

ТЕОРИЯ И МОДЕЛИРОВАНИЕ

УДК 541.64:539.2

О ВЛИЯНИИ ВТОРИЧНЫХ СТРУКТУР В ПЕПТИДАХ И БЕЛКАХ НА ИХ ХРОМАТОГРАФИЧЕСКОЕ РАЗДЕЛЕНИЕ

© 2020 г. А. В. Горшков^а, И. Г. Роздина^а, М. Л. Придатченко^а, В. В. Евреинов^{а,*}

^а Федеральный исследовательский центр химической физики им. Н.Н. Семенова Российской академии наук
119991 Москва, ул. Косыгина, 4, Россия

*e-mail: evreinov@polymer.chph.ras.ru

Поступила в редакцию 09.08.2019 г.

После доработки 03.02.2020 г.

Принята к публикации 17.02.2020 г.

В рамках решеточной модели адсорбции и хроматографии макромолекул рассмотрен вопрос о влиянии остаточных вторичных структур в пептидах и белках на их разделение. Показано, что учет в модели вторичных структур, таких как спирали, складки или сшивки, может приводить к заметному изменению объема удерживания. Вместе с тем, для получающихся в результате ферментативного гидролиза белков коротких пептидов с числом аминокислотных звеньев порядка двадцати, воздействие вторичных структур, например, остаточных *S*–*S*-сшивок, на разделение в модельном хроматографическом эксперименте незначительно. Для таких пептидов основное влияние на разделение оказывает их аминокислотный состав, и адекватное предсказание значений объема удерживания может быть получено как в приближении линейной цепи, так и в аддитивной модели адсорбции.

DOI: 10.31857/S2308112020040069

ВВЕДЕНИЕ

Макромолекулы белков в живых организмах существуют в виде трехмерных глобулярных структур. Такие структуры представляют собой плотную пространственную укладку α -спиралей, β -листов, и других элементов вторичной структуры, а также и поперечных сшивок между удаленными по цепи звеньями за счет дисульфидных *S*–*S*-мостиков. Хотя последние формально не являются элементами вторичных структур, для упрощения изложения в дальнейшем под элементами вторичных структур подразумеваются и *S*–*S*-сшивки.

В условиях хромато-масс-спектрометрического эксперимента (LC–MS), который обычно проводится в водно-органических средах при pH 2.0 и температуре 50°C, происходит денатурация белка, и его глобулярная структура разрушается. Макромолекула разворачивается, приобретая конформацию, близкую к конформации клубка, при этом аминокислотные остатки, “спрятанные” внутри глобулы, выходят наружу, и для них появляется возможность взаимодействовать с поверхностью. Тем самым, при адсорбционном взаимодействии вся последовательность цепи становится доступной для “чтения” ее текста. Такая картина подтверждается экспериментальными данными: объем удерживания V_R коррелирует со всем аминокислотным составом белка, а не толь-

ко с той его частью, которая находится на поверхности глобулы. По этой причине модель хроматографического разделения BioLCCC [1], в основе которой лежит предположение о макромолекуле как о клубке, представленном цепью случайных блужданий на решетке, в целом правильно отображает наблюдаемые экспериментальные закономерности.

В состоянии клубка все звенья имеют возможность взаимодействовать с поверхностью, создавая наиболее выгодные условия для “чтения” текста последовательности методами хроматографии. Точечные мутации последовательности, такие как перестановка, замена аминокислот и их химическая модификация относительно легко “читаемы” в макромолекуле, находящейся в клубковом состоянии. Связанность аминокислотных остатков в цепь формирует специфическое сильное взаимодействие между удаленными по цепи звеньями, при котором небольшое изменение адсорбционных свойств какого-то одного звена (например, в результате его модификации) влияет на взаимодействие с поверхностью всех остальных и, следовательно, на адсорбцию всей макромолекулы и ее хроматографические свойства. Зависимость адсорбции от последовательности звеньев в клубке имеет, как правило, энтропийную природу, и связана с наличием в адсорбированной макромолекуле петель и хвостов, распростирающихся в раствор и не контактирующих с

поверхностью. При очень сильном адсорбционном взаимодействии (практически не реализуемом в LC), когда макромолекула полностью “распластана” на поверхности, последовательность звеньев не играет роли.

