



УДК 577.112.083.3:573.6:591.147.7

## АНАЛИТИЧЕСКАЯ БИОТЕХНОЛОГИЯ РЕКОМБИНАНТНЫХ ПЕПТИДОВ И БЕЛКОВ. II. КОНТРОЛЬ ПЕРВИЧНОЙ СТРУКТУРЫ ГИБРИДНОГО БЕЛКА, СОДЕРЖАЩЕГО ЧЕЛОВЕЧЕСКИЙ ПРОИНСУЛИН, И ОПТИМИЗАЦИЯ ЕГО ПРОТЕОЛИЗА ТРИПСИНОМ

© 2000 г. Н. В. Сергеев, Н. С. Глухова, И. В. Назимов<sup>#</sup>,  
В. А. Гуляев, И. А. Донецкий, А. И. Мирошников

Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,  
117871, г. Москва, ГСП-7, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 13.01.2000 г. Принята к печати 13.03.2000 г.

С использованием комплекса аналитических микрометодов разделения (микроколоночная оф-ВЭЖХ) и качественной идентификации (масс-спектрометрия, аминокислотное секвенирование) изучена кинетика протеолиза трипсином гибридного белка (ГБ), включающего в свой состав проинсулин человека. Показано, что начальной стадией гидролиза является расщепление ГБ на лидерный фрагмент и проинсулин, который далее подвергается трансформации в ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулин через отщепление C-пептида. Исследовано влияние температуры на выход побочного продукта дез-Thr<sup>B30</sup>-инсулина. Методом пептидного картирования подтверждена структура ГБ и выявлено отсутствие N-концевого метионина в лидерном фрагменте.

*Ключевые слова:* проинсулинсодержащий гибридный белок, анализ; обращенно-фазовая ВЭЖХ; масс-спектрометрия; пептидное картирование; биотехнология аналитическая; инсулин.

### ВВЕДЕНИЕ

В настоящее время биотехнологический метод получения рекомбинантного инсулина человека (РИЧ) через природный предшественник – проинсулин является наиболее экономически эффективным [2]. В рекомбинантных штаммах *E. coli*, проинсулин экспрессируется в составе гибридного белка (ГБ), т.е. молекулы проинсулина, соединенной через остаток аргинина [3] или метионина [4] с лидерным фрагментом. Структура лидерного фрагмента влияет на биосинтез и накопление белка в клетке, что позволяет достичь высокого уровня экспрессии целевого ГБ [5]. Пространственная структура проинсулина обеспечивает предпочтительное замыкание дисульфидных связей в положениях, соответствующих нативной структуре инсулина, поэтому замыкание дисульфидных связей проводят, как правило, после отщепления лидерного фрагмента [4]. В результате триптического расщепления проинсулина с последующей [6] или одновременной обработкой карбоксипептидазой В [7] образуется инсулин, полностью идентичный природному гормону.

В нашей стране создан бактериальный штамм-суперпродуцент ГБ [8], содержащего проинсулин

человека, к N-концу которого присоединен лидерный фрагмент, в состав которого входит IgG-связывающий домен белка А из *Staphylococcus aureus* и пептидный линкер His<sub>6</sub>-Gly-Ser-Arg (рис. 1). Пространственная структура указанного ГБ обеспечивает высокий выход реакции образования правильно замкнутых дисульфидных связей непосредственно в молекуле ГБ, что позволяет исключить стадию получения проинсулина и, таким образом, снизить технологические потери.

При выборе способа дальнейшей трансформации ГБ необходимо учитывать, что при трипсинолизе проинсулина возможно образование до 25% побочного продукта дез-Thr<sup>B30</sup>-инсулина [9], что требует дополнительной стадии очистки и снижает выход целевых полупродуктов ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина и Arg<sup>B31</sup>-инсулина, а при совместном гидролизе проинсулина трипсином и карбоксипептидазой В содержание дез-Thr<sup>B30</sup>-инсулина в получаемом препарате может достигать 6%, что превышает требования фармакопейных статей [10, 11] к содержанию в субстанции инсулина родственных пептидов.

В связи с тем что дез-Thr<sup>B30</sup>-инсулин является одной из наиболее трудноудаляемых примесей из субстанции РИЧ [12], способ двухстадийной трансформации ГБ кажется более предпочтительным. При этом необходимо подобрать оптимальные условия реакции трипсинолиза ГБ, что-

Сообщение I см. [1].

