



По материалам доклада на VII Всероссийской конференции  
“Протеолитические ферменты: структура, функции, эволюция”  
(30 июня – 4 июля 2014 года, г. Петрозаводск)

## ИЗУЧЕНИЕ ПЕРВИЧНОЙ СПЕЦИФИЧНОСТИ ПРОТЕИНАЗ СТАТИСТИЧЕСКИМ АНАЛИЗОМ МАЛДИ МАСС-СПЕКТРОВ ПРОДУКТОВ ПРОТЕОЛИЗА

© 2015 г. Н. Л. Еремеев<sup>#</sup>, А. В. Борзенкова

Московский государственный университет им. М.В. Ломоносова, Химический факультет,  
119991, Москва, Ленинские Горы, 1, стр. 3

Поступила в редакцию 09.06.2014 г. Принята к печати 22.07.2014 г.

Проведена экспериментальная верификация метода изучения первичной специфичности протеолитических ферментов статистическим анализом масс продуктов протеолиза белковых субстратов на протеиназах с известной субстратной специфичностью (глутамилэндопептидаза и трипсин). Предлагаемый метод не требует прямого определения аминокислотной последовательности продуктов протеолиза, достоверно определяет протеиназы с узкой субстратной специфичностью и относительно толерантен к наличию в МАЛДИ масс-спектрах продуктов протеолиза пиков посторонних загрязнений. Показано, что для исключения ложноположительных результатов необходимо использовать набор белковых субстратов с последующим усреднением получаемых статистических данных.

*Ключевые слова:* протеолитические ферменты, первичная специфичность, матрично активированная лазерная десорбция-ионизация (МАЛДИ), масс-спектрометрия.

DOI: 10.7868/S0132342315010042

### ВВЕДЕНИЕ

Ранее [1] нами были теоретически исследованы возможности и ограничения метода изучения первичной специфичности протеолитических ферментов статистическим анализом МАЛДИ-масс-спектров продуктов протеолиза белковых субстратов (под первичной специфичностью подразумевают избирательность фермента по отношению к аминокислоте, образующей гидролизуюмую амидную связь своей карбоксильной группой – P1-положение субстрата по классификации Шехтера-Бергера [2]). Идея предложенного метода заключается в следующем. Проводят гель-электрофорез в восстанавливающих и денатурирующих условиях белка-субстрата с известной аминокислотной последовательностью. Полосу геля с субстратом подвергают протеолизу изучаемой протеиназой. МАЛДИ-масс-спектр реакционной смеси (по-

скольку в методе МАЛДИ образуются практически только однозарядные ионы [3]) дает набор масс продуктов протеолиза.

С помощью программы FindPept, теоретически можно найти все пептиды белкового субстрата, совпадающие по массе с тем или иным экспериментально полученным значением в пределах некоторого диапазона ошибки измерения (генерируемый набор пептидов). Местоположение этих пептидов в аминокислотной последовательности субстрата позволяет выделить для каждого из них две аминокислоты (для C- и N-концевых пептидов одну), после которых потенциально произошел гидролиз с образованием данного продукта (со статистической точки зрения событие произошло). Подобную процедуру проводят для каждой массы в масс-спектре, в результате чего получают набор произошедших событий. Специфичные для данной протеиназы аминокислоты в обязательном порядке будут присутствовать в наборе генерированных пептидов для каждой теоретически полученной массы, в то время как неспецифичные будут присутствовать случайным образом. Следовательно, первичная специфич-

Сокращения: МАЛДИ – матрично-активированная лазерная десорбция/ионизация; НССА – 4-гидрокси- $\alpha$ -цианокоричная кислота; ДНВ – 2,5-дигидроксibenзойная кислота.

<sup>#</sup> Автор для переписки (тел.: +7 (495) 939-34-17, факс: +7 (495) 939-54-17; эл.почта: eremeev@enzyme.chem.msu.ru).

ность протеиназы должна проявляться в максимальной частоте встречаемости той или иной аминокислоты при статистическом анализе полученного набора.

