. 20, N 18. P. 3551–3567.
9. Макаров В.В., Лютвинский Я.И., Веренчиков А.Н. Алгоритм IPLEX-2D для извлечения информации о компонентах пробы из массива данных (ВЭЖХ-МС)-экспериментов протеомики // Этот выпуск. С. 107–112.
10. Karlsson K.E., Novotny M. Separation efficiency of slurry-packed liquid chromatography micro-columns with very small inner diameters. // *Anal. Chem.* 1988. V. 60, N 17. P. 1662–1665.
11. Shen Y. et al. Ultrahigh-throughput proteomics using fast RPLC separations with ESI-MS/MS // *Anal. Chem.* 2005. V. 77. P. 6692–6701.
12. Norbeck A.D. et al. The utility of accurate mass and LC elution time information in the analysis of complex proteomes // *J. Am. Soc. Mass. Spectrom.* 2005. V. 16, N 8. P. 1239–1249.

*Институт аналитического приборостроения РАН,
Санкт-Петербург*

Материал поступил в редакцию 27.06.2006.

ESTIMATION OF CAPACITY OF HIGH RESOLUTION MASS SPECTRA FOR ANALYSIS OF COMPLEX MIXTURES

**Ya. I. Lutvinsky, V. V. Makarov, B. N. Kozlov, E. P. Podolskaya, A. S. Trufanov,
M. V. Apatskaya, I. A. Krasnov, N. V. Krasnov, A. N. Verentchikov**

Institute for Analytical Instrumentation RAS, Saint-Petersburg

Using multireflecting time-of-flight mass spectrometers (MR TOF MS) with an improved resolution R , significant improvement in peak capacity and information content of mass spectra is demonstrated. A model digest of two protein mixtures was analyzed in two modes — infusion and LC-MS. Computer simulations show the advantage of MR TOF MS compared to conventional TOF MS for analysis of the whole bacteria digests using the method of AMT Tag.