<sup>#</sup> Автор для переписки (e-mail: nazimov@ibch.ru; тел.: (095) 330-75-92).

бы до минимума снизить количество этого побочного продукта.

Известно [9], что скорость расщепления проинсулина по остатку Arg<sup>B32</sup> на порядок выше, чем по Lys<sup>B29</sup>. Именно эти кинетические закономерности обуславливают преимущественное образование ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина при трипсинолизе проинсулина. Дальнейшее расщепление ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина трипсином по сайтам Lys<sup>B29</sup> и Arg<sup>B31</sup> приводит к образованию дез-Thr<sup>B30</sup>-инсулина и Arg<sup>B31</sup>-инсулина соответственно, при этом скорость реакции по сайту Arg<sup>B31</sup> ниже, чем по Lys<sup>B29</sup> [9].

Цель работы – изучить влияние температуры трипсинолиза ГБ на относительный выход указанных продуктов, определить оптимальную температуру и продолжительность реакции, при которых выход побочного продукта дез-Thr<sup>B30</sup>-инсулина был бы минимален.

Для исследования кинетики расщепления ГБ трипсином мы использовали методологию, описанную в нашей предыдущей работе [1]. В ходе исследования предполагалось, также подтвердить первичную структуру ГБ, так как совместное использование обращенно-фазовой ВЭЖХ (офВЭЖХ) и методов качественной идентификации (аминокислотное секвенирование, масс-спектрометрия (МС), аминокислотный анализ) позволяет достоверно идентифицировать пептидные фрагменты гидролизата и по их составу восстановить структуру исходного белка.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Исследование кинетики трипсинолиза гибридного белка

Триптический гидролиз ГБ проводили при трех температурах 10, 25, 37°C и фиксированном соотношении фермент/субстрат (1 : 1000), что позволило исследовать влияние температуры на выход дез-Thr<sup>B30</sup>-инсулина.

Мы показали, что при трипсинолизе ГБ в первую очередь гидролизуется связь Arg<sup>74</sup>-Phe<sup>B1</sup>, что приводит к образованию лидерного фрагмента (рис. 2, пик 2) и проинсулина (рис. 2, пик 8). Протеолиз проинсулина, вероятно, далее протекает по механизму, описанному в работах [6, 9], что подтверждается обнаружением интермедиатов расщепления проинсулина Int I и Int II (рис. 2, пики 5, 6), имеющих одинаковую молекулярную массу (таблица). Пептид, соответствующий пику 5 на рис. 2 представляет собой продукт расщепления проинсулина по связи Arg<sup>106</sup>-Glu<sup>107</sup> (Int I), а интермедиат Int III (рис. 2, пик 6) образуется при гидролизе по связи Arg<sup>139</sup>-Gly<sup>140</sup> [9]. Дальнейшее расщепление интермедиатов приводит к образованию ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина (рис. 2, пик 4) и С-пептида (рис. 2, пик 1). Пик 3 (рис. 2) представ-

