

УДК 541.64:543.544

О ПРИМЕНИМОСТИ КОНЦЕПЦИИ КРИТИЧЕСКОЙ ХРОМАТОГРАФИИ К ЗАДАЧАМ ПРОТЕОМИКИ. ЗАВИСИМОСТЬ ВРЕМЕНИ УДЕРЖИВАНИЯ ОТ ПОСЛЕДОВАТЕЛЬНОСТИ АМИНОКИСЛОТНЫХ ОСТАТКОВ В ЦЕПИ¹

© 2007 г. А. В. Горшков*, В. В. Евреинов*, И. А. Таракова**, М. В. Горшков**

*Институт химической физики им. Н.Н. Семёнова
Российской академии наук
119991 Москва, ул. Косыгина, 4

**Институт энергетических проблем химической физики
Российской академии наук
119334 Москва, Ленинский пр., 38, корп. 2

Поступила в редакцию 18.05.2006 г.
Принята в печать 10.11.2006 г.

Предложена модель разделения пептидов, основанная на методе критической хроматографии макромолекул. В рамках модели рассмотрена совокупность экспериментальных данных по разделению пептидов. Определены основные феноменологические параметры модели – эффективные энергии адсорбции аминокислотных остатков, что позволило предсказать влияние характера их чередования в цепи на времена удерживания. Модель применена для оценки возможности разделения в разных хроматографических режимах пептидов с одинаковым аминокислотным составом, различающихся последовательностью расположения звеньев в цепи, а также содержащих изомеры аминокислот, зеркальные последовательности с разными концевыми группами.

ВВЕДЕНИЕ

Одной из задач химии биомакромолекул и протеомики является определение первичной структуры цепи, включая как саму последовательность аминокислотных остатков, так и наличие в ней в определенных местах функциональных или модифицированных групп, “перестановку” аминокислот и другие особенности строения. Аналогичная задача, существующая в химии синтетических полимеров, в настоящее время успешно решается применением метода критической хроматографии в сочетании с масс-спектрометрическим детектированием. Возможно ли применение идей и методов критической хроматографии к задаче определения последовательности

биомакромолекул и, если возможно, то какие результаты следует ожидать? Эти вопросы обсуждаются в настоящей работе на примере разделения пептидов – относительно коротких макромолекул, состоящих из 10–50 аминокислотных остатков. В рамках модели критической хроматографии рассмотрена совокупность экспериментальных данных по разделению пептидов на обращенной фазе и определены основные феноменологические параметры модели – эффективные энергии адсорбции аминокислотных остатков. Таким образом оказывается возможным теоретически предсказать времена (объемы) удерживания для пептида с любой заданной последовательностью и оценить применимость критической хроматографии для решения задач разделения пептидов с различными первичными последовательностями, разным числом, типом и местом модифицированных (функциональных) групп в цепи, пептидов, содержащих изомеры лейцина (L) и изолейцина (I), а также зеркальных последовательностей. Экспериментальная проверка предсказаний модели, а также объединение

¹ Работа выполнена при финансовой поддержке Отделения химии и наук о материалах РАН (программа “Создание эффективных методов химического анализа и исследования структуры веществ и материалов”), а также Российского фонда фундаментальных исследований (код проекта 06-04-49632) и фонда INTAS (гранты 04-83-2643 и Genomics 05-10000004-7759).

E-mail: evreinov@polymer.chph.ras.ru (Евреинов Виктор Викторович).

хроматографических данных с масс-спектрометрическими в рамках единого комплекса для "чтения" последовательностей пептидов и других биомакромолекул будут предметом дальнейших исследований.

ОТЛИЧИЕ АДСОРБЦИИ МАКРОМОЛЕКУЛ ОТ АДСОРБЦИИ НИЗКОМОЛЕКУЛЯРНЫХ ВЕЩЕСТВ

Основной измеряемой величиной в хроматографии макромолекул является объем удерживания V_R , определяемый коэффициентом распределения K_d , равным отношению концентраций макромолекул в подвижной фазе объемом V_0 и в порах объемом V_p :

$$V_R = V_0 + K_d V_p \quad (1)$$

Коэффициент распределения в свою очередь зависит от изменения свободной энергии макромолекулы ΔF при ее попадании из межчастичного пространства в пору неподвижной фазы

$$K_d = \frac{C_p}{C_0} = \frac{V_0}{V_p} e^{-\Delta F/kT} = \frac{V_0 Z_p}{V_p Z_0}$$

и, следовательно, определяется отношением статистических сумм макромолекул в поре Z_p и в межчастичном "неограниченном" пространстве Z_0 .

В случае градиента состава растворителя коэффициент распределения зависит от объема V прокачанного через колонку растворителя $K_d = K_d(V)$. При этом считается, что при медленном изменении состава элюента локальное равновесие в колонке успевает установиться, и для определения объема удерживания можно применять квазиравновесный подход. Тогда объем удерживания определяется следующим образом:

$$\int_0^{V_R - V_0} \frac{dV}{V_p K_d(V)} = 1 \quad (2)$$

Чтобы количественно описать разделение макромолекул, необходимо задать зависимость $K_d(V)$. Подобный расчет может быть сделан в рамках простой модели адсорбции цепи. Тем са-

мым оказывается возможным теоретически предсказать характер разделения макромолекул.

Прежде чем перейти к рассмотрению модели адсорбции и экспериментальному определению необходимых для расчета параметров, резюмируем качественные различия адсорбции макромолекул по сравнению с адсорбцией низкомолекулярных соединений.

Принципиальным отличием макромолекул от низкомолекулярных соединений (мономеров) является связанность звеньев в цепь [1]. В процессах адсорбции звенья цепи взаимодействуют с поверхностью коллективно. Это взаимодействие отражает не только индивидуальные свойства (химическое строение) звена, но и коллективные свойства всей цепи (ее последовательность). Если адсорбция несвязанных мономеров проистекает плавно в достаточно широком диапазоне температур, то для макромолекул она осуществляется скачком в узком температурном диапазоне вблизи критической температуры T_{cr} , при которой потеря энтропии связи вблизи поверхности компенсируется энергией притяжения. Если $T > T_{cr}$, то макромолекулы находятся в растворе, и их доля на поверхности исчезающе мала, при $T < T_{cr}$ они резко локализуются на поверхности. Адсорбция макромолекулы носит характер фазового перехода в системе связанных в цепь звеньев. Связанность звеньев в цепь, называемая "линейной памятью", означает, что равновесное состояние по отношению к адсорбционному взаимодействию реализуется с учетом ограничений, накладываемых ковалентными химическими связями. Поскольку энергия этих связей намного превышает энергию адсорбционного взаимодействия мономерного звена, в процессе адсорбции цепь не разрушается, следовательно, при адсорбции должен проявляться и характер связаннысти звеньев.

