

УДК 577.322

## ОПРЕДЕЛЕНИЕ АМИЛОИДОГЕННЫХ УЧАСТКОВ, ВХОДЯЩИХ В ОСТОВ ФИБРИЛЛ ИНСУЛИНА

© 2019 А.К. Сурин<sup>1,2,3</sup>, С.Ю. Гришин<sup>1,4</sup>, О.В. Галзитская<sup>1\*</sup>

<sup>1</sup> Институт белка РАН, 142290 Пущино, Россия; электронная почта: ogalzit@vega.protres.ru

<sup>2</sup> Государственный научный центр прикладной микробиологии и биотехнологии, 142279 Оболensk, Россия; электронная почта: alan@vega.protres.ru

<sup>3</sup> Национальный исследовательский центр эпидемиологии и микробиологии им. Н.Ф. Гамалеи Минздрава России, 123098 Москва, Россия

<sup>4</sup> Московский государственный университет им. М.В. Ломоносова, 119991 Москва, Россия; электронная почта: syugrishin@gmail.com

Поступила в редакцию 15.06.18

После доработки 03.09.18

Принята к публикации 03.09.18

Для обнаружения конформационных переходов, приводящих к образованию фибрилл инсулина, требуется определение амилоидогенных участков в структуре мономеров белка. Ранее были предложены различные модели фибриллогенеза инсулина, однако конкретные участки, отвечающие за формирование амилоидных фибрилл, не были определены. В настоящей работе с помощью биоинформатических программ предсказания амилоидогенных участков мы выявили несколько общих амилоидогенных последовательностей в структуре мономеров инсулина. Методами ограниченного протеолиза и масс-спектрометрического определения фрагментов белковой цепи, устойчивых к действию протеаз, мы обнаружили аминокислотные последовательности в структуре инсулина, предположительно входящие в остов фибрилл. Полученные результаты не противоречат ранее предложенной нами модели формирования фибрилл из кольцевых олигомеров и могут быть использованы для разработки аналогов инсулина, устойчивых к амилоидогенезу.

**КЛЮЧЕВЫЕ СЛОВА:** амилоид, амилоидогенные участки, FoldAmyloid, инсулин, кольцевые олигомеры, масс-спектрометрия, протеолиз белков.

**DOI:** 10.1134/S0320972519010111

Процесс формирования амилоидов, как специфически структурированных фибриллярных агрегатов белков, возможен за счет наличия в аминокислотной последовательности так называемых амилоидогенных участков. Известно, что при изменении химических и физических условий окружающей среды некоторые белки могут со временем приобретать ненативную конформацию, что приводит к накоплению амилоидов и проявлению связанных с ними заболеваний [1, 2]. Основную вклад в амилоидогенез белка вносят аминокислотные остатки (а.о.), которые способствуют более плотной упаковке белковой структуры [3]. Предсказание амилоидогенных участков с последующим их экспериментальным определением — важный этап на пути понимания особенностей амилоидообразования конкретного белка.

Успешное определение амилоидогенных участков может быть проведено для хорошо изу-

Принятые сокращения: а.о. — аминокислотный остаток; ЖХ-МС — жидкостная хроматография и масс-спектрометрия;  $M_r$  — молекулярная масса;  $m/z$  — отношение массы иона к его заряду.

\* Адресат для корреспонденции.

ченных белков, у которых известна структура мономера, а также подтверждена возможность образования амилоидных фибрилл [4]. Кроме того, важную информацию о влиянии отдельных а.о. на процесс образования амилоидов получают путем замены соответствующих а.о. в структуре белка дикого типа и последующего анализа полученных мутантов [5].

Инсулин, как всесторонне исследованный, образующий полиморфные агрегаты амилоидогенный белок, может служить модельным объектом для изучения амилоидогенеза. Известно, что для инсулина, первого среди белков, определили полную аминокислотную последовательность [6, 7], а некоторые особенности формирования фибрилл изучали еще с начала 1940-х гг. [8].

В растворе инсулин находится в основном в виде смеси мономеров, димеров и гексамеров, состав которой изменяется в зависимости от рН раствора и концентрации белка [9]. Мономерный инсулин имеет преимущественно  $\alpha$ -спиральную структуру, образованную двумя полипептидными А- и В-цепями, соединенными между собой дисульфидными (S—S) связями через остатки цистеина (в положениях А7—В7 и

A20–B19). В структуре А-цепи третья S–S-связь расположена между а.о. A6 и A11 [10]. А-Цепь состоит из 21 а.о., имеет две  $\alpha$ -спиральные области: одну на N-конце (A1–A8), другую – на C-конце (A12–A19). В-Цепь состоит из 30 а.о., имеет центральную  $\alpha$ -спираль (B9–B19). Третичная структура этого глобулярного белка образована из обозначенных трех  $\alpha$ -спиралей, сложенных напротив N- и C-концевых сегментов В-цепи.

После биосинтеза и секреции *in vivo* [11] инсулин хранится в виде гексамера, координированного ионами цинка ( $Zn^{2+}$ ). При формировании димера C-концевые участки образуют антипараллельный  $\beta$ -лист, в структуру которого входят а.о. B8, B9, B12, B13, B16, B20, B21, B23–B28 [12]. Гексамер с двумя ионами  $Zn^{2+}$  имеет тройную ось симметрии и состоит из трех димеров молекул инсулина, попарно объединяющих мономеры инсулина. Ионы цинка связываются с остатками His10 В-цепи – каждый цинк с тремя гистидинами, а также с молекулами воды [13]. Известно, что ионы  $Zn^{2+}$  замедляют процесс фибриллогенеза инсулина [14], и в фармакологических составах инсулин обычно представлен в виде гексамера, координированного ионами цинка. Однако для быстрой диффузии в инсулиновых помпах белок должен находиться в виде свободных мономеров. Склонность мономеров инсулина при длительном хранении в растворе при температуре выше комнатной к формированию амилоидных фибрилл служит основным препятствием для использования пептида в инсулиновой помпе [15].

Выявление механизмов образования амилоидных фибрилл позволит в перспективе разработать способы управления процессом фибриллогенеза. Стратегия исследования фибриллогенеза основана на создании упрощенных моделей формирования фибрилл *in vitro*. Nielsen et al. построили модель образования фибрилл инсулина, согласно которой росту амилоидов предшествует диссоциация димеров/гексамеров инсулина до мономеров, образующих в дальнейшем промежуточную структуру, соответствующую ядру нуклеации фибриллы [16]. Ahmad et al. показали, что первой стадией фибриллогенеза инсулина является переход от нативной к частично развернутой, промежуточной конформации мономера [17]. Vestergaard et al. предположили, что основной механизм удлинения фибрилл инсулина состоит в добавлении олигомеров к растущей фибрилле [18].