Особенно заметно влияние последовательности вблизи адсорбционной критической точки, в которой адсорбированный клубок как раз характеризуется наличием длинных петель и хвостов. В критической точке имеет место тонкая компенсация потерь энтропии цепи вблизи поверхности, обусловленных уменьшением числа способов укладки, энергией адсорбционного взаимодействия звеньев, и макромолекула оказывается весьма восприимчива к любому нарушению этой компенсации. В силу фазового характера адсорбции макромолекул, эксперимент LC–MS, осуществляемый при изменении во времени состава водно-органического растворителя, автоматически приводит к тому, что разделение происходит вблизи критической точки адсорбции со стороны адсорбционного режима. Тем самым, разделение в градиенте должно отображать не только аминокислотный состав макромолекулы, но и характер чередования остатков в цепи, т.е. ее последовательность.

Для коротких цепей пептидов вклад петель и хвостов незначителен, и основную роль здесь играет аминокислотный состав. Для них весьма успешными являются аддитивные модели [2–5], дополненные поправками на длину цепи и влияние соседей.

Такая идеализированная картина, однако, не вполне точная. Существующие в глобулярном состоянии белка вторичные структуры могут не полностью разрушаться в эксперименте. В этом случае конформацию денатурированного белка можно представить в виде небольшого количества остатков α -спиралей и β -листов, соединенных гибкими цепями, сшитыми в некоторых местах дисульфидными S–S-мостиками. Остаточные вторичные структуры меняют адсорбционные свойства всей макромолекулы и должны учитываться моделью разделения. В настоящей работе в рамках модели BioLCCC теоретически проанализировано влияние остаточных вторичных структур на адсорбционные свойства белков и пептидов и указанное влияние сопоставлено с влиянием последовательности.

МОДЕЛЬ BioLCCC С ВТОРИЧНЫМИ СТРУКТУРАМИ

На рис. 1 показаны элементарные типы вторичных структур, которые могут существовать в частично денатурированном белке. В масштабе клубка α -спираль и β -лист можно представить

жестким стержнем, а соединяющие их гибкие участки — цепью случайных блужданий.

Денатурированный белок может содержать и несколько таких структур, соединенных между собой (рис. 1, тип 5 и 6). Рассмотрим, каким образом вторичные структуры возможно интегрировать в матричную модель адсорбции макромолекулы в пространстве щелеобразной поры, лежащей в основе BioLCCC.

Напомним, что измеряемый в хроматографии объем удерживания V_R есть не что иное, как статистическая сумма макромолекулы в колонке, поэтому для понимания закономерностей разделения макромолекул необходимо найти их статистические суммы в представлении цепи блужданий со вставками в некоторых местах жестких стержней и сшивок. В дальнейшем выберем нормировку энергии таким образом, чтобы в неограниченном пространстве в отсутствие адсорбционных поверхностей статсумма макромолекулы Z была равна единице, а энергию адсорбции звеньев будем выражать в единицах kT .

Клубковая часть, как обычно, моделируется цепью, блуждающей по узлам кубической решетки, переходная матрица $W(\epsilon_i)$ которой имеет стандартный вид [1]:

$$W(\epsilon_i) = \begin{bmatrix} 2/3e^{\epsilon_i} & 2/3e^{\epsilon_i} & 0 & 0 & \dots & 0 \\ 1/6 & 2/3 & 1/6 & 0 & \dots & 0 \\ 0 & 1/6 & 2/3 & 1/6 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1/6 & 2/3 & 1/6 \\ 0 & 0 & \dots & 0 & 1/6e^{\epsilon_i} & 2/3e^{\epsilon_i} \end{bmatrix} \quad (1)$$

Здесь ϵ_i — энергия адсорбции i -го звена в цепи (для определенности будем отсчитывать его от C-конца), зависящая от типа соответствующего аминокислотного остатка. Для топологии поры типа щели с размером D , ограниченной в одном направлении, при расчете статсуммы можно исключить интегрирование по двум пространственным координатам в плоскости, параллельной поверхности. Тогда статистическую сумму цепи, начинающуюся в слое k и заканчивающуюся в слое l , запишем в виде:

$$Z_n(k, l) = U_l^T \times \prod_{i=1}^n W(\epsilon_i) \times U_k,$$

где $U_{k,l}$ — векторы, все компоненты которых нули, за исключением $k = 1$ и $l = 1$.