Равновесные свойства адсорбирующейся макромолекулы определяют объем (время) удерживания в хроматографии. Функция распределения, описывающая адсорбционное взаимодействие низкомолекулярных веществ $e^{-\epsilon(\Gamma)/kT}$ ($\epsilon(\Gamma)$ – энергия взаимодействия с поверхностью в состоянии Γ), для макромолекул заменяется на $e^{-\epsilon(\Gamma)/kT} \prod_i g_i$, где функции g_i – связи между соседними звеньями.

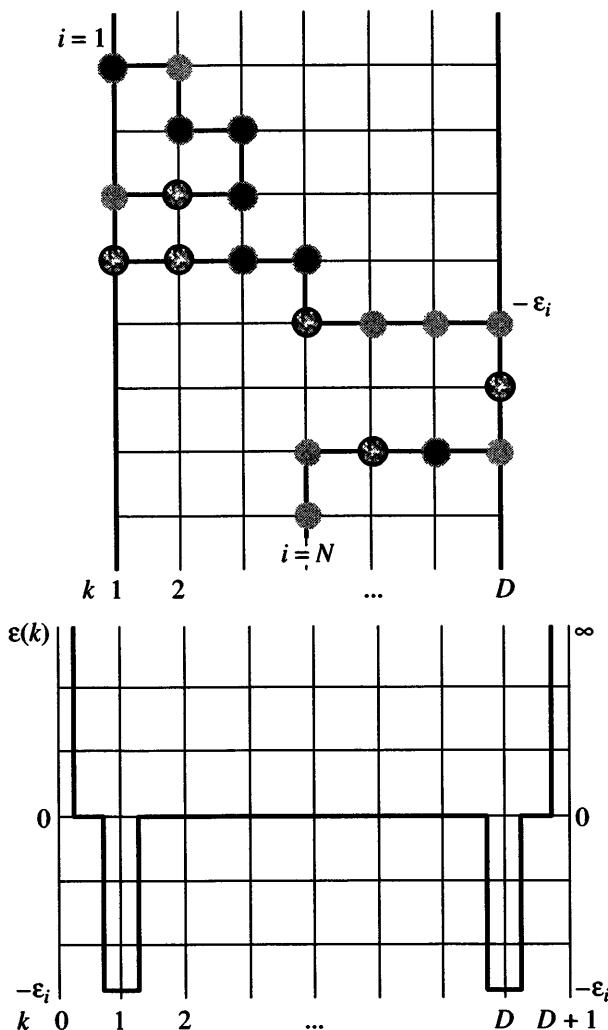


Рис. 1. Решеточная цепь случайных блужданий в щелеобразной поре как модель хроматографического разделения макромолекул пептидов.

Тогда статистическая сумма адсорбирующейся макромолекулы представляется в виде

$$Z = \int e^{-\epsilon(\Gamma)/kT} \prod_i g_i d\Gamma = \int \prod_i e^{-\epsilon_i/kT} g_i d\Gamma \quad (3)$$

Специфическое сильное взаимодействие между мономерами цепи за счет химических ковалентных связей и связанные с ним ограничения в реализации равновесного состояния макромолекулы играют большую роль и в других процессах с участием макромолекул, таких, например, как переход клубок–глобула. Для описания этих явлений применяются современные методы физики конденсированного состояния, изложенные в работах [2–4].

Несмотря на сложность адсорбции макромолекул, ее можно количественно описать с помощью простой, но не тривиальной модели. Как и в теории растворов, наиболее наглядной оказывается решеточная модель, в которой макромолекула представляется цепью случайных блужданий [5]. Данная модель дает, безусловно, приближенное описание реальных макромолекул. Однако она с успехом применяется для описания закономерностей хроматографии полимеров [6, 7], в том числе и для практического поиска условий, наилучших для их разделения по тем или иным структурным особенностям. Оказывается, что при соответствующем экспериментальном определении феноменологических параметров, описывающих адсорбционное взаимодействие звеньев цепи, результаты расчетов близки к экспериментальным данным [8].

МОДЕЛЬ ХРОМАТОГРАФИИ МАКРОМОЛЕКУЛ И ЕЕ СЛЕДСТВИЯ

В решеточной модели раствора непрерывное пространство заменяется, например, простой кубической решеткой, узлы которой заняты либо молекулами растворителя, либо звеньями цепи (рис. 1). Адсорбент представляется в виде набора щелеобразных пор размера D , а каждой конфигурации макромолекулы ставится в соответствие цепь блужданий по узлам решетки. Если звенья цепи попадают на поверхность, то они взаимодействуют с ней с энергией $-\epsilon$, что приводит к множителю $e^{\epsilon/kT}$ в конфигурационном интеграле. В силу симметрии задачи имеет смысл отслеживать блуждания цепи только по координате x , перпендикулярной адсорбционным поверхностям, т.е. модель, по сути, одномерна. Как показывает эксперимент, для синтетических полимеров выбор щелеобразной поры хорошо соответствует реальной картине разделения. По-видимому, истинная фрактальная размерность порового пространства адсорбентов на основе силикагеля близка к двум.

Связанность звеньев в цепь описывается переходной матрицей (оператором) $W(\epsilon)$, элементы которой имеют смысл условной вероятности попадания звена i цепи в точку x_i при условии, что звено $i-1$ находилось в точке x_{i-1} :

$$W(\epsilon) = \begin{bmatrix} 2/3\epsilon & 1/6\epsilon & 0 & 0 & \dots & \dots \\ 1/6 & 2/3 & 1/6 & 0 & \dots & \dots \\ 0 & 1/6 & 2/3 & 1/6 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & 1/6 & 2/3 & 1/6 \\ \dots & \dots & 0 & 0 & 1/6\epsilon & 2/3\epsilon \end{bmatrix}_{D \times D} \quad (4)$$

Для простоты будем считать энергию ϵ выраженной в единицах kT . Размерность матрицы (в единицах размера звена) определяется шириной поры. Обычно принимают размер звена $a \approx 10 \text{ \AA}$, так что для реальных адсорбентов размерность переходной матрицы невелика: $D = 10-30$.