В начале 2000-х гг. структура фибрилл инсулина была смоделирована на основе данных криоэлектронной микроскопии [19]. В последнее время на основе представлений о наличии

общего структурного элемента в амилоидных фибриллах были предложены модели фибрилл инсулина, в которых повторяющейся структурной единицей являются  $\beta$ -арки (beta arch) [20] либо кольцевые олигомеры [11, 21]. Недавно нашей группой были получены данные о фибриллообразовании безцинковых препаратов инсулина и его аналога лизпро и предложена упрощенная схема формирования полиморфных фибрилл [21].

Существование различных взглядов на структуру амилоидов инсулина может быть объяснено наличием полиморфизма фибрилл и отсутствием понимания конкретных стадий фибриллогенеза инсулина. В этой связи представляет интерес обнаружение в аминокислотной последовательности мономера инсулина амилоидогенных фрагментов, ответственных за формирование остова фибриллы. Для достижения этой цели могут быть использованы биоинформатические подходы предсказания амилоидогенных участков белковой цепи в сочетании с их экспериментальным определением. Для амилоидов характерна высокая устойчивость к химическим и физическим воздействиям, таким как изменение температуры и давления [22], а также к обработке протеазами [23–25]. Последнее свойство может быть использовано для масс-спектрометрического определения амилоидогенных участков в составе фибриллы инсулина с помощью метода, разработанного нами ранее для такого типа задач [26].

## МЕТОДЫ ИССЛЕДОВАНИЯ

**Подготовка образцов инсулина.** Безцинковый препарат инсулина был любезно предоставлен сотрудниками АО «НПК Биоран», Россия. Для приготовления всех растворов использовали деионизованную воду Milli-Q. Конечные растворы получали непосредственно перед экспериментом или из стоковых растворов, хранившихся при температуре +2–4 °С или –20 °С в соответствии с требуемыми условиями. Все препараты инсулина, если не указано иное, были приготовлены в буферном растворе: 140 мМ NaCl, 20% (v/v) уксусной кислоты (pH 2,0) [27]. Буферный раствор, содержащий 100 мМ  $NH_4HCO_3$  (pH 7,5), использовали для оптимальной работы протеаз – трипсина и химотрипсина («Sigma-Aldrich», Германия); для протеиназы К («AppliChem», США) в его состав включали 5 мМ  $CaCl_2$ . Значения pH измеряли с помощью pH-метра Mettler Toledo MP220 («Mettler Toledo», США). Оптическую плотность препаратов инсулина определяли с использованием спектрофотометра Cary 100

(«Varian», США). Концентрацию инсулина рассчитывали с учетом коэффициента экстинкции  $\epsilon_{276\text{ нм}} = 1,0$  ( $\text{мл} \cdot \text{мг}^{-1} \cdot \text{см}^{-1}$ ) [28]. Наличие амилоидных фибрилл в исследуемых образцах контролировали с помощью электронной микроскопии, как описано ранее [27].

**Биоинформатические подходы для определения амилоидогенных участков белковой цепи.** Для поиска амилоидогенных участков в инсулине мы исследовали последовательности А- и В-цепей, представленные в FASTA-формате в открытой базе данных Uniprot ([www.uniprot.org/uniprot/P01308](http://www.uniprot.org/uniprot/P01308)). Для предсказания амилоидогенных последовательностей инсулина нами были использованы пять программ, предназначенных для поиска амилоидогенных участков в белковой цепи: FoldAmyloid, ArchCandy, PASTA2.0, Waltz и AGGRESCAN (табл. 1).

Для теоретического определения амилоидогенных участков в структуре инсулина нами были применены стандартные настройки программ, рекомендованные разработчиками и доступные при открытии программ с помощью соответствующих онлайн-сервисов. Используемые программы отличаются по принципам обнаружения амилоидогенной последовательности в белковой цепи [3, 29–34]. На примере инсулина представляет интерес сравнение между собой результатов биоинформатического предсказания амилоидогенных участков и сопоставление теоретических данных с последовательностями остова фибрилл, полученными методом ограниченного протеолиза амилоидов.

**Протеолитический гидролиз белков.** Известно, что остовы амилоидных фибрилл устойчивы к действию протеаз [35, 36]. Ферментативный гидролиз фибрилл инсулина смесью протеаз может позволить выявить участки молекул, вовлеченные в формирование остова амилоидной фибриллы, защищенные от протеолиза, в то время как остальная часть полипептидной цепи

будет гидролизована. В настоящем исследовании мы применяли три протеазы – трипсин, химотрипсин и протеиназу К.

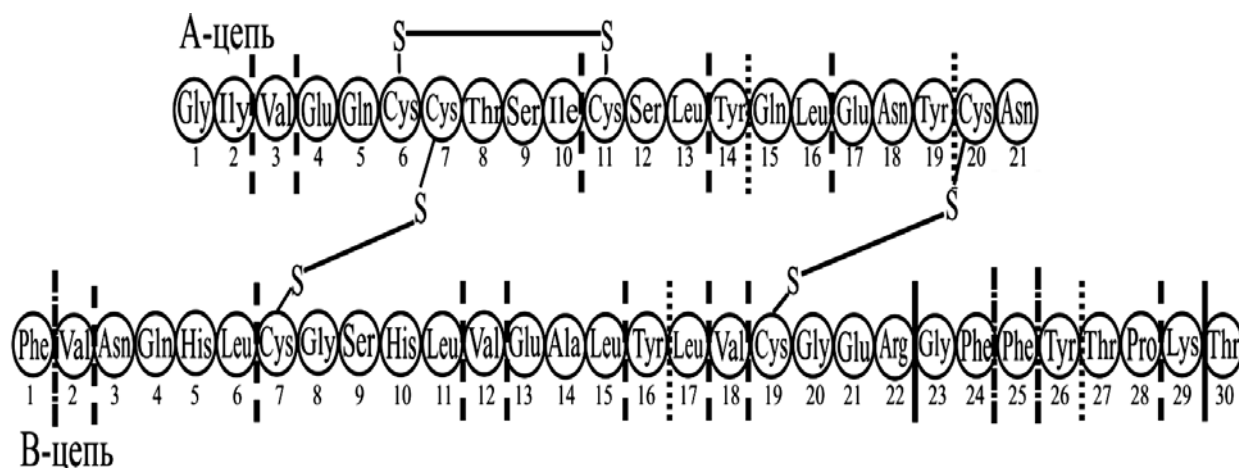
Использованные протеазы характеризуются специфичной (трипсин и химотрипсин) или преимущественной активностью по отношению к пептидным связям, в образовании которых участвуют те или иные а.о. (рис. 1).

Трипсин катализирует гидролиз пептидных связей, расположенных после конкретных а.о. – Lys или Arg [37]. Химотрипсин специфично катализирует разрыв пептидных связей, расположенных после ароматических а.о. (Phe, Tyr, Trp) [38]. Из трех протеаз только протеиназа К не является специфичной – фермент расщепляет белки преимущественно после гидрофобных а.о. (Val, Leu, Ile, Pro, Met, Phe) [39]. Белки полностью расщепляются при длительной инкубации или при высокой концентрации протеиназы К [40].