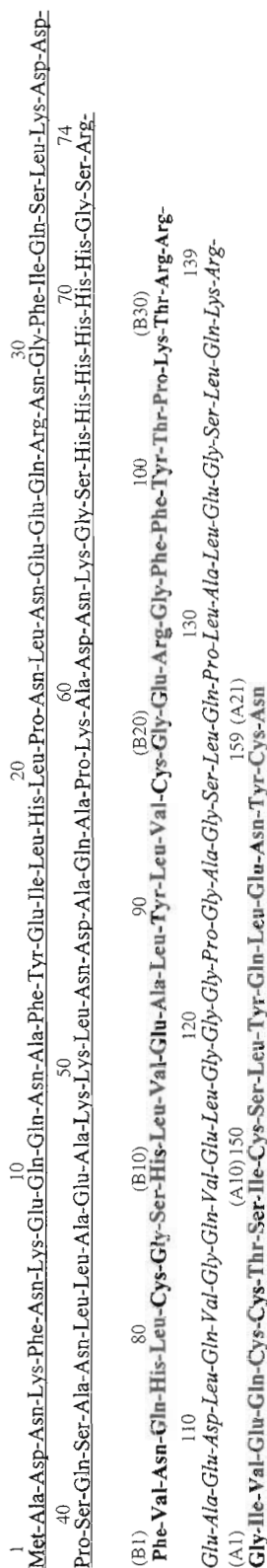
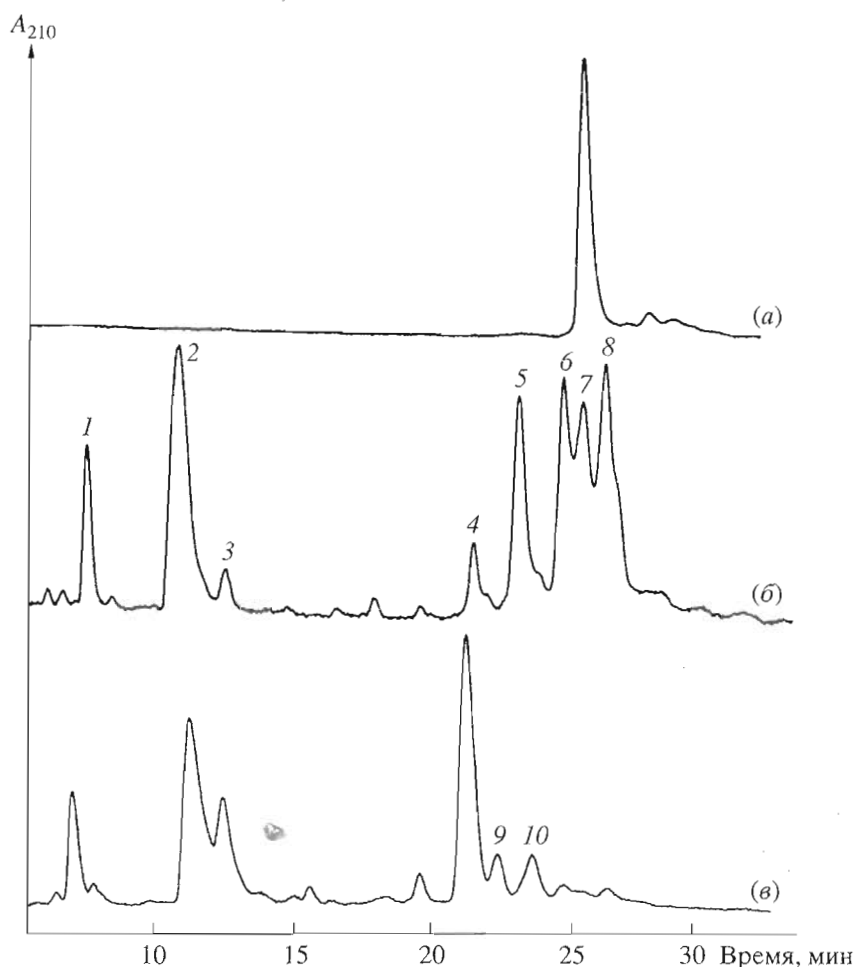


Рис. 1. Аминокислотная последовательность ГБ. Нумерация АК-остатков в индивидуальном инсулине показана в скобках. Подчеркнут лидерный фрагмент. Жирным шрифтом показана АК-последовательность цепей ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина (Phe<sup>75</sup>-Arg<sup>106</sup> – В-цепь, Gly<sup>140</sup>-Asn<sup>160</sup> – А-цепь). Курсивом показан С-пептид.



**Рис. 2.** Анализ продуктов протеолиза трипсином ГБ при 10°C через 0 (а), 1 (б) и 4 ч (в). Хроматограф Милихром А-02; колонка Vydac С4-300 (2 × 75 мм); градиент ацетонитрила (23–32%) в 0.1% TFA за 32 мин; скорость потока 100 мкл/мин; максимальное давление 1.5 МПа; температура 40°C. Идентификацию пиков 1–10 см. в тексте.