Определение первичной специфичности является обязательным этапом характеристики любой новой протеазы и проводится с использованием синтетических или природных субстратов. К синтетическим субстратам относятся амиды *N*-замещенных аминокислот, а также ди-, три- и олигопептиды с *C*-концевой амидной группой. В этом случае определяют константы гидролиза в сериях гомологичных субстратов и на основании их сопоставления делают вывод о преимущественном предпочтении данным ферментом определенной аминокислоты в *P1*-положении [4–6]. Данный метод требует хорошей синтетической базы, длительной кинетической работы и не дает информации о вторичной специфичности фермента.

При использовании природных белковых субстратов требуется разделение продуктов протеолиза с последующим определением их структуры. Это позволяет выявить гидролизующие пептидные связи и в результате статистического анализа определить не только первичную, но и вторичную специфичность протеиназы [7]. Однако для определения структуры полипептидов *N*-концевым секвенированием [8–10] необходимо хорошее хроматографическое разделение продуктов протеолиза и большие количества реагентов. Использование же для этой цели tandemной масс-спектрометрии [11–13] не всегда оправдывается экономически.

Следует отметить, что попытки прямого определения местоположения продукта протеолиза белкового субстрата исключительно по его массе (а только такую информацию и дает МАЛДИ-масс-спектр) наталкиваются на ряд препятствий. Существует так называемый дефект массы пептидов, заключающийся в разнице между их номинальной расчетной и экспериментально определяемой моноизотопной массами. Эта разница линейно возрастает с увеличением массы пептида и, по разным оценкам, имеет тангенс угла наклона от  $4.55 \times 10^{-4}$  [14] до  $5.7 \times 10^{-4}$  [15] (в среднем  $4.99 \times 10^{-4}$  [16]). Так, например, в работе [17] эта разница достигала 8 Да, в силу чего, для уточнения последовательности продуктов протеолиза, авторы данной статьи прибегали к tandemной масс-спектрометрии или секвенированию по Эдману. Предлагаемый же нами метод не требует прямого определения аминокислотной последовательности продуктов протеолиза и весьма прост инструментально.

Цель настоящей работы заключалась в экспериментальной верификации метода изучения первичной специфичности протеолитических ферментов статистическим анализом масс продуктов протеолиза белковых субстратов на примере известных протеиназ с различной субстратной спе-

цифичностью – трипсине и глутамилэндопептидазе.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Проведенный ранее теоретический анализ предлагаемого метода определения первичной специфичности протеиназ выявил ряд его ограничений [1]. Во-первых, этот подход не может быть реализован для ферментов с относительно широкой субстратной специфичностью. Во-вторых, молекулярная масса белкового субстрата не может превышать 80 кДа. В-третьих, возможно получение ложноположительных результатов для неспецифичных аминокислот, особенно если их содержание в субстрате мало. В этой связи для экспериментальной верификации метода были выбраны протеиназы с узкой субстратной специфичностью: – трипсин (положительно заряженные аминокислоты – аргинин и лизин) и глутамилэндопептидаза (отрицательно заряженные аминокислоты – глутаминовая и аспарагиновая кислоты) соответственно. В качестве белковых субстратов были выбраны относительно низкомолекулярные куриный лизоцим (М.м. около 16 кДа) и катепсин L камчатского краба *Paralithodes camtschaticus* (М.м. 35.4 кДа).

Масс-спектры продуктов протеолиза конкретного субстрата обеими выбранными протеиназами воспроизводились достаточно хорошо. Можно отметить разницу в относительных интенсивностях тех или иных пиков в различных спектрах (данные не приведены). Не все массы проявляются в повторных экспериментах (это связано с высоким заданным уровнем отношения сигнал/шум – не менее 10), однако коэффициент взаимопересечения достоверно определяемых масс между тремя параллельными спектрами для каждой пробы составлял не менее 70–80%. Полученные наборы масс продуктов протеолиза приведены в таблице.