Препараты трипсина, химотрипсина и протеиназы К в равных долях растворяли в буфере, содержащем 100 мМ  $\text{NH}_4\text{HCO}_3$  (рН 7,5) и 5 мМ  $\text{CaCl}_2$ , до концентраций 0,0267 мг/мл. Препарат инсулина растворяли в буфере, содержащем 140 мМ  $\text{NaCl}$  и 20% (v/v) уксусной кислоты (рН 2,0), до концентрации 2 мг/мл. Подготовленный раствор инсулина инкубировали при 37 °С в течение 48 ч, полученные фибриллы осаждали центрифугированием в течение 20 мин при 10 000 об./мин на центрифуге Eppendorf 5418R («Eppendorf», Германия) и дважды промывали в ~1 мл 100 мМ  $\text{NH}_4\text{HCO}_3$  (рН 7,5). Осадок нерастворимых фибрилл ресуспендировали в 100 мМ  $\text{NH}_4\text{HCO}_3$  (рН 7,5) с добавлением 5 мМ  $\text{CaCl}_2$ , содержащего смесь протеаз (соотношение протеазы к инсулину 1/25 (m/m)), и инкубировали в течение 8 ч при 37 °С со скоростью перемешивания 400 об/мин в термостате Thermomixer Comfort («Eppendorf», Германия). Реакцию протеолиза останавливали добавлением в раствор 0,02% (v/v)

**Таблица 1.** Характеристика использованных программ для определения амилоидогенных участков

Программа	Описание	Ссылка на интернет-ресурс, условия предоставления доступа
FoldAmyloid	использует ожидаемую плотность упаковки а.о. и вероятность образования водородных связей [3, 29]	<a href="http://bioinfo.protres.ru/fold-amyloid/">http://bioinfo.protres.ru/fold-amyloid/</a> , свободный доступ
ArchCandy	исследует вероятность образования различных конформаций $\beta$ -арок [30]	<a href="mailto:andrey.kajava@crbm.cnrs.fr">andrey.kajava@crbm.cnrs.fr</a> , доступ предоставляется бесплатно по запросу
PASTA2.0	сравнивает известные $\beta$ -структуры с последовательностью а.о. [31, 32]	<a href="http://protein.bio.unipd.it/pasta2/">http://protein.bio.unipd.it/pasta2/</a> , свободный доступ
Waltz	отличает настоящие амилоиды от других агрегатов [33]	<a href="http://waltz.switchlab.org/index.cgi">http://waltz.switchlab.org/index.cgi</a> , свободный доступ
AGGRESCAN	выясняет склонность к агрегации отдельного а.о. в белковой последовательности [34]	<a href="http://bioinf.uab.es/aggrescan/">http://bioinf.uab.es/aggrescan/</a> , свободный доступ



**Рис. 1.** Схема, иллюстрирующая на примере мономера инсулина потенциальные сайты гидролиза пептидных связей тремя протеазами: сплошной линией обозначены связи, специфично гидролизуемые трипсином, пунктирной линией – химотрипсином, штриховыми линиями – протеиназой К, а штрих-пунктирными линиями – пептидные связи, гидролизуемые как химотрипсином, так и протеиназой К

концентрированной муравьиной кислоты с последующим осаждением в течение 20 мин при 10 000 об./мин на центрифуге Eppendorf 5418R, осадки дважды промывали ~1 мл 100 мМ  $\text{NH}_4\text{HCO}_3$  (рН 7,5). Полученные осадки растворяли в ~0,2 мл концентрированной муравьиной кислоты и высушивали с помощью вакуумного концентратора Eppendorf 5301 («Eppendorf», Германия).

Подготовленные таким образом гидролизаты затем использовали для обращенно-фазовой жидкостной хроматографии и масс-спектрометрического анализа (ЖХ-МС).

**Жидкостная хроматография и масс-спектрометрия (ЖХ-МС).** С целью очистки образцов, полученных после ферментативного гидролиза, от ассоциатов гидролизатов фибрилл мы использовали фильтрацию на хроматографических колонках, набитых в лабораторных условиях обращенной фазой C18 с размерами 5 мкм и порами 100 Å («Phenomenex», США). Разделение белковых гидролизатов проводили на комбинированной колонке, состоящей из концентрирующей предколонки Acclaim PepMap 100 75 мкм × 2 см с фазой C18 3 мкм 100 Å и разделяющей колонки Acclaim PepMap RSLC 50 мкм × 15 см с фазой C18 2 мкм 100 Å («Thermo Scientific», США) на нанопотоковом хроматографе Easy nLC 1000 («Thermo Scientific», США). В качестве детектора использовали масс-спектрометр высокого разрешения с орбитальной ловушкой ионов Orbitrap Elite («Thermo Scientific», Германия). Панорамные спектры регистрировали в диапазоне 300–2000 m/z с разрешением 240 000, фрагментацию ионов проводили в камере HCD, спектры фрагментации записывали при разрешении 60 000. Масс-спектры гидролизатов образцов, выходящих с колонки с интервалом 1,2 с, анализировали по данным

масс-хроматограммы. Статистический анализ и обработку полученной информации проводили с помощью программ Xcalibur («Thermo Scientific», США) и PEAKS Studio 7.5 («Bioinformatics Solution Inc.», Канада). Пептиды, для которых автоматически измеренная интенсивность ионного тока превышала  $10^6$ , считали достоверно идентифицированными в соответствии с параметрами программы PEAKS Studio 7.5 [26].

Далее последовательности А- и В-цепей инсулина, устойчивые к действию протеаз и определенные в ходе ЖХ-МС, сопоставляли с амилоидогенными фрагментами, теоретически предсказанными программами FoldAmyloid, Waltz и AGGRESKAN.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

**Предсказание амилоидогенных участков аминокислотной последовательности инсулина с помощью биоинформатических программ.** Перечисленные в табл. 1 биоинформатические инструменты, как видно из их описания, различаются используемыми подходами обнаружения амилоидогенной последовательности. Можно предположить, что результаты определения амилоидогенных фрагментов в А- и В-цепях инсулина, полученные с использованием разных программ, будут отличаться.

Как было рассмотрено выше (рис. 1), структура инсулина представляет собой две пептидные цепи, соединенные между собой двумя дисульфидными связями. В связи с двухцепочечным строением инсулина, а также работой программ только с неразветвленной аминокислотной последовательностью, мы анализировали

А- и В-цепи по отдельности. В ходе биоинформатического анализа с помощью пяти программ были выявлены амилоидогенные последовательности для А- и В-цепей инсулина (табл. 2).

Последовательности, которые определяются той или иной программой как амилоидогенные, выявляются с точностью до одного а.о., однако, как следует из табл. 2, нет одинаковых предсказаний амилоидогенных фрагментов разными программами. В то же время выравнивание амилоидогенных последовательностей, обнаруженных несколькими программами (табл. 2), позволяет выявить схожие результаты для А-цепи (участок А8–А15) и В-цепи (участки В12–В19 и В23–В27) инсулина.