В данной модели радиус действия адсорбционных сил δ ограничен размером аминокислотного остатка a , т.е. предполагается короткодействующим. Обычно размер a , диаметр α -спирали d и размер молекул растворителя r имеют один по-

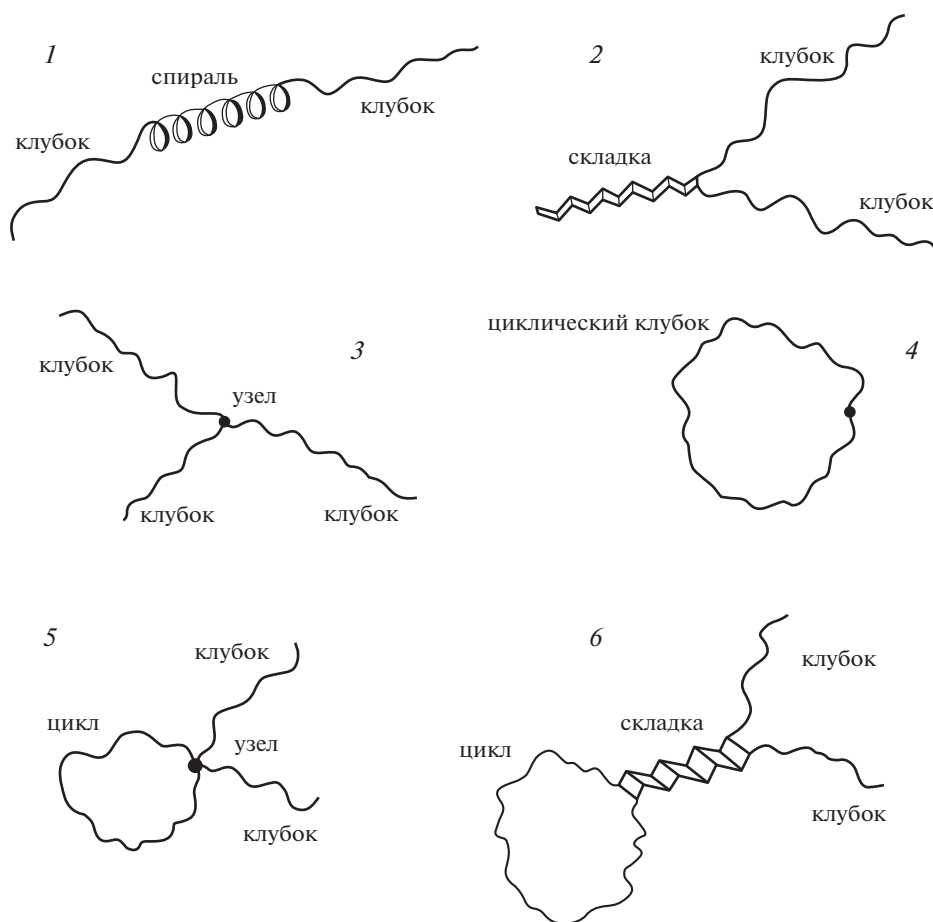


Рис. 1. Элементы вторичных структур (1–4) и их комбинации (5, 6) в макромолекуле: 1 – спираль, 2 – складка, 3 – сшивка, 4 – цикл.

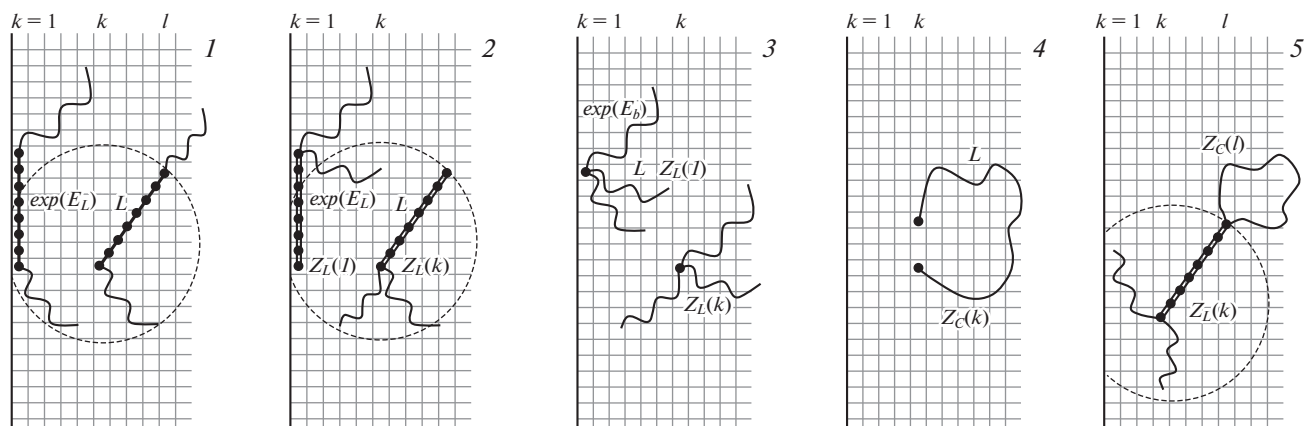


Рис. 2. Представление спирали (1), складки (2), сшивки (3) и цикла (4) в решеточной модели, а также их комбинаций (5).