Для каждой экспериментально определенной массы в используемом субстрате находили соответствующие по массе пептидные последовательности в пределах разброса  $\pm 1$  Да. Аминокислотам, найденным в *P1*-положении, присваивали значение  $a_i = 1$  ( $i$  – индекс конкретной аминокислоты) и суммировали эти значения по всему набору генерированных пептидов  $A_i = \sum a_i$ . Полученную сумму событий нормировали на содержание данной аминокислоты  $n_i$  в используемом субстрате:

$$\bar{A}_i = \frac{A_i}{n_i}$$

С целью сравнения результатов, полученных таким образом для различных ферментов и белков-субстратов, величины  $\bar{A}_i$  дополнительно нормировали на максимальное значение, получаемое в конкретном эксперименте, и выражали в процентах.

Массы продуктов гидролиза лизоцима и катепсина L трипсином и глутамилэндопептидазой

Субстрат	Массы продуктов после гидролиза	
	трипсином	глутамилэндопептидазой
Лизоцим	1007.367; 1030.567; 1045.545; 1347.725; 1428.663; 1474.186; 1505.636; 1675.676; 1706.421; 1754.074; 1784.164; 1803.964	1182.674; 1201.076; 1429.411; 1476.025; 1767.856; 1977.562; 2035.698; 2161.321
Катепсин L	1042.561; 1066.896; 1235.062; 1271.815; 1494.195; 1635.408; 1834.580; 2002.620; 2384.597; 2395.957; 2500.071; 2588.300; 2871.368; 3001.865; 3011.693	1004.276; 1016.311; 1060.448; 1095.469; 1216.720; 1307.720; 1357.889; 1415.935; 1571.585; 1602.039; 1745.163; 2313.622; 2518.415; 2525.743; 2666.587; 2688.582; 3740.307

$$\bar{A}_{i,\text{норм}} = \frac{\bar{A}_i}{A_{i,\text{max}}} \times 100\%.$$

Полученные результаты (рисунок, *a–z*), позволяют говорить о следующем. Во-первых, при обработке данных по каждой паре фермент/субстрат максимальные статистические веса имеют именно специфичные для данной протеиназы аминокислоты. Однако если для трипсина при протеолизе обоих белковых субстратов максимальный статистический вес имеет лизин по сравнению с аргинином (рисунок, *a–b*), то для глутамилэндопептидазы максимально статистически значимая аминокислота зависит от гидролизуемого белка: глутаминовая кислота (рисунок, *b*) при протеолизе лизоцима и аспарагиновая кислота (рисунок, *c*) при протеолизе катепсина L. Во-вторых, в ряде случаев появляются ложноположительные результаты. Так, при гидролизе лизоцима трипсином (рисунок, *a*) достаточно большой вес имеет пролин, при гидролизе катепсина L глутамилэндопептидазой (рисунок, *c*) большой статистический вес имеют лизин и метионин. Тем не менее первичная специфичность использованных ферментов проявляется достаточно четко, особенно при усреднении результатов по двум белковым субстратам (рисунок, *d–e*).