Программа ArchCandy не обнаружила амилоидогенных участков, т.к. значение  $S_i$ -функции оценки амилоидогенности последовательности оказалось ниже порогового значения, равного 0,578, что указывает на неамилоидогенность последовательности или неприменимость метода по отношению к данному объекту исследования [30]. Кроме того, программа PASTA2.0 не обнаружила амилоидогенных участков в В-цепи инсулина.

Из полученных результатов на примере исследуемого объекта можно сделать вывод, что эти программы лучше не использовать в качестве основного метода определения амилоидогенных участков, тогда как экспериментальные методы являются основными. Экспериментальным методом обнаружения амилоидогенных участков, образующих остов амилоидных фибрилл, является ферментативный гидролиз последних смесью протеаз и выявление защищенных от протеолиза фрагментов [26, 41]. Для определения амилоидогенных участков в инсулине мы проанализировали гидролизаты фибрилл методом жидкостной хроматографии совместно с масс-спектрометрией (ЖХ-МС).

**Определение амилоидогенных участков фрагментов инсулина методом ограниченного протеолиза с последующим анализом ЖХ-МС.** Метод масс-спектрометрии позволяет проводить расчет отношения массы иона к его заряду ( $m/z$ ) для выходящих с колонки гидролизатов независимо, что позволяет разрешить перекрывающиеся хроматографические пики двух и более образцов (рис. 2).

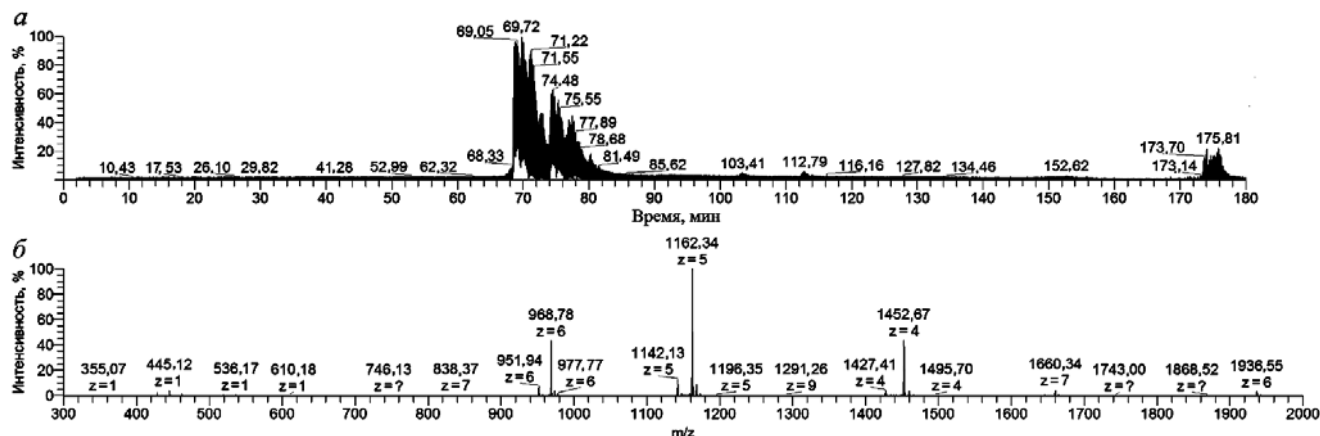
Как видно из представленного на рис. 2 масс-спектра, полученного для элюированного с разделяющей колонки образца гидролизата фибрилл инсулина, основной вклад в суммарный ионный ток вносят многозарядные ионы с диапазоном  $m/z$  900–1500. Пики с величинами  $m/z$  968,78, 1162,34, 1452,67 и соответствующими положительными зарядами (6, 5, 4) опреде-

ляют ионы одной и той же молекулы – мономера инсулина ( $M_r = 5,8$  кДа), в то время как минорные пики с  $m/z$  1660,34 ( $z = 7$ ) и 1936,55 ( $z = 6$ ) соответствуют димеру белка ( $M_r = 11,6$  кДа). Ионы, соответствующие тримерам, гексамерам и более крупным ассоциатам, не обнаруживаются из-за особенностей проведения эксперимента, т.к. они подвергаются диссоциации в ходе электрораспыления, кроме того, значения  $m/z$  для них превышают верхний порог диапазона (2000). Мы сравнили значения  $m/z$ , полученные для панорамных масс-спектров всех детектируемых ионизированных гидролизатов фибрилл инсулина, разделенных на обращенной фазе С18. В масс-спектрах присутствовал набор как одно-, так и многозарядных ионов. Пики с наибольшей интенсивностью были идентичны негидролизованному целым последовательностям белка, а пики с низкой интенсивностью соответствовали крупным фрагментам ( $M_r > 5$  кДа). Для сигналов многозарядных ионов в диапазоне  $m/z$  900–1500 с помощью программного обеспечения Xcalibur и PEAKS Studio 7.5 были определены молекулярные массы и аминокислотные последовательности для негидролизированных пептидов из препаратов фибрилл. Следует подчеркнуть, что на рис. 2 представлен стандартный (неприведенный) масс-спектр при ионизации электроспреем. Молекулы одного и того же вещества при этом типе ионизации способны образовывать несколько по-разному заряженных ионов, соответствующих одной и той же последовательности фрагмента. В табл. 3 приведены только 4-зарядные ионы, которые были выбраны для представления результатов определения устойчивых к протеолизу двухцепочечных фрагментов фибрилл инсулина.

Как видно из табл. 3 и рис. 2, в продуктах гидролиза фибрилл присутствуют ионы, харак-

**Таблица 2.** Предсказанные биоинформатическими программами амилоидогенные участки для А- и В-цепей инсулина

Программа	Амилоидогенные участки	
	А-цепи	В-цепи
FoldAmyloid	A11–A15	B1–B5, B12–B19, B23–B27
ArchCandy	не обнаружены	не обнаружены
PASTA2.0	A2–A12	не обнаружены
Waltz	A10–A21	B12–B18
AGGRESKAN	A8–A15	B7–B19, B23–B27



**Рис. 2.** Хроматограмма (а) и панорамный масс-спектр (б) гидролизатов фибрилл инсулина, полученных при соотношении смеси протеаз к инсулину 1/25 (*m/m*). Для масс-спектра гидролизата образца фибрилл инсулина время выхода составляло 69,05 мин

терные для целостной последовательности белка и двухцепочечных фрагментов. Причем относительная интенсивность пиков молекулярных ионов, соответствующих целому негидролизованному инсулину, превосходит относительную интенсивность пиков молекулярных ионов, которые идентичны фрагментам белка. Полученный результат можно объяснить с помощью данных электронной микроскопии [42], демонстрирующих, что фибриллы инсулина склонны образовывать агрегаты (кластеры, пучки, ленты) из отдельных тонких фибрилл. Таким образом, амилоидные фибриллы, расположенные внутри агрегатов, и, соответственно, образующие их молекулы инсулина, будут недоступны воздействию протеаз.