рядок величины. Поэтому будет естественно величины, имеющие размерность длины, выразить в единицах размера аминокислотного звена $a = 5 \text{ \AA}$. Тогда размерность матрицы равна размеру поры адсорбента, представленному в единицах размера звена (например, для поры с размером $300 \text{ \AA}$ размерность матрицы $D = 60$).

Для спирального участка, моделируемого стержнем, шаг по решетке зависит от его длины L , которая связана с числом n входящих в него аминокислотных остатков как $L = n - 1$. Первый аминокислотный остаток, входящий в спираль, логично считать последним звеном клубковой части (рис. 2). При выбранной симметрии поры и

уровне энергии для расчета статсуммы существенны только положения конца стержня в плоскостях, параллельных адсорбционным поверхностям.

Стержень с длиной L (обычно для белков спираль содержит не менее четырех звеньев, т.е. $L > 4$), начинающийся в слое k , может располагаться в слоях $k \pm 1, \dots, k \pm L$, общее число которых $2L + 1$.

$$W_{\alpha}(L) = \begin{bmatrix} 1/(2L+1)e^{E_{\alpha}} & 1/(2L+1)e^{\varepsilon_n} & 1/(2L+1)e^{\varepsilon_n} & 0 & \dots & 0 \\ 1/(2L+1) & 1/(2L+1) & 1/(2L+1) & 1/(2L+1) & \dots & 0 \\ 1/(2L+1) & 1/(2L+1) & 1/(2L+1) & 1/(2L+1) & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1/(2L+1) & 1/(2L+1) & 1/(2L+1) \\ 0 & 0 & \dots & 1/(2L+1)e^{\varepsilon_n} & 1/(2L+1)e^{\varepsilon_n} & 1/(2L+1)e^{E_{\alpha}} \end{bmatrix} \quad (2)$$

Для упрощения показана матрица для $L = 2$, в общем случае ее структура очевидна. Энергия адсорбционного взаимодействия тогда, когда стержень целиком попадает в адсорбционный слой, есть сумма значений энергии взаимодействия всех входящих в него аминокислотных остатков,

$E_{\alpha} = \sum_2^n \varepsilon_i$. Если только конец стержня n попадает в адсорбционный слой, энергия адсорбции равна энергии адсорбции концевых аминокислотного остатка ε_n ; будем считать, что все остальные звенья стержня в такой “наклонной” конфигурации с поверхностью не взаимодействуют. Длина стержня $L = 1$, по сути, соответствует одному шагу цепи блужданий, однако, матрица (2) в этом случае несколько отличается от переходной матрицы цепи блужданий (1). Данное отличие не существенно, и для реальных длинных стержней статсумма, рассчитанная по выражению (2), мало отличается от точной модели адсорбции стержня [6, 7].

Таким образом, статсумму цепи общей длиной N , содержащей спиральный участок между звеньями m и k , представим в виде

$$Z_N = U^T \times \prod_{i=k+1}^N W(\varepsilon_i) \times W_{\alpha}(k-m) \times \prod_{i=1}^m W(\varepsilon_i) \times U, \quad (3)$$

где U — единичный вектор. Обобщение на случай нескольких спиральных участков в цепи очевидно.