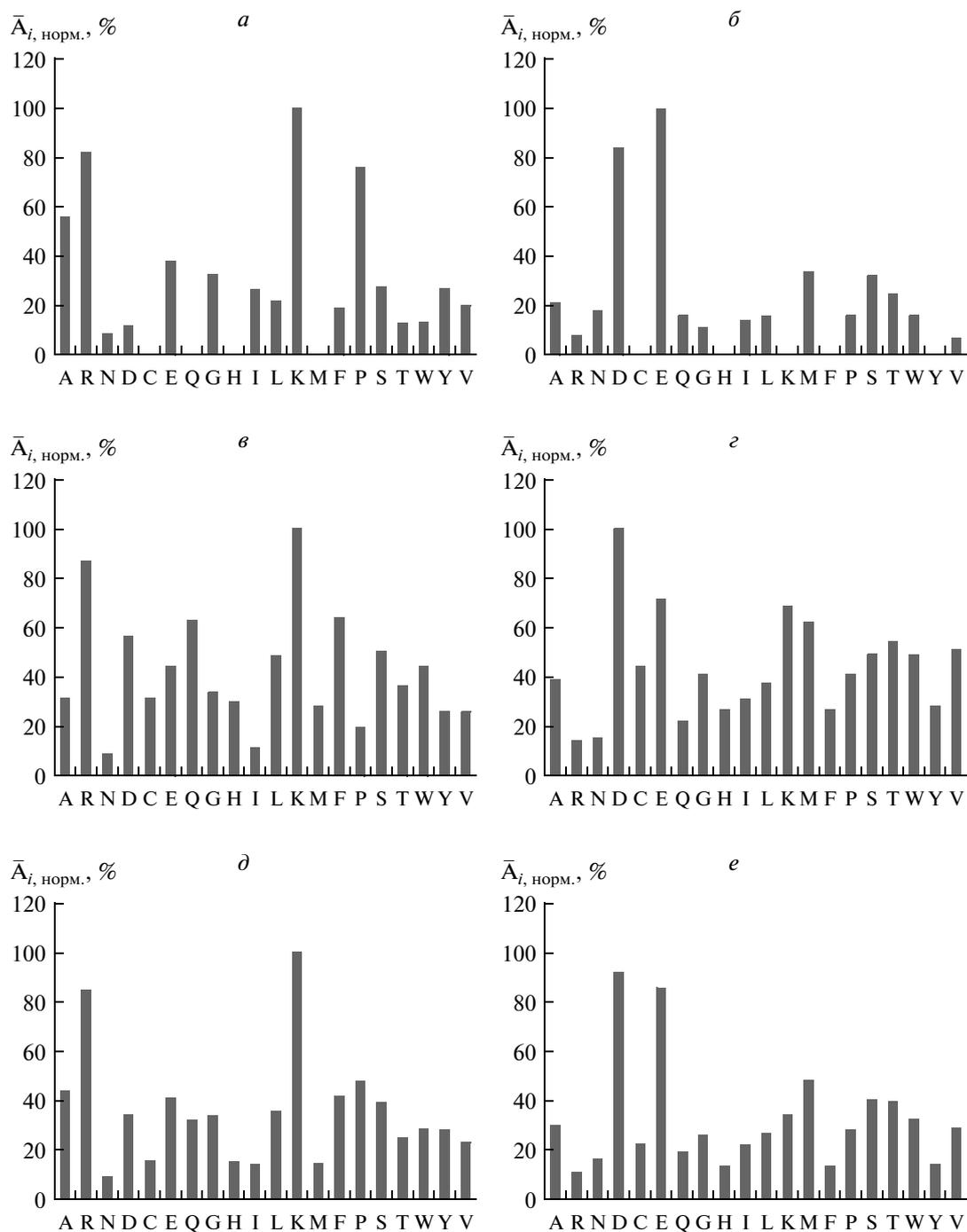
Необходимо отметить следующий факт. Как трипсин, так и глутамилэндопептидаза присутствуют в качестве расщепляющих агентов в программе “Mascot” ([www.matrixscience.com](http://www.matrixscience.com)), используемой для идентификации белков по анализу масс продуктов их гидролиза. Обработка полученных нами масс-спектров в данной программе с использованием заданного нами диапазона разброса масс –  $\pm 1$  Да – во всех случаях привела к однозначной идентификации использованного белкового субстрата. При этом, однако, до 10–15% экспериментальных масс являлись посторонними (то есть не могли быть получены при гидролизе данного белка ферментом данной специфичности – разница между номинальной расчетной и экспериментально

определяемой моноизотопной массами превышает 1 Да). Однако показанные выше результаты статистической обработки масс-спектров однозначно указывают на правильность определения первичной специфичности протеиназы. Это говорит об относительной толерантности предложенного метода к наличию в масс-спектрах продуктов протеолиза пиков посторонних загрязнений.