Для цепи длиной N функция распределения последнего звена в цепи по координатам внутри поры имеет вид

$$\mathbf{P}_N = W^N \mathbf{P}_0, \quad (5)$$

где $\mathbf{P}_0 = \mathbf{U}$ – стартовый вектор, все компоненты которого равны единице, описывающий равновероятное начало цепи в любой точке внутри поры. Прямым перемножением матриц можно убедиться, что компоненты вектора $\mathbf{P}_N(x_k)$ действительно содержат все возможные конфигурации цепи, начинающейся внутри поры и заканчивающейся в точке x_k . Для нахождения статистической суммы Z_p необходимо просуммировать \mathbf{P}_N по всем точкам x_k – координатам конца цепи:

$$Z_p = \mathbf{U}^T W^N \mathbf{P}_0 \quad (6)$$

Очевидно, что это есть не что иное, как запись конфигурационного интеграла (3) в дискретном пространстве узлов решетки (\mathbf{U} – единичный вектор). Статистическая сумма цепи нормирована таким образом, что в неограниченном пространстве (межчастичном пространстве объемом V_0) $Z_0 = 1$ для любой длины цепи N . С учетом нормировки для получения коэффициента распределения K_d необходимо Z_p в уравнении (6) разделить на размер поры D , выраженный в единицах размера звена. Тогда

$$K_d = e^{-\Delta F/kT} = \frac{1}{D} Z_p \quad (7)$$

Таким образом, модель случайных блужданий позволяет элементарно рассчитать коэффициент

распределения и предсказать закономерности хроматографического разделения макромолекул в соответствии с основным уравнением (1). Значения V_0 и V_p – константы колонки, объем межчастичного пространства (подвижной фазы) и суммарный объем пор (неподвижной фазы). Экспериментально определяемый объем удерживания V_R есть, по сути, статистическая сумма макромолекулы в колонке. В отличие от времени удерживания, объем удерживания имеет глубокий физический смысл. На практике обычно используют время удерживания t_R , которое зависит от объемного расхода растворителя и пропорционально V_R . В дальнейшем мы будем говорить в основном об объемах удерживания.

КРИТИЧЕСКИЕ УСЛОВИЯ И ЕДИНЫЙ МЕХАНИЗМ ХРОМАТОГРАФИИ МАКРОМОЛЕКУЛ

Напомним, какие следствия для хроматографии полимеров вытекают из предложенной модели в важнейших частных случаях. Если взаимодействие звеньев цепи с поверхностью равно нулю ($\epsilon = 0$), то, очевидно, реализуется эксклюзионный режим хроматографии макромолекул. Это непосредственно следует из формы переходной матрицы в уравнении (4). При $\epsilon = 0$ сумма элементов матрицы в первой и последней строке меньше единицы, и их последовательное перемножение приводит к тому, что компоненты вектора $\mathbf{P}_N(x_k)$ уравнения (5) вблизи стенок быстро стремятся к нулю. Следовательно, при увеличении длины цепи часть пространства поры вблизи стенок оказывается “исключенной”, и калибровка (в стандартных координатах $\lg N - V_R$) характерна эксклюзионному режиму разделения (рис. 2). С ростом энергии отталкивания $\epsilon \rightarrow -\infty$ картина практически не меняется, за исключением того, что эффективный размер поры становится равным $D - 2$, что несущественно для широких пор: $D \gg 1$.

Другой предельный случай (адсорбционный режим) соответствует сильной энергии притяжения, такой, что цепь оказывается адсорбированной на какой-либо из поверхностей поры. В адсорбционном режиме статистическая сумма, коэффициент распределения K_d и объем удерживания V_R экспоненциально сильно зависят от длины цепи. Это также можно заметить из

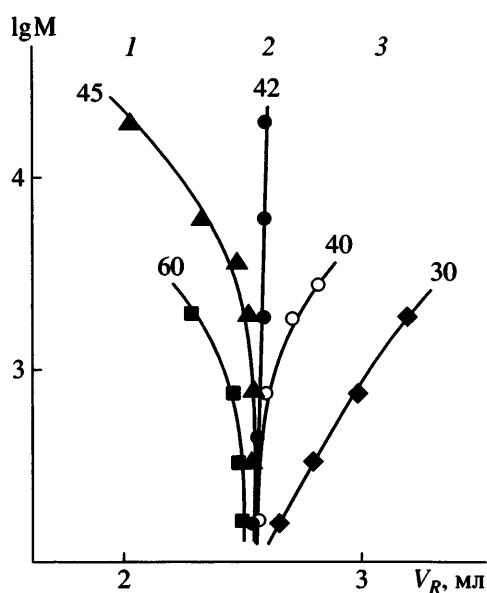


Рис. 2. Режимы хроматографии макромолекул на примере полиэтиленоксида: переход от эксклюзионного (1) к адсорбционному (3) режиму через критический режим (2) на обращенной фазе $-C_{18}$ в бинарном растворителе вода–ацетонитрил при изменении его состава. Цифры у кривых – содержание компонента В, об. %.

формы переходных матриц: при их перемножении основной вклад будут давать экспоненциально большие члены. В пределе $\varepsilon \rightarrow \infty$ каждое звено цепи взаимодействует с поверхностью.

Кроме этих двух режимов, обычных для хроматографии макромолекул, существует еще один, особый режим разделения, называемый критическим. Из вида уравнения (4) легко заметить, что при энергии взаимодействия звена с поверхностью $\varepsilon_{cr} = \ln(6/5)$ (при компенсации потерь энтропии энергией притяжения) суммы компонентов в первой и последней строке W равны единице. Следовательно, равномерное распределение P_0 сохраняется для цепей любой длины, статистическая сумма при этом в точности равна объему поры D , коэффициент распределения равен единице, $K_d = 1$, а объем удерживания

$$V_R = V_0 + V_p \quad (8)$$

В критическом режиме размер макромолекул и ММР как бы “исчезают”, что создает наилучшие условия для исследования других, отличных от ММР, типов структуры макромолекул: числа и типа функциональных групп в цепи, ее топологии, распределения по составу и архитектуре со-

полимеров. Для синтетических полимеров ММР существует всегда, и в эксклюзионном или адсорбционных режимах маскирует какие-либо другие структурные особенности цепи. Для биомакромолекул ММР отсутствует, так что в определенном смысле хроматография биополимеров проще: возможности эксклюзионного и адсорбционного режимов разделения для исследования их структуры также реализуемы.

Критическая хроматография была введена в практику хроматографии в работе [9] и в настоящее время широко используется для разделения макромолекул по разным типам их структуры [10–13]. Особенности критического режима есть следствие масштабной инвариантности системы вблизи критической точки адсорбции. Они имеют глубокую связь с общей теорией фазовых переходов [14], которую в рамках настоящей работы обсуждать не представляется возможным.

Критический режим объединяет эксклюзионный и адсорбционный режимы в рамках единого механизма хроматографии. Как показано в настоящее время в работах по хроматографии полимеров, рассматриваемая модель адсорбции цепи случайных блужданий (или ее аналог, сводящийся к уравнению теплопроводности) адекватно описывает основные закономерности разделения макромолекул.

Напомним также, как проявляются особенности адсорбции макромолекул при их градиентном элюировании, тем более что обычно разделение биомакромолекул происходит в условиях градиента состава растворителя.