Рассмотрим другие типы вторичных структур, представленные на рис. 2. Для начала запишем очевидное тождество для цепи блужданий с длиной N , которую можно представить в качестве цепи с длиной n , пришедшую в слой k , а затем покинувшего его цепь с длиной $N - n$. Если просуммировать по всем слоям $k = 1 \dots D$, то в итоге должна

Поскольку расстояние между плоскостями одинаковое, то вероятность попадания конца стержня на поверхность соответствующего шарового слоя также одинаковая, т.е. вероятность попадания второго конца стержня в любой доступный слой равна $1/(2L + 1)$. Таким образом, переходная матрица, описывающая стержень спирали в поре, имеет вид:

получиться просто статсумма цепи блужданий с длиной N :

$$\begin{aligned} Z_N &= \sum_{k=1}^D \left(U^T \times \prod_{n+1}^N W(\varepsilon_i) \times U_k^T \times U_k \times \right. \\ &\quad \left. \times \prod_1^n W(\varepsilon_i) \times U \right) = \\ &= U^T \times \prod_{n+1}^N W(\varepsilon_i) \times E \times \prod_1^n W(\varepsilon_i) \times U \end{aligned} \quad (4)$$

Здесь E — единичная матрица. Если в звене n имеется, например, узел разветвления (рис. 2, схема 3) и к нему подходит еще одна цепь, то вместо единичной матрицы в уравнении (4) необходимо поставить диагональную матрицу, элементы которой представляют собой статистические суммы боковой цепи с длиной N_b , начинающейся в слое k :

$$W_b(k, k) = Z_{N_b}(k) = U^T \times \prod_{i=1}^{N_b} W(\varepsilon_i) \times U_k$$

Таким образом, статистическую сумму цепи с узлом разветвления запишем в виде

$$\begin{aligned} Z_N &= U^T \times \prod_{n+1}^N W(\varepsilon_i) \times \\ &\times \begin{bmatrix} Z_{N_b}(1) & 0 & \dots & 0 \\ 0 & Z_{N_b}(2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & Z_{N_b}(D) \end{bmatrix} \times \prod_1^n W(\varepsilon_i) \times U \end{aligned}$$

Если к звену n примыкает складка (рис. 2, схема 2), простирающаяся до l -го звена, тогда элементы диагональной матрицы суть статсуммы стержня длиной $L = [l - m]/2$, один конец которого находится в слое k :

$$W_{\beta}(k, k) = Z_{\beta}(k) = U^T \times W_{\alpha}(L) \times U_k$$

Статистическую сумму цепи со складкой можно представить как

$$Z_N = U^T \times \prod_{i=l+1}^N W(\epsilon_i) \times W_{\beta}(l-m) \times \prod_{i=1}^m W(\epsilon_i) \times U$$

Наконец, цепь может содержать в своей структуре цикл, статистическая сумма которого

$$\begin{aligned} Z_c &= A \sum_1^D Z_N(k, k) = \\ &= A \sum_1^D \left(U_k^T \times \prod_{i=1}^N W(\epsilon_i) \times U_k \right) \end{aligned}$$

Множитель A учитывает отличие статсуммы цикла в неограниченном пространстве от единицы. Для кубической решетки этот множитель быстро становится близким к статсумме одномерной гауссовой циклической цепи:

$$\frac{1}{A} \sim \sqrt{\frac{3}{2\pi N}}$$

Для рассматриваемой одномерной модели поры можно не учитывать то, что конец цикла не попадает в исходную точку (рис. 2, схема 4), а оказывается лишь в исходном слое — по координатам, параллельным адсорбционным поверхностям, звенья цепи не имеют ограничений.

Рассмотренные выше элементарные типы вторичных структур позволяют записать статсумму и для более сложной структуры, например, изображенной на схеме 5 рис. 2:

$$Z_N = U^T \times \prod_{i=l+1}^N W(\epsilon_i) \times \Lambda \times \prod_{i=1}^m W(\epsilon_i) \times U$$

Здесь Λ — диагональная матрица, элементы которой есть статистические суммы складки, начинающейся в слое k , с примыкающим к ней на другом конце циклом:

$$\Lambda(k, k) = U^T \times W_c \times W_{\beta}(L) \times U_k$$

Для завершения модели необходимо определить значения энергии адсорбции аминокислотных остатков в спиральном состоянии. Это сложно выполнимая задача: для нее потребуются модели пептидов в заведомо спиральном состоянии. В настоящей работе для оценки влияния вторичных структур использован более простой вариант. Поскольку механизм адсорбции на обращенной фазе обусловлен в основном гидрофобным взаимодействием, в первом приближении будем считать, что энергия аминокислотных звеньев в спиральном состоянии близка к энергии адсорбции аминокислот в коротких пептидах, моделируемых жестким стержнем [7]. Если принять такое предположение, то между значениями энер-

гии звеньев в клубке ϵ_c и в спирали ϵ_h существует зависимость, близкая к линейной. Это означает, что более гидрофобная аминокислота в клубковом состоянии будет более гидрофобной и в спиральном состоянии. Такое приближение значительно сокращает набор параметров в модели: вместо двадцати дополнительных значений энергии имеем два дополнительных коэффициента линейного преобразования.