Примечательно, что не было обнаружено значимого количества ионов, соответствующих фрагментам белка с  $M_r = 1-5$  кДа. Данный ре-

зультат можно интерпретировать следующим образом: в ходе инкубации с протеазами не происходит гидролиз внутренней аминокислотной последовательности инсулина, а протеазы отщепляют пептиды по краям молекулы, поэтому образуются, соответственно, мелкие ( $M_r < 1$  кДа) и крупные ( $M_r > 5$  кДа) фрагменты, которые в сумме примерно соответствуют  $M_r$  мономера. Таким образом, структура белка, входящая в состав фибриллы, сохраняется защищенной от действия протеаз.

В качестве контроля результатов ограниченного протеолиза фибрилл мы использовали обработанный аналогичной смесью протеаз раствор инсулина, не содержащий фибрилл. Соответствующие результаты разделения продуктов гидролиза инсулина и масс-спектрометрического детектирования фрагментов мономера белка анализировали с помощью программ PEAKS

**Таблица 3.** Двухцепочечные фрагменты пептидов, соответствующие сигналам ионного тока, полученным после анализа данных ЖХ-МС для гидролизатов фибрилл инсулина

Последовательность в А-цепи	Последовательность в В-цепи	Обозначение а.о.	m/z (z = 4)	$M_r$ , Да
GIVEQCCTSICSLYQLENYCN	FVNQHLCGSHLVEALYLVCGERGFFYTPKT	A1-A21 и B1-B30	1452,67	5806,7
GIVEQCCTSICSLYQLENYCN	FVNQHLCGSHLVEALYLVCGERGFFYTPK	A1-A21 и B1-B29	1427,41	5705,6
GIVEQCCTSICSLYQLENYCN	FVNQHLCGSHLVEALYLVCGERGFFYTP	A1-A21 и B1-B28	1395,38	5577,5
IVEQCCTSICSLYQLENYCN	NQHLCGSHLVEALYLVCGERGFFYTPKT	A2-A21 и B3-B30	1377,37	5505,5
GIVEQCCTSICSLYQLENYCN	FVNQHLCGSHLVEALYLVCGERGFFYT	A1-A21 и B1-B27	1370,87	5479,5
EQCCTSICSLYQLENYCN	FVNQHLCGSHLVEALYLVCGERGFFYTPK	A4-A21 и B1-B29	1359,58	5434,3
GIVEQCCTSICSLYQLENYCN	FVNQHLCGSHLVEALYLVCGERGFFY	A1-A21 и B1-B26	1345,86	5379,4
GIVEQCCTSICSLYQLENYCN	HLCGSHLVEALYLVCGERGFFYTPKT	A1-A21 и B5-B30	1330,86	5319,4
GIVEQCCTSICSLYQLENYCN	FVNQHLCGSHLVEALYLVCGERGFF	A1-A21 и B1-B25	1305,09	5216,3
VEQCCTSICSLYQLENYC	FVNQHLCGSHLVEALYLVCGERGFFY	A3-A20 и B1-B26	1275,07	5096,3

Studio 7.5 и Xcalibur. По данным масс-спектров, полученных для элюированных с разделяющей колонки гидролизатов мономеров инсулина, основной вклад в суммарный ионный ток вносили однозарядные ионы с диапазоном  $m/z$  от 300 (нижнее пороговое регистрируемое значение в нашем эксперименте) до 1000 (рис. 3).

Относительная интенсивность ионного тока для пиков с величинами  $m/z > 900$  для одно- и многозарядных ионов была незначительна по сравнению с данными, полученными в ходе анализа гидролизатов фибрилл. По соотношениям  $m/z$ , выявленным для фрагментов хроматографического разделения продуктов гидролиза мономеров инсулина, мы сделали вывод, что смесь протеаз полностью расщепляла белок, протеолитический гидролиз был глубоким, и в смеси гидролизатов практически отсутствовали пептиды с  $M_r > 1$  кДа.

После ограниченного протеолиза неагрегированного инсулина детектировались только мелкие фрагменты с  $M_r < 1$  кДа. У фибрилл инсулина после ограниченного протеолиза помимо мелких фрагментов ( $M_r < 1$  кДа) обнаруживались также крупные фрагменты ( $M_r > 5$  кДа). Из этого следует, что фрагменты белка с  $M_r > 5$  кДа, представленные в табл. 3, содержат амилоидогенные последовательности, входящие в остов фибрилл. Особенностью метода ограниченного протеолиза, применимо к нашему объекту исследования, является то, что участки молекулы инсулина, задействованные в образовании связей амилоидных агрегатов, не способны подвергаться действию используемых протеаз, в отличие от мономерной формы инсулина.

Сопоставление последовательностей, представленных в табл. 3, показывает, что фрагменты фибрилл инсулина, защищенные от протеолиза,

могут отличаться на несколько а.о. Кроме того, как следует из рис. 1, часть последовательностей не соответствует сайтам гидролиза пептидных связей протеазами. Полученные результаты можно объяснить некоторым полиморфизмом амилоидных фибрилл инсулина и низкой специфичностью протеиназы К. Мы проанализировали полученные данные и выбрали общие минимальные последовательности в А- и В-цепях для всех фрагментов. Таким образом, по результатам ограниченного протеолиза для инсулина были предложены следующие амилоидогенные участки молекулы: А4–А20 и В5–В25.

Полученные результаты не противоречат представленной нами ранее модели формирования амилоидных фибрилл инсулина [11, 21]. На основе данных трансмиссионной электронной микроскопии было показано, что зрелые фибриллы построены из повторяющегося общего структурного элемента – кольцевого олигомера [21]. Для инсулина внешний диаметр олигомеров (диаметр фибриллы) составляет ~6–7 нм, внутренний ~2 нм. При этом кольцевые олигомеры уложены в фибрилле либо «кольцо к кольцу» своими боковыми сторонами, либо слегка перекрывают друг друга («домино»), что объясняет полиморфизм фибрилл. Такое строение фибриллы может объяснить особенности результатов ограниченного протеолиза фибрилл. В зависимости от ориентации кольцевых олигомеров, слагающих фибриллу, протеазы имеют разную доступность участков молекулы инсулина с пригодными для гидролиза пептидных связей сайтами. С помощью ЖХ–МС были выявлены негидролизированные целые последовательности белка и устойчивые к протеолизу крупные фрагменты молекулы инсулина с небольшими отличиями по длине последовательности.