Скачкообразный переход макромолекул в адсорбированное состояние значительно упрощает картину их разделения. Обычно градиент растворителя осуществляется таким образом, что эффективная энергия взаимодействия звена с поверхностью уменьшается с увеличением объема прокачанного растворителя. До тех пор, пока энергия взаимодействия звена больше критической ($\varepsilon > \varepsilon_{cr}$), макромолекула экспоненциально сильно адсорбируется и практически не движется вдоль колонки. Как только величина энергии уменьшается до критического значения, макромолекула резко десорбируется и перемещается по колонке со скоростью растворителя. Таким образом, градиентная хроматография макромолекул неявно предполагает переход через кри-

тическую точку адсорбции. Следовательно, особенности критической точки также важны для понимания закономерностей градиентной хроматографии макромолекул. Эти закономерности заметно упрощаются для узких пор: в данном случае существенным становится не характер чередования различных мономеров, а их количество в цепи.

В градиенте растворителей макромолекулы распределяются по составам смесей, в которых реализуются соответствующие критические энергии, и выходят в узких зонах вблизи своих критических точек. Критические энергии зависят от химического строения звеньев цепи. С учетом ступенчатого характера перехода в адсорбированное состояние даже макромолекулы, весьма близкие по строению звеньев (например, изомеры), в градиенте растворителя могут разделяться. Градиент используется и для быстрого поиска критических точек адсорбции макромолекул в различных хроматографических системах.

ЭФФЕКТИВНАЯ ЭНЕРГИЯ АДСОРБЦИИ ЗВЕНА

Основным феноменологическим параметром представленной выше модели является микроскопическая эффективная энергия взаимодействия звена цепи (для пептидов – аминокислотного остатка). В общем случае эта энергия есть функция химического строения звена, поверхности адсорбента, химического строения и состава растворителя и температуры колонки. Ясно, что рассчитать такую энергию исходя из первых принципов методами квантовой химии и молекулярной динамики вряд ли возможно. Однако эффективная энергия может быть определена экспериментально в рамках корреляционной теории, например теории Снайдера [15]. Подробное описание применения теории Снайдера для моделирования разделения олигомеров с функциональными группами дано в работах [10, 11], ниже сформулированы краткие принципы такого подхода. Для оценки эффективной энергии можно использовать любую другую корреляционную теорию, ее связь с приведенной выше формальной теорией адсорбции макромолекул осуществляется аналогично. Вместе с тем теория Снайдера имеет в своей основе простую лэнгмюровскую модель адсорбции компонентов растворителя и

весьма просто “вписывается” в рассматриваемую решеточную модель адсорбции цепи.

Имея в виду приложение модели к разделению пептидов, в первую очередь необходимо определить зависимость эффективной энергии от состава бинарного растворителя. Для простоты зафиксируем тип адсорбента (обращенная фаза типа $-C_{18}$) и компоненты бинарного растворителя (для пептидов обычно это вода и ацетонитрил с добавками 0.1–0.2% трифторуксусной кислоты, так, чтобы pH составлял 2.0).

В силу малости средней концентрации мономеров в клубке весьма эффективным оказывается применение приближения системы разорванных звеньев [1]. Энергия взаимодействия звена с поверхностью принимается такой же, как и для его низкомолекулярного аналога. Эта энергия ϵ отражает равновесие в многокомпонентной системе звено–поверхность–растворитель. В соответствии с теорией Снайдера ее можно представить в виде

$$\epsilon = X^0 - \epsilon_0(N_B) \quad (9)$$

Здесь X^0 – абсолютная энергия взаимодействия звена с поверхностью, $\epsilon_0(N_B)$ – энергия десорбции растворителя (обычно называемая его элюирующей силой), зависящая от состава N_B (мольной доли компонента B бинарного растворителя A – B). Уравнение (9), основанное на пренебрежении взаимодействием звеньев с растворителем и молекул растворителя друг с другом, достаточно хорошо обосновано для полярных адсорбентов типа силикагеля. Для обращенной фазы возможно, что такое упрощение не столь хорошо обосновано. Однако надо иметь в виду, что в нашем случае корреляционная теория применяется в рамках упрощенной модели цепи. Усложнение только корреляционной теории и включение в нее взаимодействий в растворе без соответствующего усложнения, в том числе и модели цепи, вряд ли оправдано. Можно сказать, что выражение (9) для эффективной энергии просто постулируется, не вдаваясь в его физическое обоснование.

Рассмотрим цепь блужданий на решетке в бинарном растворителе с компонентами A (вода) и B (ацетонитрил). Модель предполагает короткодействующее (точечное) взаимодействие звена с поверхностью, ограниченное размером порядка

размера звена $a \approx 10 \text{ \AA}$. Попадание звена на поверхность вызывает десорбцию либо молекул А, либо молекул В.

В растворе вся поверхность занята молекулами растворителя. В силу малости средней концентрации звеньев в клубке можно предположить, что адсорбция звеньев макромолекулы слабо влияет на равновесное распределение компонентов растворителя на поверхности и в растворе. Тогда усредненная матрица перехода $\langle W \rangle$ может быть представлена в виде

$$\langle W \rangle = (1 - \theta_B)W(X^0 - \varepsilon_A) + \theta_B W(X^0 - \varepsilon_B),$$

где θ_B – степень заполнения поверхности компонентом В. Легко заметить, что $\langle W \rangle$ есть

$$\langle W \rangle = W(\varepsilon^{\text{eff}}) = W(X^0 - \varepsilon_{AB})$$

Здесь

$$\varepsilon_{AB} = \varepsilon_A - \ln[(1 - \theta_B) + \theta_B \exp(\varepsilon_A - \varepsilon_B)],$$

а ε_A и ε_B – соответствующие элюирующие силы “чистых” компонентов А и В. Следовательно, в бинарном растворителе цепь представляется блужданием по решетке с матрицей перехода, имеющей ту же структуру, что и в однокомпонентном растворителе, но с заменой элюирующей силы на эффективную величину ε_{AB} . Введем константу равновесия компонентов растворителя на поверхности:

$$K(P, T) = \frac{\theta_B(1 - N_B)}{(1 - \theta_B)N_B} = \exp(\varepsilon_B - \varepsilon_A)$$

Тогда для элюирующей силы смеси получаем

$$\varepsilon_{AB} = \varepsilon_A + \ln[1 - N_B + N_B \exp(\varepsilon_B - \varepsilon_A)] \quad (10)$$

Это есть уравнение Снайдера [15] для элюирующей силы бинарного растворителя. Таким образом, изменение состава растворителя приводит к изменению эффективной энергии адсорбции звена цепи. Корреляционная теория позволяет непрерывным образом описать переход от адсорбции к эксклюзии через критический режим при изменении состава растворителя и связанные с данным переходом изменения в характере разделения макромолекул.