ПРИМЕРЫ РАСЧЕТА И ОЦЕНКА ВЛИЯНИЯ ВТОРИЧНЫХ СТРУКТУР

Рассмотрим особенности, к которым в адсорбционном поведении может приводить учет в клубках вторичных структур. На практике необходимо знать, какие аминокислотные остатки входят в клубковые участки, а какие в спиральные. В приложении к определенной макромолекуле белка, скорее всего, это можно определить экспериментально, используя различные физические методы. В общем случае произвольной последовательности, для теоретического анализа необходимо пользоваться математическим моделированием.

Существует множество алгоритмов, предсказывающих наличие вторичных структур в белках, которые дают примерно одинаковый результат. В настоящей работе была использована модель PSIPRED [8]. Для заданной последовательности цепи эта модель определяет в каком состоянии (c — клубок, h — спираль) находится тот или иной аминокислотный остаток. Например, для белка Cytochrome модель PRISPED предсказывает четыре спиральных участка, включающие звенья 3–18, 50–53, 61–69 и 88–101. Таким образом, если предположить, что в денатурированной макромолекуле Cytochrome спиральные участки сохраняются, то ее можно представить в виде четырех стержней разной длины, соединенных гибкими цепями. Важно, что алгоритмы предсказания вторичных структур неявно предполагают нормальные условия, в то время как хроматографический эксперимент осуществляется обычно в кислой среде (pH 2.0) при температуре $\sim 50^\circ\text{C}$, значит, некоторые из спиралей могут разрушиться.

В рамках такого приближения были рассчитаны значения объема удерживания макромолекул белков как с учетом вторичных структур, так и без учета последних. В расчетах, как и ранее в работе [7], были использованы стандартные условия LC-эксперимента: растворитель вода—ацетонитрил с добавлением кислоты для достижения pH 2.0, градиент от 0 до 100% ацетонитрила в течение 100 мин, колонка 250×4.6 мм, эффективность колонки ~ 4000 теоретических тарелок.

Аминокислотные остатки в спиральном состоянии могут иметь отличную от клубкового состо-

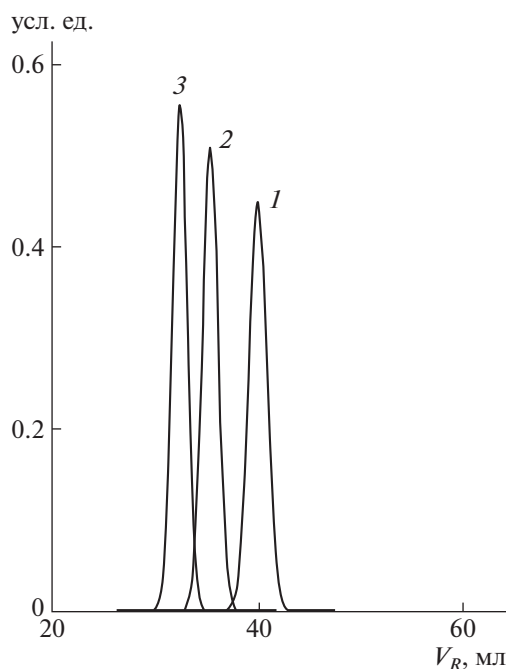


Рис. 3. Модельная хроматограмма белка Cytochrome в представлении клубка (1) и с учетом вторичных структур типа спиралей, включающих звенья 3–18, 50–53, 61–69 и 88–101, при энергии адсорбции остатков в спиральном состоянии, соответствующей энергии в модели адсорбции стержня (2) и клубка (3).

яния энергию, и для более точного моделирования необходимо вводить дополнительную таблицу значений эффективной энергии остатков в состоянии спирали. В данной работе будем считать, что значения энергии остатков в спиральном состоянии близки к значениям эффективной энергии остатков в модели адсорбции пептидов как жестких стержней. Эти значения, а также значения энергии в состоянии клубка, приведены в работе [7].

На рис. 3 показана модельная хроматограмма белка Cytochrome, рассчитанная как в предположении о наличии в структуре клубка спиралей, так и для чисто клубкового состояния. Видно, что наличие вторичных структур в денатурированном белке оказывается весьма существенным и влечет заметное изменение объема удерживания по сравнению с представлением всей цепи в виде клубка. При этом, вопреки ожиданию, макромолекула со спиральными участками элюируется раньше, чем клубок, даже с учетом того, что значения эффективной энергии остатков в спирали больше таковых в клубке.