Суммируя вышесказанное, можно сделать следующие выводы. Узкая первичная субстратная специфичность протеиназ достаточно хорошо может быть определена предложенным нами подходом. Однако, в зависимости от белка-субстрата и наличия пиков посторонних загрязнений в масс-спектрах, после статистической обработки могут иметь достаточно большой вес аминокислоты, не относящиеся к специфичности используемой протеиназы. Таким образом, для получения достоверных результатов необходимо использовать набор белковых субстратов с последующим усреднением получаемых статистических данных. Преимущество предлагаемого нами метода изучения первичной специфичности протеиназ мы видим в том, что он не требует прямого определения аминокислотной последовательности продуктов протеолиза и весьма прост инструментально. Дорогостоящие стадии секвенирования заменяются при этом на компьютерную обработку получаемых МАЛДИ-масс-спектров.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

В работе использовали соли, реактивы для проведения гель-электрофореза и компоненты буферных растворов производства фирмы “Serva” (Германия). Белки-маркеры для гель-электрофореза PageRuler (11–170 кДа) и буфер для нанесения белков на гель производства “Fermentas” (Канада). Куриный лизоцим, трипсин “proteomic grade” и матрицы для масс-спектрометрического анализа: 4-гидрокси- $\alpha$ -цианокоричная кислота (НССА) и 2,5-дигидроксибензойная



Результаты статистической обработки генерированных наборов пептидов для лизоцима (*a*, *b*) и катепсина L (*в*, *г*) по нормированной частоте встречаемости аминокислот ( $\bar{A}_{i, \text{норм.}}$ , %) при обработке МАЛДИ-масс-спектров продуктов их протеолиза трипсином (*a*, *в*, *д*) и глутамилэндопептидазой (*б*, *г*, *е*). (*д*) – Среднее арифметическое соответствующих значений  $\bar{A}_{i, \text{норм.}}$ , %, рис. (*a*) и (*в*); (*е*) – среднее арифметическое соответствующих значений  $\bar{A}_{i, \text{норм.}}$ , %, рис. (*б*) и (*г*).

кислота (DHB) – (Sigma, США). Ацетонитрил, ацетон и трифторуксусная кислота (HPLC-grade) производства “Криохром” (Россия). Катепсин L из *P. camtschaticus* и глутамилэндопептидаза из *Bacillus intermedius* любезно предоставлены Г.Н. Руденской (МГУ, Москва).

Все растворы для масс-спектрометрического анализа готовились на деионизованной воде, полученной на аппарате Millipore (Millipore, США).

Электрофорез в полиакриламидном геле по Лэммли [18] проводили на приборе Mini-Prote-

an®3 (Bio-Rad, США). Белковые субстраты растворяли в воде (2 мг/мл), смешивали с 62 мМ Трис-НСl-буфером рН 6.8, содержащим 2% (w/w) SDS, 10% (v/v) глицерина, 5% (v/v) β-меркаптоэтанол, 0.01% (w/w) бромфенилового синего, в соотношении 1 : 1.2 (v/v) и кипятили 5 мин. Полученный раствор наносили на дорожки геля. Использовали 4% концентрирующий и 12% разделяющий полиакриламидные гели. По окончании электрофореза гель фиксировали в течение 1 ч в системе метанол/уксусная кислота/вода (4 : 1 : 5), окрашивали в течение 30 мин 0.1% (w/w) раствором Ку-масси бриллиантового синего G-250 в 10% (v/v) растворе CH<sub>3</sub>COOH с 0.5% (w/w) сульфата меди, содержащем 27% (v/v) этанола, и отмывали 24 ч в системе метанол/уксусная кислота/вода (2 : 1 : 7).

**Протеолиз белков из геля** проводили по модифицированной методике производителя трипсина для протеомных исследований [19]. Из интересующей белковой полосы геля вырезали скальпелем кусочки 1 × 1 мм и по одному помещали в пробирки.

Для удаления SDS в пробирки добавляли по 150 мкл раствора 40% (v/v) водного этанола и 5% CH<sub>3</sub>COOH (v/v), инкубировали 15 мин при комнатной температуре. После удаления раствора шприцем кусочки геля промывали водой (150 мкл) в течение 5 мин при комнатной температуре. Краситель удаляли с помощью раствора 200 мМ NH<sub>4</sub>HCO<sub>3</sub> в 40% (v/v) ацетонитриле (200 мкл раствора, инкубация 30 мин при 37°C), отмывые от красителя образцы снова промывали водой (150 мкл). После дегидратации геля ацетонитрилом (150 мкл, 10 мин) растворитель удаляли и высушивали образцы в вакуумном эксикаторе.