Для применения корреляционной теории и описания характера разделения макромолекул при любых изменениях состава растворителя необходимо определить три основных феноменологических параметра: энергию адсорбции звена X^0 и элюирующие силы растворителей ε_A и ε_B . Для этого необходимо исследовать переход от эксклюзионного режима к адсорбционному при нескольких фиксированных составах бинарного растворителя. Фактически для определения параметров корреляционной теории достаточно трех составов, один из которых соответствует критическому режиму, а два других – адсорбционному и эксклюзионному. Критический состав N_{cr} дает простейшее уравнение, связывающее феноменологические параметры модели:

$$\varepsilon_{cr} = X^0 - \varepsilon_{AB}(N_{cr}) = \ln 6/5 \approx 0.18$$

При этом $K_d = 1.0$ и $V_R = V_0 + V_p$. Для других составов необходимо рассчитывать V_R по уравнениям (1), (7), (10) и сравнивать их с экспериментально наблюдаемыми величинами. Варьируя X^0 , ε_A , ε_B и добиваясь наилучшего совпадения теории и эксперимента, оказывается возможным определение параметров модели, необходимых для описания с хорошей точностью перехода от эксклюзии к адсорбции. Знание этих параметров дает возможность предсказывать особенности разделения макромолекул известной структуры в данном бинарном растворителе А и В при любом его составе и при любом характере изменения состава во времени (градиенте).

Рассмотрим, как эту модель можно применить для предсказания времен удерживания пептидов с разной последовательностью.

ГЕТЕРОПОЛИМЕР И ЭФФЕКТИВНЫЕ ЭНЕРГИИ АДСОРБЦИИ АМИНОКИСЛОТНЫХ ОСТАТКОВ

Принципиальным отличием биомакромолекул от макромолекул синтетических полимеров является гетерогенность их цепей и отсутствие ММР. В рамках предложенной модели гетерогенность цепи учитывается следующим образом:

$$Z = \mathbf{U}^T W(\varepsilon_N) W(\varepsilon_{N-1}) \dots W(\varepsilon_2) W(\varepsilon_1) \mathbf{P}_0 \quad (11)$$

Здесь $W(\varepsilon_i)$ – переходная матрица i аминокислотного остатка, имеющего энергию взаимодействия

ствия с поверхностью ϵ_i . Для определения коэффициента распределения теперь необходимо знать не только степень полимеризации пептида, но и последовательность аминокислотных остатков в его цепи, его "текст" $\{S\}$.

Уже сама форма записи статистической суммы в виде (11) позволяет сделать принципиальный вывод о ее зависимости от последовательности составляющих цепь аминокислотных остатков. Переходные матрицы $W(\epsilon_i)$ некоммутативны:

$$W_i W_j \neq W_j W_i$$

и цепи, имеющие абсолютно одинаковый аминокислотный состав, но различающиеся характером чередования звеньев, могут иметь разные коэффициенты распределения. Следовательно, предлагаемая модель органически содержит в себе различие адсорбционных свойств биомакромолекул, отличающихся последовательностью составляющих их аминокислотных остатков.

Запись эффективной энергии взаимодействия ϵ_{AB} в виде уравнения (10) и "расцепление" свойств звена и растворителя позволяют применить корреляционную теорию и для гетерополимеров. Для описания разделения биомакромолекул по их первичной последовательности $\{S\}$ необходимо задать последовательность феноменологических параметров $\{X_i^0\}$, характеризующих эффективное взаимодействие аминокислотных остатков цепи с поверхностью (будем называть это адсорбционным энергетическим профилем последовательности). Указанные параметры определяются экспериментально, после чего модель позволяет оценить объем удерживания пептида с любым известным "текстом" $\{S\}$.

Наиболее простой способ определения эффективных энергий адсорбции аминокислотных остатков – исследование перехода от эксклюзии к адсорбции и определение критических точек для синтетических гомополимеров полиаминокислот. Однако далеко не все полиаминокислоты можно синтезировать, кроме того, многие из них не растворимы в смеси вода–ацетонитрил. Тем не менее, достаточно лишь для одной растворимой полиаминокислоты i экспериментально определить эффективную энергию адсорбции X_i^0 остат-

ка и константы растворителей ϵ_A и ϵ_B . Затем, заменяя в этой последовательности всего одну, например, крайнюю аминокислоту, можно найти эффективные энергии взаимодействия $\{X_i^0\}$ и для всех других аминокислотных остатков.

Идейно такая постановка задачи аналогична задаче об адсорбции макромолекул с одной концевой функциональной группой. В критических условиях для растворимой полиаминокислоты, имеющей в последовательности на конце N другую аминокислоту с энергией адсорбции ϵ_N , коэффициент распределения K_d однозначно связан с разностью эффективных энергий взаимодействия этих аминокислот (см. выражение (6)):

$$\begin{aligned} K_d &= \frac{1}{D} (\mathbf{U}^T W_N(\epsilon_N) W^{N-1}(\epsilon_i = \epsilon_{cr}) \mathbf{P}_0) = \\ &= \frac{1}{D} (\mathbf{U}^T W_N(\epsilon_N) \mathbf{P}_0) = 1 + \frac{2}{D} (e^{\frac{X_N^0 - X_i^0}{D}} - 1) \end{aligned} \quad (12)$$

В предлагаемой корреляционной теории параметры растворителя ϵ_A , ϵ_B не зависят от свойств макромолекулы, они предполагаются одинаковыми для макромолекул любой природы. Эти параметры могут быть определены вообще независимо от пептидов, например, по исследованию перехода от эксклюзии к адсорбции макромолекул полиэтиленоксида.

Для завершения построения корреляционной теории необходимо учесть различие адсорбционных свойств адсорбентов одного типа (например $-C_{18}$) в разных колонках или одного и того же адсорбента (одной колонки) со временем. Ясно, что активность колонки со временем меняется, что в обычной хроматографии низкомолекулярных соединений учитывается коррекцией наблюдаемых объемов удерживания на объем удерживания известного стандартного вещества. Такой же подход применим и для макромолекул: для эффективной энергии взаимодействия можно записать

$$\epsilon = \alpha(X^0 - \epsilon_{AB})$$

Параметр α интерпретируется, аналогично теории Снайдера, как "активность" адсорбента и определяется по времени удерживания стандартного пептида с известной последовательностью аминокислотных остатков.

Кроме эффективных энергий адсорбции аминокислотных остатков в корреляционную теорию необходимо включить энергию адсорбции концевых групп $X_{N,C}^0$, а также энергию адсорбции модифицированных аминокислотных остатков X_m^0 для предсказания времен удерживания модифицированных пептидов.