Чтобы выделить данный факт, был рассчитан объем удерживания макромолекулы в предположении о равенстве энергии адсорбции остатков в спиральном и клубковом состоянии. Объем удерживания такой макромолекулы еще меньше.

Скорее всего, это связано с тем, что переход стержня в адсорбированное состояние, в зависимости от его длины, носит резкий и немонотонный характер [6]. Возможно, что такое поведение отражает особенности модели, в которой энергия “наклонных” состояний спиральных участков принята равной нулю.

Ранее в широком ряду белков обнаружена хорошая корреляция между экспериментальными и теоретическими величинами объемов удерживания в предположении о чисто клубковом состоянии макромолекулы. По-видимому, такая корреляция свидетельствует о том, что в условиях LC-эксперимента вторичные структуры разрушаются и их влияние становится незначительным. Однако, если эксперимент проводить в нормальных условиях при pH 7.0, влияние вторичных структур может оказаться существенным и определять картину хроматографического разделения.

Другой пример — белок группы инсулина Relaxin, состоящий из двух цепей A и B, соединенных тремя связями S–S (рис. 4). Скорее всего, денатурированный белок имеет клубкообразную, сшитую этими связями структуру, включающую циклы и хвосты. При исследовании белки подвергают ферментативному гидролизу, в результате чего они распадаются на фрагменты, в которых связи S–S не разрушены. Образуются как линейные пептиды, так и пептиды сложной формы — разветвленные и циклические [9].

Расчет в рамках рассматриваемой модели и условий LC-эксперимента [9] показывает удовлетворительную корреляцию между экспериментальным и теоретическим порядком выхода этих фрагментов (рис. 4, темные точки). Однако, если заменить разветвленные и циклические структуры их линейными аналогами, имеющими тот же набор аминокислот, то расчетные времена мало изменятся (светлые точки), а корреляция будет не менее хорошей. Следовательно, для коротких пептидов главную роль в разделении играет их аминокислотный состав, а топология цепи мало значима. Это не удивительно, так как короткие цепи адсорбируются практически целиком, и роль петель, энтропия которых определяет зависимость от последовательности, незначительна.

Еще один пример для рассмотрения — однопечный белок IGF-I, сшитый дисульфидными мостиками (рис. 5). При его гидролизе протеазой V8, наряду с линейными, также образуются разветвленные и циклические структуры [9].

Как и в предыдущем случае, расчет в целом правильно предсказывает порядок выхода фрагментов. Если снова заменить сложные разветвленные структуры обычными линейными цепочками с тем же набором аминокислот, то результат корреляции получается не хуже.