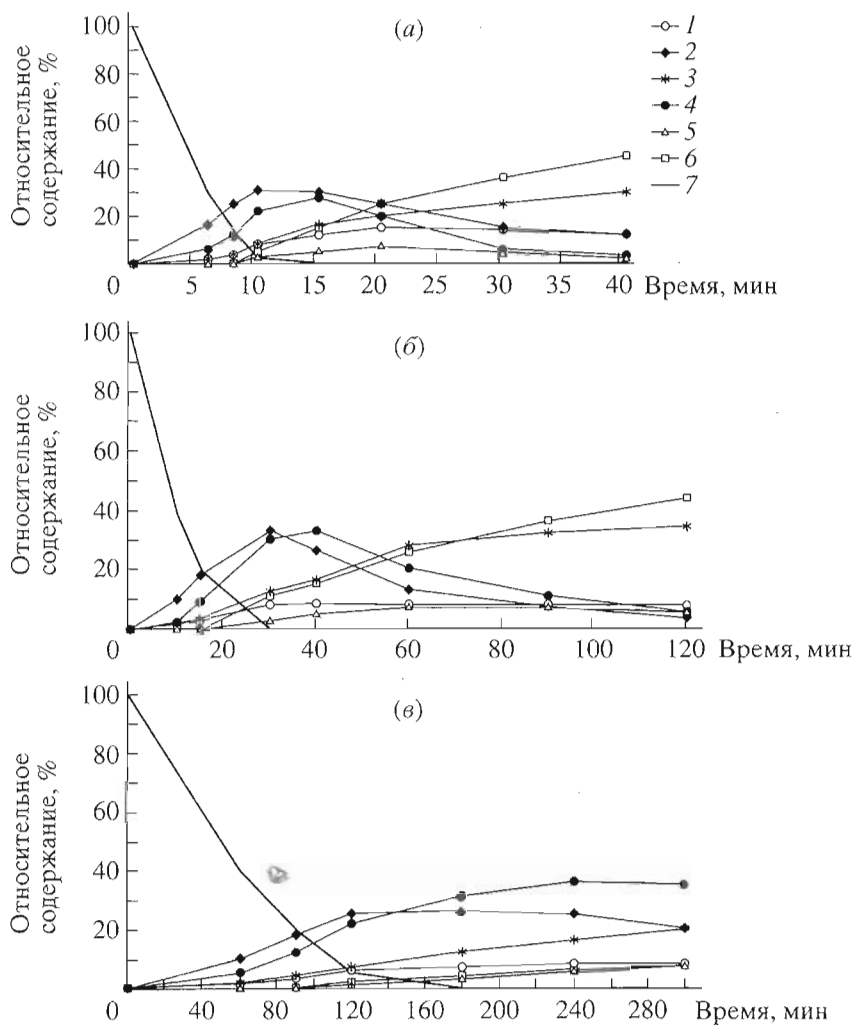
ляет собой продукт протеолиза лидерного белка (ЛФ-1) с  $m/z$  7022.5 (таблица).

Как оказалось, связь Arg<sup>74</sup>-Phe<sup>75</sup> является не единственной, хотя и преимущественно гидролизуемой в момент начала реакции. Анализ кинети-

ческих данных накопления продуктов реакции показал, что скорость образования С-пептида в начале реакции превышает скорость образования ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина. Эти данные позволяют предположить, что механизм реакции трипсиноли-

#### Характеристика пептидных фрагментов, полученных при протеолизе ГБ трипсином

| Номер пика на рис. 2 | Пептид   | Фрагмент в последовательности ГБ   | Молекулярная масса, Да |           |
|----------------------|--|--|------------------------|-----------|
|                      |  |  | по данным МС           | расчетная |
| 1                    | С-KR (С-пептид)                                    | Glu <sup>107</sup> -Arg <sup>139</sup>   | 3302.8                 | 3302.7    |
| 2                    | Лидерный фрагмент                                  | Ala <sup>2</sup> -Arg <sup>74</sup>  | 8290.5                 | 8289.1    |
| 3                    | ЛФ-1   | Ala <sup>2</sup> -Lys <sup>63</sup>  | 7022.5                 | 7021.6    |
| 4                    | ди-Arg <sup>B31</sup> -Arg <sup>B32</sup> -инсулин | Phe <sup>75</sup> -Arg <sup>106</sup> + Gly <sup>140</sup> -Asn <sup>160</sup> | 6121.8                 | 6120.6    |
| 5                    | Int-I  | Phe <sup>75</sup> -Arg <sup>139</sup> + Gly <sup>140</sup> -Asn <sup>160</sup> | 9408.5                 | 9407.4    |
| 6                    | Int II   | Phe <sup>75</sup> -Arg <sup>106</sup> + Glu <sup>107</sup> -Asn <sup>160</sup> | 9408.5                 | 9407.7    |
| 7                    | ГБ   | Ala <sup>2</sup> -Asn <sup>160</sup>   | 17663.6                | 17784.6   |
| 8                    | Проинсулин   | Phe <sup>75</sup> -Asn <sup>160</sup>  | 9390.6                 | 9389.4    |
| 9                    | Arg <sup>B31</sup> -инсулин                        | Phe <sup>75</sup> -Arg <sup>105</sup> + Gly <sup>140</sup> -Asn <sup>160</sup> | 5963.1                 | 5963.9    |
| 10                   | дез-Thr <sup>B30</sup> -инсулин                    | Phe <sup>75</sup> -Lys <sup>103</sup> + Gly <sup>140</sup> -Asn <sup>160</sup> | 5707.1                 | 5707.5    |