Прежде чем приступить к проверке, насколько адекватно предлагаемая модель описывает существующие экспериментальные данные по хроматографии пептидов, несколько переформулируем ее. Примем, что цепь пептида начинается справа налево (с С-конца к N-концу). Будем считать первым мономером решеточной цепи первый аминокислотный остаток с С-конца. Тогда “стартовый” вектор цепи P_0 , описывающий распределение первого звена в поре, будет иметь вид

$$P_0 = \begin{bmatrix} e^{\epsilon_1} \\ e \\ 1 \\ 1 \\ \dots \\ e^{\epsilon_1} \end{bmatrix}$$

Статистическая сумма пептида с заданным “текстом” $\{S\}$ общей длиной N запишется следующим образом:

$$Z_p(\{S\}) = U^T \prod_{i=2}^N W(\epsilon_i) P_0 \quad (13)$$

Обычно в корреляционной теории выбирается наиболее слабый растворитель (пентан для нормальной фазы или вода для обращенной фазы), для которого элюирующая сила полагается равной нулю ($\epsilon_A = 0$). Такая нормировка ничего не меняет, поскольку на изменение элюирующей силы от состава растворителя влияет лишь разность $\epsilon_B - \epsilon_A$. Соответствующей перенормированной энергией взаимодействия можно всегда выбрать элюирующую силу одного из компонентов (более слабого) равной нулю. Будем считать такую нормировку выполненной, так что элюирующая сила слабого компонента – воды $\epsilon_A = 0$.

Наконец, для определения объема удерживания при градиенте состава раствора интеграл (2) удобно записать в дискретном виде, аппрокси-

мируя непрерывную функцию $K_d(V)$ ступенчатой $K_d(V_i)$ где V_i – порция растворителя постоянного состава:

$$V_R - V_0 = \sum_{i=1}^{n-1} V_i + K_d(V_n) V_p \left(1 - \sum_{i=1}^{n-1} \frac{V_i}{K_d(V_i) V_p} \right) \quad (14)$$

Здесь n – номер порции растворителя, в котором макромолекула выходит из колонки, определяемый условием

$$\sum_{i=1}^{n-1} \frac{1}{K_d(V_i) V_p} < 1 < \sum_{i=1}^n \frac{1}{K_d(V_i) V_p}$$

Применим модель для описания экспериментальных данных по разделению пептидов, представленных в серии работ [16–19].

В этих работах также предложен корреляционный подход для предсказания объемов удерживания пептидов. Он основывается на принципе аддитивности вкладов аминокислотных остатков во время удерживания пептида:

$$t_R = \sum_i n_i C R_i + t_0, \quad (15)$$

где $C R_i$ – коэффициент удерживания i -го типа аминокислотного остатка, число которых n_i (степень полимеризации пептида $N = \sum_i n_i$). Параметр t_0 позволяет учесть небольшие различия в свойствах адсорбентов (активность) и определяется по стандартам S₁–S₅ (Peptide retention standards, Pierce Biotechnology Inc.). Коэффициенты удерживания находятся экспериментально по модельным синтетическим пептидам для обращенной фазы, бинарного растворителя вода (A) – ацетонитрил (B) с добавками трифтормукусной кислоты и pH 2.0, в градиенте 0–100 об. % В с наклоном 1 об. % В/мин. Стандартная скорость подачи раствора 1 мл/мин. Теоретические расчеты, представленные ниже, предполагают эти условия, если не оговаривается иное.

Уравнение (15) соответствует приближению к описанию разделения пептидов “со стороны” низкомолекулярных соединений. Очевидно, что оно не содержит информацию о последовательности цепи $\{S\}$, а только лишь о ее аминокислотном составе. Такое приближение достаточно хорошо

обосновано для коротких пептидов, которые адсорбируются на поверхности всеми звеньями.

При увеличении длины цепи аддитивность должна нарушаться, что связано с "включением" энтропии цепи и появлением в конфигурации адсорбированной цепи петель и хвостов, не имеющих контакта с поверхностью. Для учета этого обстоятельства в уравнение (15) вводятся дополнительные корреляционные поправки [18], также не имеющие физического смысла. Более сложные варианты такого подхода не ограничиваются линейной частью, а используют в уравнении (15) полиномы и, следовательно, дополнительные наборы корректирующих коэффициентов.

Как будет показано ниже, наблюдающиеся экспериментально отклонения от аддитивности, возникающие при увеличении длины цепи, автоматически вытекают из предлагаемой нами модели и не требуют введения дополнительных поправок. Кроме того, она в отличие от уравнения (15) или его полиномиальных аналогов явно учитывает не только аминокислотный состав, но и последовательность цепи $\{S\}$.

Модель цепи блужданий соответствует принципу аддитивности в предельных случаях коротких цепей или очень сильной энергии адсорбции звеньев $\varepsilon_i \gg 1$. При этом цепь фактически находится в единственном основном состоянии, когда она полностью распластана на поверхности и взаимодействие зависит от ее аминокислотного состава, а не от последовательности. В градиенте растворителя при переходе от адсорбционного к эксклюзионному режиму даже короткая макромолекула не взаимодействует с поверхностью всеми звеньями. Из-за вклада других состояний цепи должна появляться зависимость ее времени удерживания от последовательности. Данный эффект в общем случае не очень велик, и целью работы является оценка его масштаба, который, очевидно, будет зависеть от конкретной последовательности звеньев в цепи. Такие оценки при корректном определении феноменологических параметров модели могут давать количественное представление о возможностях хроматографии для решения разных задач протеомики.

СРАВНЕНИЕ МОДЕЛИ С ЭКСПЕРИМЕНТАЛЬНЫМИ ДАННЫМИ

Опишем в рамках предлагаемой модели совокупность экспериментальных данных работ [16–19]. В расчетах по уравнениям (7), (13), (14) необходимо задать константы колонки V_0 и V_p , а также размер пор адсорбента D . Поскольку указанные параметры в работах [16–19] не определялись экспериментально, мы приняли для них следующие величины: $V_0 \approx V_p \approx 1.3$ мл для колонки 4.1×250 мм и ≈ 1.7 мл для колонок 4.6×250 мм. Размер пор D , выраженный в единицах звеньев, принят $D = 30$ для адсорбентов с порами 300 \AA и 10 для пор 100 \AA . Обозначения аминокислотных остатков, используемые ниже, соответствуют общепринятым.

Использованные в работе [16] модели пептидов Ac-GXXLLKK-amid, где X – аминокислотный остаток, позволяют из экспериментальных хроматографических данных по их разделению определить необходимую шкалу энергий X_i^0 .

Параметры растворителя $\varepsilon_A = 0$ и $\varepsilon_B = 2.40$ выбирались такими, чтобы совокупность экспериментальных времен удерживания модельных пептидов при найденных величинах X_i^0 , а также пептидов разной длины описывалась наилучшим образом. При этом энергии концевых групп Ac-(CH₃-CO-) и -amide (-CO-NH₂) принимались равными нулю ($\varepsilon_{Ac} = \varepsilon_{amide} = 0$), что является естественным приближением для взаимодействия этих групп с обращенной фазой.