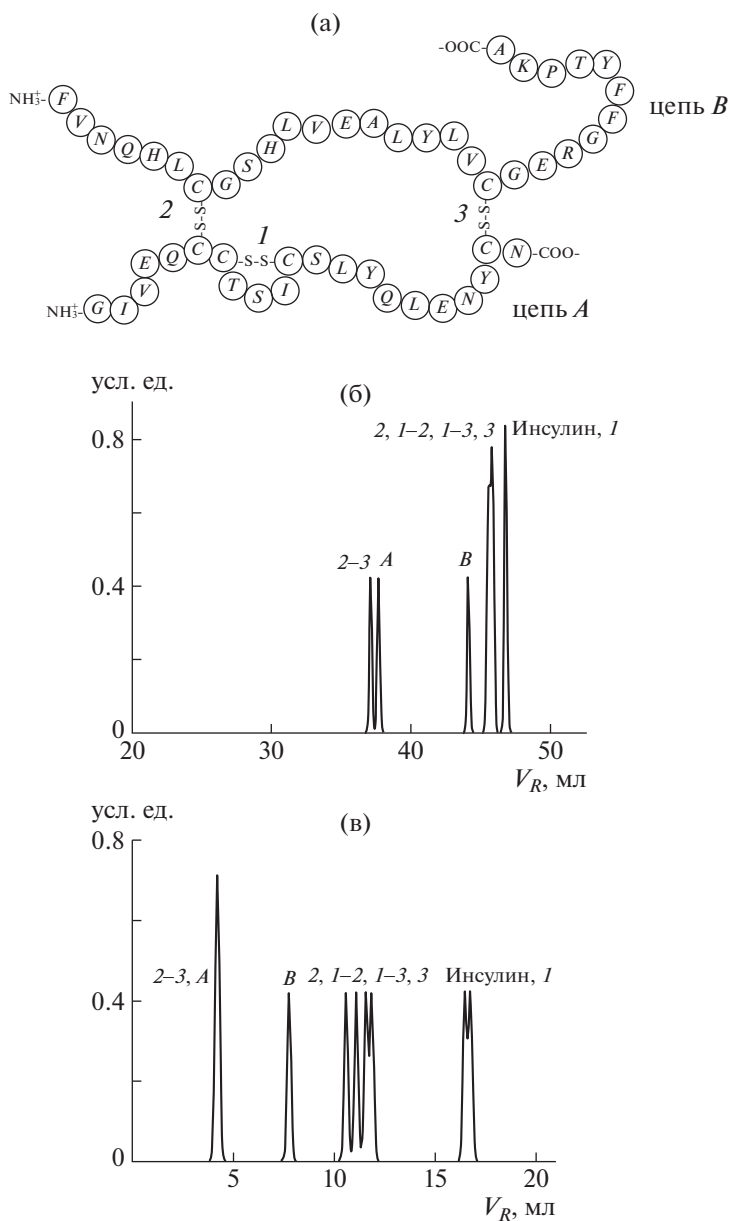


Рис. 6. Схема макромолекулы Инсулин (а), а также модельные хроматограммы в градиентном (б) и изократическом (в) режимах. Цифрами обозначены макромолекулы, получающиеся после однократного (1, 2, 3) и двухкратного (1–2, 1–3, 2–3) расщепления соответствующих S–S-сшивков; А, В – отдельные цепи.

тические условия (рис. 6б), однако и в этом случае полного разделения всех шести структур вряд ли можно добиться.

ЗАКЛЮЧЕНИЕ

Модель хроматографического разделения, основанная на представлении макромолекулы в виде цепи блужданий, допускает несложное обобщение на случай наличия в ней разных типов вторичных структур: спиральных участков, складок, S–S-сшивков. Для практического применения эту

модель необходимо дополнить набором значений эффективной энергии аминокислотных остатков в спиральном или складчатом состоянии, а также S–S-сшивков. Это не простая экспериментальная задача, так как для ее реализации необходимо иметь ряд моделей пептидов в заведомо известном (например, спиральном) состоянии в условиях хроматографического эксперимента. Учет влияния вторичных структур, возможно, поможет в интерпретации данных по хроматографическому разделению денатурированных белков, не подвергнутых ферментативному гидролизу.

Что касается пептидов, обычно используемых в протеомных экспериментах для идентификации белков, то для них усложнение модели, учитывающей влияние вторичных структур, оказывается малозначимым. Для интерпретации хроматографического разделения пептидов, в дополнение к их анализу методами масс-спектрометрии, вполне достаточно модели клубка и простых аддитивных моделей [1–5].

Работа выполнена в соответствии с Государственным заданием 0082-2019-0003 (№ АААА-А20-120021090129-9).

СПИСОК ЛИТЕРАТУРЫ

1. Gorshkov A.V., Evreinov V.V., Tarasova I.A., Gorshkov M.V. // Polymer Science B. 2007. V. 49. № 3–4. P. 93.
2. Guo D., Mant C.T., Taneja A.K., Parker J.M.R., Hodges R.S. // J. Chromatogr. 1986. V. 359. P. 499.
3. Krokhin O.V. // Anal. Chem. 2006. V. 78. № 22. P. 7785.
4. Reimer J., Spicer V., Krokhin O.V. // J. Chromatogr. A. 2012. V. 1256. P. 160.
5. Gilar M., Xie H.W., Jaworski A. // Anal. Chem. 2010. V. 82. № 1. P. 265.
6. Жулина Е.Б., Горбунов А.А., Скворцов А.М. // Высокомолек. соед. А. 1984. Т. 26. № 5. С. 915.
7. Gorshkov A.B., Goloborod'ko A.A., Pridatchenko M.L., Tarasova I.A., Rozdina I.G., Evreinov V.V., Gorshkov M.V. // Polymer Science A. 2017. V. 59. № 3. P. 354.
8. <http://bioinf.cs.ucl.ac.uk/psipred>
9. Canova-Davis E., Kessler T.J., Ling V.T. // Anal. Biochem. 1991. V. 196. № 1. P. 39.