К сухим кусочкам геля добавляли по 20 мкл раствора трипсина или глутамилэндопептидазы (20 мкг белка на 1 мл раствора 0.1 мМ HCl и 36 мМ NH<sub>4</sub>HCO<sub>3</sub> в 8.1% (v/v) ацетонитриле), инкубировали 10 мин при 5°C, добавляли по 50 мкл 40 мМ раствора NH<sub>4</sub>HCO<sub>3</sub> в 9% (v/v) ацетонитриле и инкубировали реакционную смесь в течение ночи при 37°C. После инкубации отбирали жидкость над гелем и помещали ее в отдельную пробирку. Для дополнительной экстракции пептидов в каждую пробирку добавляли по 50 мкл 0.1% раствора TFA в 50%-ном (v/v) ацетонитриле и инкубировали 30 мин при 37°C. Жидкость после дополнительной экстракции совмещали с жидкостью, полученной на предыдущем шаге. Этот раствор использовали в качестве образцов для МАЛДИ-масс-спектрометрического анализа.

**Приготовление образцов для масс-спектрометрии.** В работе использовали смесь растворов матриц HCCA и DHB в соотношении 1 : 1 по объему. Исходные растворы матриц в концентрации 20 мг/мл готовили в смеси ацетонитрила с 0.1% (v/v) водным раствором TFA (7 : 3). Образцы наносили на мишень

методом высушенной капли. Для этого к 5 мкл образца добавляли 5 мкл раствора матрицы и перемешивали. Далее 0.7–1.5 мкл смеси наносили на стальную плашку-мишень MPT 384 (Bruker Daltonik, Германия) и высушивали на воздухе.

**МАЛДИ-масс-спектры** получали на времяпролетном масс-спектрометре Autoflex II (Bruker Daltonik, Германия) с азотным лазером ( $\lambda = 337$  нм, энергия лазерного импульса 100 мкДж, частота 50 Гц). Все измерения проводили в режиме положительно заряженных ионов, которые вытягивались из области ионизации ускоряющим напряжением 19 кВ. Для калибровки прибора использовали стандартные смеси фирмы "Bruker Daltonik" (Германия). Для получения конечного спектра накапливали 200–300 индивидуальных спектров. Для каждой исследуемой пробы получали три параллельных спектра.

**Обработку спектров** проводили с помощью программы Bruker DataAnalysis, Version 1.6 g ((C) Bruker Daltonik GmbH, 1999). Из набора изотопных пиков молекулярного иона выбирали пики, соответствующие моноизотопной массе [20]. Все масс-спектры очищали от перекрестных загрязнений [21]. Как достоверно значимые интерпретировали пики с соотношением сигнал/шум не менее 10. Набор полученных масс использовали в дальнейшем для работы с базами данных.

**Генерация набора пептидных продуктов протеолиза.** Для каждой экспериментально полученной массы продукта протеолиза с помощью программы FindPept [22] в аминокислотной последовательности белкового субстрата находили пептиды, отвечающие этой массе с ошибкой измерения  $\pm 1$  Да. Аналогичную процедуру проводили для каждой теоретически полученной массы и сгенерированный набор пептидов подвергали статистическому анализу. Аминокислотные последовательности куриного лизоцима (P00698) и катепсина L из *P. camtschaticus* (E5LEX0) были взяты из базы Uniprot [23].

## БЛАГОДАРНОСТИ

Авторы благодарят Г.Н. Руденскую (МГУ, Москва) за любезно предоставленные белки и О.А. Малошицкую (МГУ, Москва) за помощь при снятии масс-спектров.

## СПИСОК ЛИТЕРАТУРЫ

1. Драчевская М.И., Борзенкова А.В., Еремеев Н.Л. // Биоорган. химия. 2012. Т. 38. № 1. С. 111–118.
2. Schechter J., Berger A. // Biochem. and Biophys. Res. Commun. 1967. V. 27. P. 157–162.
3. Roepstorff P. // Proteomics in Functional Genomics: Protein Structure Analysis, Eds Jollès P., Jörnvall H. Basel; Berlin; Boston: Birkhäuser Verlag AG, 2000. P. 81–97.