Энергия X_i^0 определялась следующим образом. На первом этапе варьированием энергий X_K^0 , X_G^0 , X_L^0 добивались совпадения экспериментальных и расчетных объемов удерживания для трех основных моделей:

Ac-GGGLLLKK-amid, Ac-GLLLLLKK-amid,

Ac-GKKLLLKK-amid

Это дает возможность найти энергии адсорбции остатков аминокислот K, G, L, содержащихся в трех модельных соединениях. Затем, варьированием уже только одной энергии X_i^0 , из объемов удерживания оставшихся моделей определялись

все остальные энергии. Найденная таким образом шкала энергий, которая является основой для

всех последующих расчетов, приведена ниже и на диаграмме рис. 3.

K	H	R	N	G	S	Q	D	T	E
0.266	0.386	0.516	0.614	0.656	0.698	0.746	0.781	0.876	0.984
A	P	C	Y	V	M	I	L	F	W
1.143	1.143	1.296	1.686	1.751	1.822	2.156	2.298	2.319	2.436

Если пептиды имеют другие концевые группы ($\text{H}_2\text{N}-$, $-\text{COOH}$), то их наличие меняет энергию адсорбции концевых аминокислотных остатков: $X_N = X_N^0 + \epsilon_{\text{H}_2\text{N}-}$, $X_C = X_C^0 + \epsilon_{-\text{COOH}}$. Использованные в работе [16] модели Y-GLLLKK-Z, где Y, Z – разные типы концевых групп, без труда позволяют найти $\epsilon_{\text{H}_2\text{N}-} = -1.69$ и $\epsilon_{-\text{COOH}} = -0.3$.

Наконец, активность колонки α определялась из времени удерживания стандарта S_4 , который был использован в работах [16–19] также для подстройки времен удерживания.

Для начала рассмотрим зависимость объемов удерживания от длины цепи и рассчитаем объемы удерживания пептидов, приведенных в табл. 1 и на рис. 4. Как видно, модель дает правильную зависимость изменения объемов удерживания при увеличении длины цепи и весьма точно описывает экспериментальные данные работы [18] без дополнительных корреляционных коэффициентов, не имеющих физического смысла. Очевидно, что наблюдаящаяся нелинейность связана именно с энтропией цепи и вытекает из свойств модели.

Следует сказать, что пептиды, приведенные в табл. 1, – это, фактически, гомополимеры, хотя и со сложной структурой повторяющейся единицы. Поэтому для них, как и для обычных гомополимеров, должны существовать критические точки, в которых зависимость времен удерживания от длины исчезает. Для примера мы рассчитали критический состав воды и ацетонитрила, вблизи которого для этих пептидов должна пропадать зависимость разделения от длины цепи: $N_{\text{cr}} = 29.5$ об. % В (1–6), 32.5 об. % В (7–11), 48.5 об. % В (12–17), 61.0 об. % В (18–20). Экспериментальное исследование адсорбционного перехода для этих пептидов и обнаружение критических точек, в которой исчезает зависимость от длины цепи, может служить серьезным дополнительным аргументом в пользу предлагаемой модели и ее физической сущности. Кроме того, такое исследование позволяет определить энергию взаимодействия концевых групп Ac– и –amide. Мы положили значение этой энергии равным нулю, что может быть не совсем верно. С другой стороны, отклонение наблюдавших зависимостей от предсказаний теории позволит понять, в каких направлениях она должна быть усовершенствована.

На рис. 5 показано изменение объемов удерживания стандартов S_1 – S_5 (табл. 2) при изменении наклона градиента. Общая формула стандартов S_2 – S_5 следующая: Ac–RGXXGLGLK–amide. Остатки XX в общей формуле есть GG в стандарте S_2 , AG (S_3), VG (S_4) и VV (S_5). Стандарт S_1 имеет ту же последовательность, что и S_3 , однако вместо концевой группы Ac– у него концевая группа $\text{H}_2\text{N}-$.

Очевидно, что уменьшение наклона градиента увеличивает объемы удерживания. Как видно, и в этом случае имеем качественное согласие теории с экспериментом [17].

На рис. 6 и в табл. 3 представлены результаты расчета и экспериментальные данные для ряда

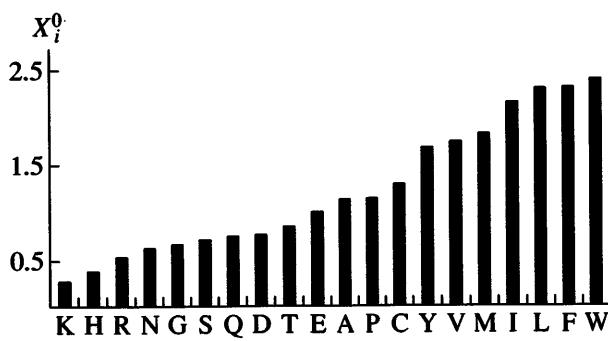


Рис. 3. Диаграмма энергий взаимодействия X_i^0 аминокислотных остатков с поверхностью обращенной фазы $-C_{18}$. Обозначения аминокислотных остатков соответствуют стандартным.

Таблица 1. Расчетные и наблюдаемые [18] времена удерживания синтетических пептидов разной длины на обращенной фазе $-C_{18}$, ($\alpha = 1.092$)

Пептид, №	Строение пептида	$t_{R_{\text{набл}}}$, мин	$t_{R_{\text{теор}}}$, мин
1	Ac-(GKGLG) ₁ -amide	10.0	6.8
2	Ac-(GKGLG) ₂ -amide	16.6	15.2
3	Ac-(GKGLG) ₄ -amide	21.9	22.8
4	Ac-(GKGLG) ₆ -amide	24.4	25.5
5	Ac-(GKGLG) ₈ -amide	26.6	27.1
6	Ac-(GKGLG) ₁₀ -amide	27.8	28.1
7	Ac-(GLGAKGAGVG) ₁ -amide	16.1	17.5
8	Ac-(GLGAKGAGVG) ₂ -amide	22.1	25.3
9	Ac-(GLGAKGAGVG) ₃ -amide	25.1	28.2
10	Ac-(GLGAKGAGVG) ₄ -amide	26.8	29.7
11	Ac-(GLGAKGAGVG) ₅ -amide	28.0	30.7
12	Ac-(LGLKA) ₁ -amide	18.3	15.7
13	Ac-(LGLKA) ₂ -amide	29.7	31.0
14	Ac-(LGLKA) ₄ -amide	39.9	40.3
15	Ac-(LGLKA) ₆ -amide	44.6	43.7
16	Ac-(LGLKA) ₈ -amide	48.4	45.4
17	Ac-(LGLKA) ₁₀ -amide	56.0	46.4
18	Ac-(LGLKL) ₁ -amide	25.2	22.8
19	Ac-(LGLKL) ₂ -amide	38.5	40.6
20	Ac-(LGLKL) ₄ -amide	49.8	50.8

пептидов, взятых из работы [17]. Для этих соединений согласие теории и эксперимента хуже, и наблюдается определенный разброс, который связан в том числе и с тем, что мы не учитывали различия в активности поверхности для разных колонок. По-видимому, авторы работы [17] проводили такую коррекцию по стандарту S_4 , однако первичная информация о времени удерживания

такого стандарта, необходимая для определения активности, в работе не приводится. Тем не менее, даже без стандартизации колонок и условий разделения для всех случаев модель правильно описывает тенденции и масштаб изменения объемов удерживания при изменении последовательности цепи пептида.