**Рис. 3.** Кинетика протеолиза ГБ трипсином при 37 (а), 25 (б) и 10°С (в). Показано относительное содержание в реакционной смеси С-пептида (1), лидерного фрагмента (2), фрагмента ЛФ-1 (3), ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина (4), Arg<sup>B31</sup>-инсулина (5), дез-Thr<sup>B30</sup>-инсулина (6), ГБ (7).

за ГБ включает, по-видимому, процесс отщепления С-пептида непосредственно из молекулы ГБ.

В пользу последнего утверждения также свидетельствует факт, что фракция, соответствующая пику 5 (рис. 2), содержит молекулярный ион с *m/z* 14404.9, и, вероятно, представляет собой ГБ с отщепленным С-пептидом.

Кинетика трипсинолиза ГБ при разных температурах показана на рис. 3. Как оказалось, варьируя температуру гидролиза, можно существенно изменить количество дез-Thr<sup>B30</sup>-инсулина, образующегося за время полной конверсии ГБ и его интермедиатов (рис. 2, пики 5, 6, 8). При 37°С соотношение ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулин : Arg<sup>B31</sup>-инсулин : дез-Thr<sup>B30</sup>-инсулин на момент окончания реакции (15 мин) составляет 55 : 5 : 45, что соответствует высокому выходу (45%) побочного продукта дез-Thr<sup>B30</sup>-инсулина. При 25°С указанное

соотношение составляет 67 : 10 : 23 (40 мин), а при 10°С – 75 : 12 : 13 (4 ч), что соответствует выходу дез-Thr<sup>B30</sup>-инсулина 23 и 13% соответственно.

Снижение выхода дез-Thr<sup>B30</sup>-инсулина при уменьшении температуры трипсинолиза объясняется, по-видимому, большей зависимостью от температуры скорости реакции образования дез-Thr<sup>B30</sup>-инсулина, чем расщепления ГБ и его интермедиатов, т.е. при пониженной температуре (10°С) скорость реакции образования ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина выше, чем скорость его превращения в побочный продукт.

#### *Определение подлинности первичной структуры гибридного белка*

Для подтверждения первичной структуры ГБ требовалось убедиться, что в структуре ГБ нет аминокислотных модификаций или замен ни в ли-

дерном фрагменте, ни, что особенно важно, в структуре инсулина. Методом масс-спектрометрии было показано, что молекулярный ион ГБ имеет  $m/z$  17633 Да, что не соответствует расчетной молекулярной массе. Однако при ширине пика молекулярного иона 800 Да невозможно оценить ни точную молекулярную массу белка, ни предположить характер аминокислотной замены. Подтверждение подлинности первичной структуры ГБ было проведено с помощью МС-анализа пептидных фрагментов, полученных в результате трипсинолиза ГБ. Заслуживает внимания тот факт, что трипсин, селективно отщепляющий С-пептид и лидерный фрагмент (соединенный через остаток аргинина с N-концом В-цепи), идеально подходит для пептидного картирования ГБ, т.к. позволяет получить информацию о структуре как лидерного фрагмента, так и проинсулиновой части молекулы ГБ.

Результаты МС-анализа пептидных фрагментов ГБ представлены в таблице. Молекулярный ион С-пептида содержащего на С-конце последовательность Lys-Arg (С-KR), (рис. 2, пик 1) имеет  $m/z$  3302.8, что практически совпадает с расчетным значением молекулярной массы. Было найдено, что в условиях избыточного протеолиза (37°C, 3 ч) С-пептид подвергается расщеплению трипсином по остатку Lys, что приводит к образованию укороченного С-пептида (С-K) с  $m/z$  3146.2 (данные не показаны).