4. Клёсов А.А. Ферментативный катализ. Том 2. М.: Изд-во Московского университета, 1984. 216 с.
5. Collins P.J., McMahon G., O'Brien P., O'Connor B. // *Int. J. Biochem. & Cell Biology*. 2004. V. 36. P. 2320–2333.
6. Guionie O., Moallic C., Niamker S., Placier G., Sine J.-P., Colas B. // *Comparative Biochemistry and Physiology*. 2003. V. 135. P. 503–510.
7. Зинченко А.А., Руми Л.Д., Антонов В.К. // *Биоорг. химия*. 1976. Т. 2. С. 803–810.
8. Баскова И.Р., Завалова Л.Л. // *Биохимия*. 2007. Т. 72. № 2. С. 270–278.
9. Milstein Celia P., Deverson D.E.V. // *Biochem. J*. 1971. V. 123. P. 945–958.
10. Nolan C., Margoliash E., Peterson J.D., Steiner D.F. // *J. Biol. Chem*. 1971. V. 246. P. 2780–2791.
11. Shilling O., Overall C.M. // *Nature Biotechnol*. 2008. V. 26. P. 585–594.
12. Keller U. auf dem, Schilling O. // *Biochimie*. 2010. V. 92. P.1705–1714.
13. Keller U. auf dem, Schilling O., Overall C. M. // *Methods Mol. Biol*. 2011 V. 753. P. 257–272.
14. Gay S., Binz P.A., Hochstrasser D.F., Appel R.D. // *Electrophoresis*. 1999. V. 20. № 18. P. 3527–3534.
15. Tabb D.L., MacCoss M.J., Wu C.C., Anderson S.D., Yates J.R. // *Anal. Chem*. 2003. V. 75. № 10. P. 2470–2477.
16. Wolski W.E., Farrow M., Emde A., Lehrach H., Lalowski M., Reinert K. // *Proteome Sci*. 2006. V. 4: 18.
17. Keiler K.C., Silber K.R., Sauer R.T., Downard K.M., Pappayannopoulos I.A., Biemann K. // *Protein Sci*. 1995. V. 4. № 8. P. 1507–1515.
18. Laemmli U.K. // *Nature*. 1970. V. 227. P. 680–685.
19. <http://www.sigmaaldrich.com/sigma-aldrich/bulletin/t6567bul.pdf>
20. Yergey J., Heller D., Hansen G.R., Cotter J., Fenselau C. // *Anal. Chem*. 1983. V. 55. P. 353–356.
21. Ding Q., Xiao L., Xiong S., Jia Y., Que H., Guo Y., Liu S. // *Proteomics*. 2003. V. 3. P. 1313–1317.
22. <http://au.expasy.org/tools/findpept.html>
23. <http://www.uniprot.org/>

## The Study of Protease Primary Specificity by Statistical Analysis of MALDI Mass-Spectra of Proteolysis Products

N. L. Ereemeev<sup>#</sup>, A. V. Borzenkova

<sup>#</sup>Phone: +7 (495)–939–34–17, fax: +7–(495)–939–54–17; e-mail: [eremeev@enzyme.chem.msu.ru](mailto:eremeev@enzyme.chem.msu.ru)  
Lomonosov Moscow State University, Faculty of Chemistry, Moscow 119991, Leninskie Gory

Experimental verification for studying of proteolytic enzymes' primary specificity by statistical analysis of MALDI (matrix-assisted laser desorption/ionization) mass spectra of products obtained by protein substrates proteolysis was done by the use of proteinases with known substrate specificity (trypsin and glutamylendopeptidase). Proposed technique not requires direct determination of proteolysis products amino acid sequences, reliably establishes proteinases with a narrow substrate specificity and shows a relative tolerance for the presence in MALDI mass-spectra peaks of contaminants. It was shown that the pseudo-positive results exception requires the use of protein substrates series with the following averaging received statistical data.

*Keywords: proteolytic enzymes, primary specificity, matrix-assisted laser desorption/ionization (MALDI), mass spectrometry*