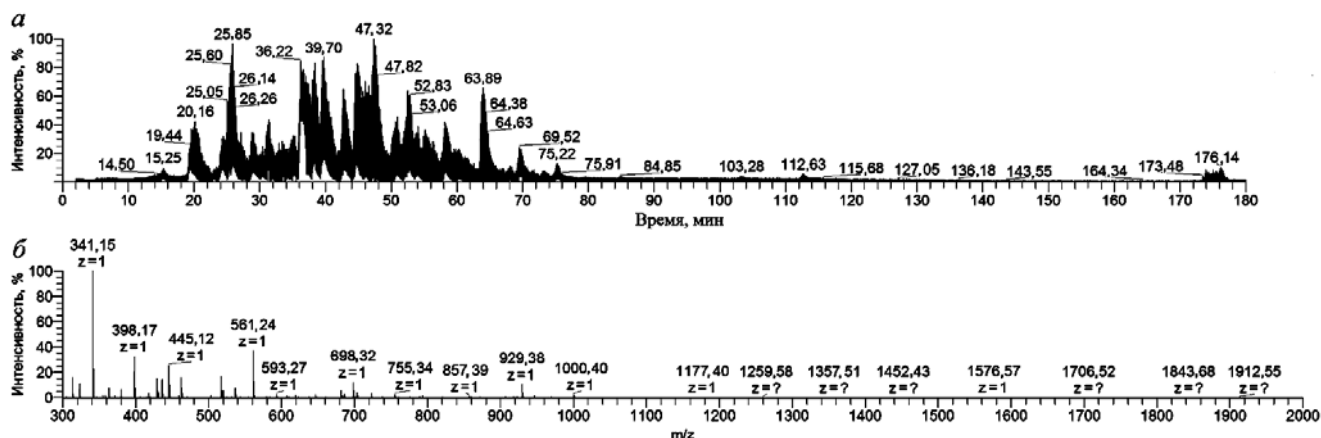


Рис. 3. Хроматограмма (а) и панорамный масс-спектр (б) гидролизатов инсулина (не содержащего фибрилл), полученных при соотношении смеси протеаз к инсулину 1/25 ( $m/m$ ). Для масс-спектра гидролизата образца инсулина время выхода составляло 69,05 мин

После экспериментального определения остова фибрилл мы в первую очередь сравнили полученные результаты с нашими данными биоинформатического определения амилоидогенных последовательностей инсулина (рис. 4).

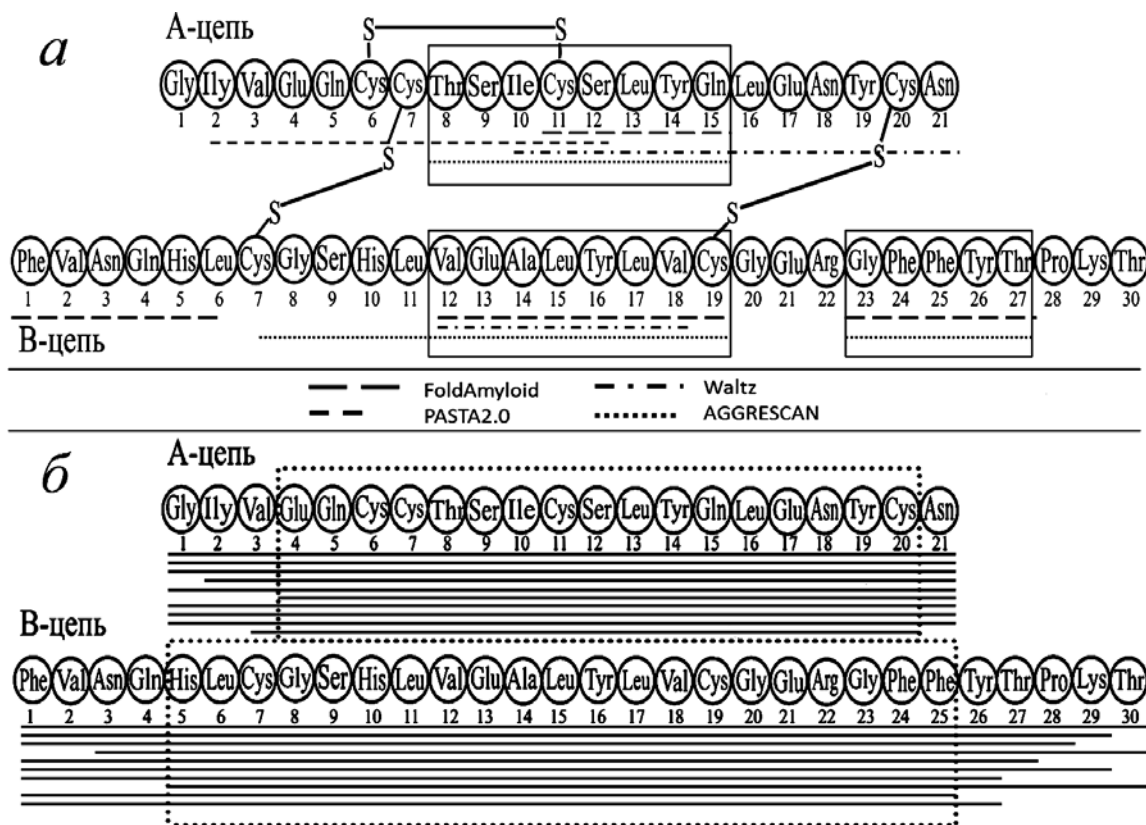
Нами было выявлено, что амилоидогенные участки в А-цепи (A8–A15) и В-цепи (B12–B19 и B23–B27) инсулина, предсказанные тремя программами (FoldAmyloid, Waltz и AGGRESCAN), пересекаются с экспериментальными данными для остова фибрилл инсулина (защищенные от протеолиза участки цепей A4–A20 и B5–B25). Общими между теоретически предсказанными и неподверженными протеолизу последовательностями являются следующие участки: A8–A15 в А-цепи и B12–B19, B23–B25 в В-цепи. Все эти участки находятся внутри участков, которые были определены экспериментальным способом.

Как показано на рис. 4, последовательности A8–A15 и B12–B19 входят в остов фибрилл инсулина и находятся между участками, ограниченными межцепочечными S–S-связями в мономере белка. В перспективе представляет интерес синтез пептидов, соответствующих определен-

ным нами амилоидогенным участкам A4–A20 и B5–B25, для оценки того, какие фибриллы они образуют, и отличается ли их морфология.

Таким образом, с помощью программ FoldAmyloid, Waltz и AGGRESCAN, предсказывающих амилоидогенные фрагменты в белковой цепи, были выявлены общие амилоидогенные участки для инсулина. По результатам протеолитического гидролиза фибрилл инсулина и последующего анализа гидролизатов с помощью жидкостной хроматографии и масс-спектрометрии были выявлены минимальные амилоидогенные участки, устойчивые к действию протеаз. На основе теоретических и экспериментальных методов в А- и В-цепях инсулина были предложены последовательности A8–A15 и B12–B19, B23–B25 соответственно. Полученные данные не противоречат модели, предложенной для фибриллообразования инсулина [11, 21].