Лидерный фрагмент (рис. 2, пик 2), также подвергается расщеплению трипсином, так как содержит 7 остатков лизина и 2 остатка аргинина (рис. 1). Было установлено, что гидролиз лидерного фрагмента идет ступенчато, начиная с С-концевой части молекулы, приводя к образованию нескольких пептидных фрагментов (данные не показаны). Как оказалось, найденные молекулярные массы для лидерного фрагмента и продуктов его протеолиза не совпадают с расчетными. При этом разница составляет примерно 131 Да для всех пептидов, что может означать отсутствие остатка метионина в лидерном белке. Анализом пяти N-концевых аминокислотных остатков ГБ, лидерного фрагмента и фрагмента ЛФ-1 (рис. 2, пик 3) было подтверждено, что в лидерном белке и, соответственно, в его фрагментах действительно отсутствует N-концевой метионин.

С учетом отсутствия метионина найденные молекулярные массы для указанных фрагментов находятся в хорошем соответствии с рассчитанными, что подтверждает отсутствие других модификаций первичной структуры лидерного белка.

Как видно из приведенных данных МС-анализа (таблица), молекулярные массы ди-Arg<sup>B31</sup>-Arg<sup>B32</sup>-инсулина (рис. 2, пик 4), Arg<sup>B31</sup>-инсулина (рис. 2, пик 9) и дез-Thr<sup>B30</sup>-инсулина (рис. 2, пик 10)

также совпадают с рассчитанными, что подтверждает отсутствие аминокислотных модификаций или замен в аминокислотной последовательности инсулина.

Таким образом, в результате проделанной работы были выбраны оптимальные температура и продолжительность трипсинолиза ГБ. В результате изучения кинетики установлен механизм протекания реакции, выделены и с помощью МС-анализа охарактеризованы все основные продукты протеолиза ГБ. С помощью комплекса инструментальных аналитических микрометодов подтверждена первичная структура ГБ и было обнаружено отсутствие N-концевого метионина в аминокислотной последовательности лидерного фрагмента.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Протеолиз гибридного белка.** К раствору 0.5 мг ГБ в 500 мкл буфера (0.1 М Трис, pH 8.1) добавляли 5 мкл раствора трипсина (ICN, США) в том же буфере (соотношение фермент/субстрат от 1 : 1000, по весу), термостатировали при 10–37°C. Аликвоты (20 мкл) отбирали через разные промежутки времени. Реакцию останавливали добавлением 10% трифторуксусной кислоты (TFA) до концентрации 0.2% TFA в образце. Гидролизат анализировали методом оФВЭЖХ как указано далее.

**Микроколоночная оФВЭЖХ.** Образцы гидролизата ГБ центрифугировали (5 мин, 12000 об/мин, микроцентрифуга Eppendorf 5410 (Eppendorf, ФРГ)). Надосадочную жидкость помещали в пробирку автосамплера хроматографа Милихром А-02 (Эконова, Новосибирск, Россия). Для анализа использовали 5 мкл раствора. Разделение проводили на колонке Vydac C<sub>4</sub> (2 × 75 мм), заполненной сорбентом (300 Å, 5 мкм, Separation Group, США) в градиенте концентрации ацетонитрила (сорт 1, "Криохром", Санкт-Петербург, Россия) от 23 до 32% за 32 мин в 0.1% трифторуксусной кислоте (филиал ГИПХ, Пермь, Россия). Скорость потока 100 мкл/мин. Температура колонки 40°C. Детекция 210 нм.

Для измерения времен элюирования и расчета площадей пиков использовали программу "Мультихром для Windows" (Амперсенд, Москва, Россия).

Фрагменты идентифицировали по значениям  $m/z$  масс-спектров собранных фракций.

**Масс-спектрометрический анализ.** Фракции, содержащие изучаемые пептиды, анализировали методом лазерного возбуждения молекул вещества на поверхности мишени с последующей время-пролетной регистрацией молекулярных ионов ионизированных веществ (метод MALDI-TOF-MS).

К 1 мкл элюата добавляли 1 мкл раствора матрицы (2,5-дигидробензойная кислота в 20% водном ацетонитриле, концентрация – 20 мг/мл) и высушивали на мишени.