Полученные в настоящей работе результаты могут быть использованы для разработки новых устойчивых к агрегации аналогов инсулина. Склонность инсулина к формированию амилоидных фибрилл при температуре выше комнат-



**Рис. 4.** Результаты определения амилоидогенных фрагментов в структуре инсулина: с использованием биоинформатических программ (а) и методом ограниченного протеолиза с ЖХ-МС (б). а – Линиями различных типов подчеркнуты амилоидогенные а.о., предсказанные соответствующими программами. Рамками выделены консенсусные последовательности A8–A15, B12–B19 и B23–B27; б – полужирными линиями подчеркнуты двухцепочечные фрагменты, устойчивые к протеолизу. Последовательности A4–A20 и B5–B25, соответствующие предполагаемому остову фибрилл, выделены пунктирными рамками



ной служит основным препятствием для применения пептида, например, в инсулиновой помпе [15]. В фармакологических составах мономер инсулина обычно используется в виде гексамера, координированного ионами цинка, но это накладывает определенные ограничения на разработку инсулиновых помп, в которых мономер должен быть свободен для быстрой диффузии, но при длительном хранении могут формировать амилоиды. Исследование амилоидогенеза может обеспечить теоретические основы создания ультрастабильных аналогов инсулина для применения в проектах по искусственной поджелудочной железе. Наш опыт работы с инсулином позволяет сделать заключение, что наиболее стабильным и устойчивым к амилоидной агрегации будет аналог инсулина, в котором отсут-

ствуют амилоидогенные аминокислоты, вносящие наибольший вклад в процесс амилоидообразования, согласно расчетным моделям, которые положены в основу работы программы Fold-Amyloid [3, 29].

#### Финансирование

Работа выполнена при поддержке РФФИ (грант № 18-14-00321). Масс-спектрометрические измерения проведены на базе оборудования ЦКП Института белка РАН.

#### Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

### СПИСОК ЛИТЕРАТУРЫ

- Kelly, J.W. (1996) Alternative conformations of amyloidogenic proteins govern their behavior, *Curr. Opin. Struct. Biol.*, **6**, 11–17.
- Dobson, C.M. (2001) Protein folding and its links with human disease, *Biochem. Soc. Symp.*, **68**, 1–26.
- Галзитская О.В., Гарбузинский С.А., Лобанов М.Ю. (2006) Поиск амилоидогенных участков белковой цепи, *Мол. биология*, **40**, 910–918.
- Sipe, J.D., Benson, M.D., Buxbaum, J.N., Ikeda, S., Merlini, G., Saraiva, M.J., and Westermarck, P. (2012) Amyloid fibril protein nomenclature: 2012 recommendations from the Nomenclature Committee of the International Society of Amyloidosis, *Amyloid*, **19**, 167–170.
- O'Donnell, C.W., Waldispuhl, J., Lis, M., Halfmann, R., Devadas, S., Lindquist, S., and Berger, B. (2011) A method for probing the mutational landscape of amyloid structure, *Bioinformatics (Oxford, England)*, **27**, 34–42.
- Sanger, F., and Tuppy, H. (1951) The amino-acid sequence in the phenylalanyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates, *Biochem. J.*, **49**, 463–481.
- Sanger, F., and Tuppy, H. (1951) The amino-acid sequence in the phenylalanyl chain of insulin. 2. The investigation of peptides from enzymic hydrolysates, *Biochem. J.*, **49**, 481–490.
- Waugh, D.F. (1941) The properties of protein fibers produced reversibly from soluble protein molecules, *A. J. Physiol.*, **133**, 484–485.
- Klunk, W.E., Pettegrew, J.W., and Abraham, D.J. (1989) Quantitative evaluation of congo red binding to amyloid-like proteins with a  $\beta$ -pleated sheet conformation, *J. Histochem. Cytochem.*, **37**, 1273–1281.
- Brange, J., Andersen, L., Laursen, E.D., Meyn, G., and Rasmussen, E. (1997) Toward understanding insulin fibrillation, *J. Pharm. Sci.*, **86**, 517–525.
- Selivanova, O.M., Grishin, S.Yu., Glyakina, A.V., Sadgjan, A.S., Ushakova, N.I., and Galzitskaya, O.V. (2018) Analysis of insulin analogs and the strategy of their further development, *Biochemistry (Moscow)*, **83**, 146–162.
- Baker, E.N., Blundell, T.L., Cutfield, J.F., Cutfield, S.M., Dodson, E.J., Dodson, G.G., Hodgkin, D.M.C., Hubbard, R.E., Isaacs, N.W., Reynolds, C.D., Sakabe, K., Sakabe, N., and Vijayan, N.M. (1988) The structure of 2Zn pig insulin crystals at 1.5 Å resolution, *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **319**, 369–456.
- Cutfield, J.F., Cutfield, S.M., Dodson, E.J., Dodson, G.G., Emdin, S.F., and Reynolds, C.D. (1979) Structure and biological activity of hagfish insulin, *J. Mol. Biol.*, **132**, 85–100.
- Frankaer, C.G., Sonderby, P., Bang, M.B., Mateiu, R.V., Groenning, M., Bukrinski, J., and Harris, P. (2017) Insulin fibrillation: the influence and coordination of  $Zn^{2+}$ , *J. Struct. Biol.*, **199**, 27–38.
- Phillips, N.B., Whittaker, J., Ismail-Beigi, F., and Weiss, M.A. (2012) Insulin fibrillation and protein design: topological resistance of single-chain analogs to thermal degradation with application to a pump reservoir, *J. Diabetes Sci. Technol.*, **6**, 277–288.
- Nielsen, L., Frokjaer, S., Brange, J., Uversky, V.N., and Fink, A.L. (2001) Probing the mechanism of insulin fibril formation with insulin mutants, *Biochemistry*, **40**, 8397–8409.
- Ahmad, A., Millett, I.S., Doniach, S., Uversky, V.N., and Fink, A.L. (2003) Partially folded intermediates in insulin fibrillation, *Biochemistry*, **42**, 11404–11416.
- Vestergaard, B., Groenning, M., Roessle, M., Kastrop, J.S., Weert, M., Flink, J.M., Frokjaer, S., Gajhede, M., and Svergun, D.I. (2007) A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils, *PLoS Biol.*, **5**, 134–146.
- Jimenez, J.L., Nettleton, E.J., Bouchard, M., Robinson, C.V., Dobson, C.M., and Saibil, H.R. (2002) The protofilament structure of insulin amyloid fibrils, *Proc. Natl. Acad. Sci. USA*, **99**, 9196–9201.
- Kajava, A.V., Baxa, U., and Steven, A.C. (2010)  $\beta$ -arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils, *FASEB J.*, **24**, 1311–1319.
- Selivanova, O.M., Suvorina, M.Y., Surin, A.K., Dovidchenko, N.V., and Galzitskaya, O.V. (2017) Insulin and lispro insulin: what is common and different in their behavior? *Curr. Protein Pept. Sci.*, **18**, 57–64.
- Meersman, F., and Dobson, C.M. (2006) Probing the pressure-temperature stability of amyloid fibrils provides new insights into their molecular properties, *Biochim. Biophys. Acta*, **1764**, 452–460.
- Malisauskas, M., Weise, C., Yanamandra, K., Wolf-Watz, M., and Morozova-Roche, L. (2010) Liability landscape and

- protease resistance of human insulin amyloid: a new insight into its molecular properties, *J. Mol. Biol.*, **396**, 60–74.
24. Kheterpal, I., Williams, A., Murphy, C., Bledsoe, B., and Wetzel, R. (2001) Structural features of the A $\beta$  amyloid fibril elucidated by limited proteolysis, *Biochemistry*, **40**, 11757–11767.
  25. Piejko, M., Dec, R., Babenko, V., Hoang, A., Szewczyk, M., Mak, P., and Dzwolak, W. (2015) Highly amyloidogenic two-chain peptide fragments are released upon partial digestion of insulin with pepsin, *J. Biol. Chem.*, **290**, 5947–5958.
  26. Surin, A.K., Grigorashvili, E.I., Suvorina, M.Y., Selivanova, O.M., and Galzitskaya, O.V. (2016) Determination of regions involved in amyloid fibril formation for A $\beta$ (1–40) peptide, *Biochemistry (Moscow)*, **81**, 762–769.
  27. Selivanova, O.M., Suvorina, M.Y., Dovidchenko, N.V., Eliseeva, I.A., Surin, A.K., Finkelstein, A.V., Schmatchenko, V.V., and Galzitskaya, O.V. (2014) How to determine the size of folding nuclei of protofibrils from the concentration dependence of the rate and lag-time of aggregation. II. Experimental application for insulin and LysPro insulin: aggregation morphology, kinetics, and sizes of nuclei, *J. Phys. Chem. B*, **118**, 1198–1206.
  28. Porter, R.R. (1953) Partition chromatography of insulin and other proteins, *Biochem. J.*, **53**, 320–328.
  29. Garbuzynskiy, S.O., Lobanov, M.Y., and Galzitskaya, O.V. (2010) FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence, *Bioinformatics*, **26**, 326–332.
  30. Ahmed, A.B., Znassi, N., Chateau, M.-T., and Kajava, A.V. (2015) A structure-based approach to predict predisposition to amyloidosis, *Alzheimers Dement.*, **11**, 681–690.
  31. Trovato, A., Seno, F., and Tosatto, S.C.E. (2007) The PASTA server for protein aggregation prediction, *Protein Eng. Des. Sel.*, **20**, 521–523.
  32. Walsh, I., Seno, F., Tosatto, S.C.E., and Trovato, A. (2014) PASTA 2.0: an improved server for protein aggregation prediction, *Nucleic Acids Res.*, **42**, W301–W307.
  33. Maurer-Stroh, S., Debulpaep, M., Kuemmerer, N., Lopez de la Paz, M., Martins, I.C., Reumers, J., Morris, K.L., Copland, A., Serpell, L., Serrano, L., Schymkowitz, J.W.H., and Rousseau, F. (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices, *Nat. Methods*, **7**, 237–242.
  34. Conchillo-Sole, O., de Groot, N.S., Aviles, F.X., Vendrell, J., Daura, X., and Ventura, S. (2007) AGGRESKAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides, *BMC Bioinformatics*, **8**, 1–17.
  35. Zurdo, J., Gujjarro, J.I., and Dobson, C.M. (2001) Preparation and characterization of purified amyloid fibrils, *J. Am. Chem. Soc.*, **123**, 8141–8142.
  36. Galzitskaya, O.V., and Selivanova, O.M. (2017) Rosetta stone for amyloid fibrils: the key role of ring-like oligomers in amyloidogenesis, *J. Alzheimers Dis.*, **59**, 785–795.
  37. Olsen, J.V., Ong, S.E., and Mann, M. (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues, *Mol. Cell. Proteomics*, **3**, 608–614.
  38. Appel, W. (1986) Chymotrypsin: molecular and catalytic properties, *Clin. Biochem.*, **19**, 317–322.
  39. Kraus, E., Kiltz, H.H., and Femfert, U.F. (1976) The specificity of proteinase K against oxidized insulin B chain, *Hoppe-Seyler's Z. Physiol. Chem.*, **357**, 233–237.
  40. Morihara, K., and Tszuki, H. (1975) Specificity of proteinase K from *Tritirachium album* limber for synthetic peptides, *Agric. Biol. Chem.*, **39**, 1489–1492.
  41. Polverino de Laureto, P., Taddei, N., Frare, E., Capanni, C., Costantini, S., Zurdo, J., Chiti, F., Dobson, C.M., and Fontana, A. (2003) Protein aggregation and amyloid fibril formation by an SH3 domain probed by limited proteolysis, *J. Mol. Biol.*, **334**, 129–141.
  42. Selivanova, O.M., and Galzitskaya, O.V. (2012) Structural polymorphism and possible pathways of amyloid fibril formation on the example of insulin protein, *Biochemistry (Moscow)*, **77**, 1237–1247.

## DEFINING AMYLOIDOGENIC REGIONS BELONGING TO THE SPINE OF INSULIN FIBRILS

A. K. Surin<sup>1,2,3</sup>, S. Yu. Grishin<sup>1,4</sup>, and O. V. Galzitskaya<sup>1\*</sup>

<sup>1</sup> Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Russia; E-mail: ogalzit@vega.protres.ru

<sup>2</sup> State Research Center for Applied Microbiology and Biotechnology, 142279 Obolensk, Russia; E-mail: alan@vega.protres.ru

<sup>3</sup> Gamaleya National Research Center of Epidemiology and Microbiology, Ministry of Health of the Russian Federation, 123098 Moscow, Russia

<sup>4</sup> Lomonosov Moscow State University, 119991 Moscow, Russia; E-mail: syugrishin@gmail.com

Received June 15, 2018

Revision received September 3, 2018

Accepted September 3, 2018

To reveal conformational changes resulting in formation of insulin amyloids it is required to define amyloidogenic regions in the structure of protein monomers. Different models of the insulin fibrillogenesis were proposed previously. However, precise regions responsible for the formation of amyloid fibrils have not been defined. Using the bioinformatics programs for predicting amyloidogenic regions, we have defined some common amyloidogenic sequences in the structure of insulin monomers. Using the methods of limited proteolysis and mass spectrometry determination of protein chain fragments resistant to action of proteases, we revealed amino acid sequences in the insulin structure that may be present in the spine of insulin fibrils. The obtained results do not conflict with the earlier proposed model of fibril formation from ring-like oligomers and can be used for designing insulin analogs resistant to amyloidogenesis.

**Keywords:** amyloid, amyloidogenic regions, FoldAmyloid, insulin, ring-like oligomers, mass spectrometry, protein proteolysis