Масс-спектры снимали в режиме положительных ионов с использованием УФ-лазера ( $\lambda$  337 нм, импульс 3 нс, энергия в импульсе 120 мкДж) на приборе Vision 2000 (Finnigan-Thermo Bioanalysis, Англия).

### СПИСОК ЛИТЕРАТУРЫ

1. Сергеев Н.В., Назимов И.В., Гавриков В.Г., Мирошников А.И. // Биоорганическая химия. 2000. Т. 26. С. 41–50.
2. Brange J. Galenics of Insulin. Berlin–Heidelberg: Springer-Verlag, 1987. P. 1–5.
3. Nilsson J., Johansson P., Samuelson E., Stahl S., Uhlen M. // J. Biotechnol. 1996. V. 48. P. 241–250.
4. Klyushnichenko V.E., Yakimov S.A., Nazimov I.V., Arutyunyan A.M., Ivanov A.E., Maltsev K.V., Wulfson A.N. // J. Chromatogr. B. Biomed. Appl. 1994. V. 662. P. 363–369.
5. Kang Y., Yoon J.W. // J. Biotechnol. 1994. V. 36. P. 45–54.
6. Howard W., Davidson J., Hutton C. // J. Biochem. 1987. V. 245. P. 575–582.
7. Jonasson P., Nilsson J., Samuelson E., Moks T., Stahl S., Uhlen M. // Eur. J. Biochem. 1988. V. 236. P. 656–661.
8. Коробко В.Г., Болдырева Е.Ф., Шингарова Л.Н., Мирошников А.И., Гавриков В.Г., Степанов А.В., Калинин Ю.Т. // Патент РФ № 99109523/13. 1999. Бюлл. № 36.
9. Миргородская О.А., Казанина Г.А., Миргородская Е.П., Шевченко А.А., Мальцев К.В., Мирошников А.И., Ройнсторфф П. // Биоорганическая химия. 1997. Т. 23. С. 91–97.
10. United States Pharmacopeia, 23, Second Supplement. Insulin Human. 1984. P. 2637–2645.
11. British Pharmacopeia 1988, H.M. Stationary Office, London, 1988. P. 312–313.
12. Health W.F., Belagaje R.M., Brooke G.S., Chance R.E., Hoffmann J.A., Long H.B., Reams S.G., Roundtree C., Shaw W.N., Slicker L.J., Sundell K.L., DiMarchi R.D. // J. Biol. Chem. 1992. V. 267. P. 419–425.

## Analytical Biotechnology of Recombinant Peptides and Proteins. II. A Confirmation of the Primary Structure of Fusion Protein Containing Human Proinsulin and Optimization of Its Proteolysis by Trypsin

N. V. Sergeev, N. S. Glukhova, I. V. Nazimov<sup>#</sup>, V. A. Gulyaev,  
I. A. Donetskii, and A. I. Miroshnikov

*Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,  
ul. Miklukho-Maklaya 16/10, GSP-7 Moscow, 117871 Russia*

The kinetics of trypsin proteolysis of the fusion protein (FP) containing human proinsulin was studied by a set of analytical micromethods. These were the microcolumn reversed-phase HPLC and the qualitative identification by MALDI-TOF mass spectrometry and amino acid sequencing. The first stage of the proteolysis was shown to be the cleavage of FP into the leader fragment and proinsulin. The subsequent splitting off of C-peptide from proinsulin results in the formation of Arg<sup>B31</sup>-Arg<sup>B32</sup>-insulin. The effect of temperature on the formation of de-Thr<sup>B30</sup>-insulin, a by-product, was also studied. The structure of FP was confirmed by the peptide mapping technique, and the leader fragment was shown to contain no N-terminal Met residue.

*Key words: analytical biotechnology; fusion protein, analysis; MALDI-TOF mass spectrometry; peptide mapping; recombinant proteins, analysis; reversed-phase HPLC*

<sup>#</sup> To whom correspondence should be addressed; phone: +7 (095) 330-7592;  
e-mail: nazimov@ibch.ru.

For the full text of English translation of this article, see *Russian Journal of Bioorganic Chemistry*, 2000, vol. 26, no. 7. It is also available (free) in PDF format through the Web Site <http://www.maik.rssi.ru/>.