Отметим, что пептиды, исследованные в работе [17], дают возможность оценить зависимость времени удерживания от слабой "дефектности" последовательности (перестановки нескольких остатков в цепи). В частности, пептидами с переставленным остатками являются следующие пары образцов: (20, 22), (39, 42), (40, 41), (35, 44) и (45, 46). Любой вариант аддитивной модели дает одинаковый объем удерживания для указанных пар. Напротив, предлагаемая нами модель даже для коротких цепей предсказывает небольшие различия во времени удерживания, связанные с перестановкой остатков (слабыми различиями в "текстах"). Масштаб этой зависимости определяется как конкретной последовательностью, так и типом "переставляемых" остатков, но важно отметить сам факт того, что модель "видит" перестановку.

Экспериментальное разделение последовательностей с переставленными аминокислотами должно наблюдаться только для пары (20, 22). Это не удивительно, поскольку перестановки в данной паре "разрушают" блок сильно взаимодействующих остатков L. Для других пар разделение в экспериментальных условиях работы [17] невозможно. Вместе с тем предсказываемые небольшие различия во временах удерживания пептидов коррелируют с наблюдаемыми временами удерживания. Наконец, следует отметить, что невозможность разделения пар не является принципиальной: изменяя градиент или проводя хроматографирование в изократическом режиме, можно добиться их разделения. Например, при

Таблица 2. Расчетные времена удерживания синтетических пептидов $S_1 \dots S_5$ на обращенной фазе $-C_{18}$ ($\alpha = 1.06$) для разного градиента состава растворителя

v, об. % В/мин	S_1	S_2	S_3	S_4	S_5
4	6.9	7.2	7.6	8.2	9.3
2	9.6	10.3	11.1	12.2	14.4
1	13.7	15.1	16.7	18.8	23.2
0.5	19.6	22.5	25.5	29.6	38.0

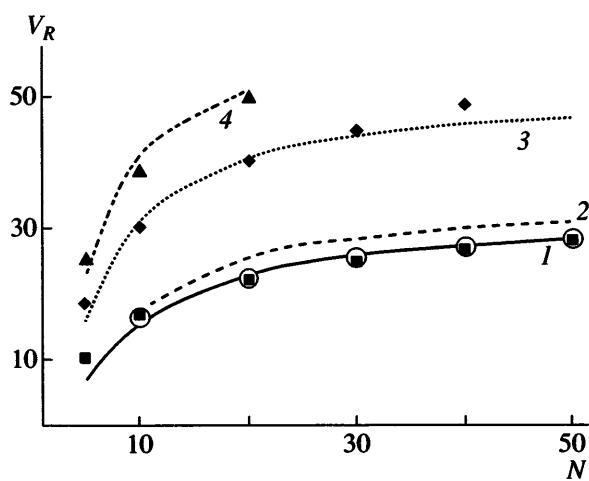


Рис. 4. Зависимость объема удерживания от длины цепи пептидов. Точками обозначены экспериментальные значения [18], кривыми – расчетные для структур 1–6 (1), 7–11 (2), 12–17 (3), 18–20 (4) (табл. 1).

уменьшении наклона градиента В до 0.25 об. % В/мин расчетный объем удерживания пары (39, 40) становится соответственно 38.2 и 37.1 мл, что может оказаться вполне достаточно для их разделения.

Таким образом, предложенная модель разделения пептидов, учитывающая связанность звеньев в цепь, в общем правильно и непротиворечиво описывает совокупность экспериментальных данных работ [16–19]. В отличие от альтернативных подходов, эта модель основана на простой физической картине адсорбции макромолекул, имеющей весьма надежное экспериментальное обоснование в хроматографии синтетических полимеров. Как и в случае с синтетическими полимерами, модель правильно схватывает суть явления и при корректном определении феноменологических параметров впервые дает возможность оценить времена удерживания пептидов, различающихся расположением звеньев в цепи, и осуществить выбор оптимальных условий разделения, необходимых для “чтения” аминокислотных последовательностей.

Примем, что предложенная модель соответствует действительности, и рассмотрим, как тонкие различия в первичной структуре пептидов могут проявляться в хроматографии.

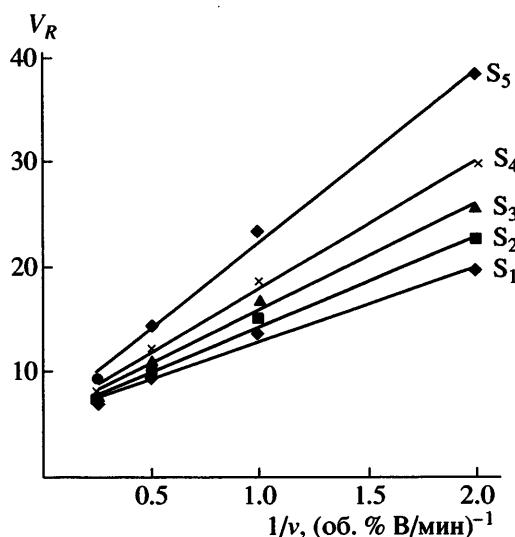


Рис. 5. Зависимость объема удерживания от наклона линейного градиента бинарного растворителя вода–ацетонитрил для стандартов S₁–S₅ (табл. 2). v – скорость изменения компонента В.

В качестве первого примера возьмем так называемые перевернутые, зеркальные последовательности GEADALNLLDGGIYTAG и GATYIYGGDLNLADAEG. Очевидно, что разделение таких последовательностей не следует из аддитивной модели (15). Не вытекает оно и из предлагаемой нами модели. Однако, если эти последовательности содержат разные концевые группы, отличающиеся энергией взаимодействия с поверхностью (например, NH₂ – ~~~ –amide), то картина полностью меняется. Наша модель предсказывает разделение указанных выше последовательностей, отличающихся перестановкой концевых групп.

Различие во взаимодействии зеркальных последовательностей с переставленными концевыми группами есть следствие связанности аминокислотных остатков в цепь. Концевая группа меняет условия взаимодействия с поверхностью не только крайних звеньев, но в том числе и удаленных по цепи аминокислот. Поскольку в зеркальных последовательностях концевая группа H₂N– находится в разном окружении аминокислотных остатков, ее вклад в общую энергию взаимодействия пептидов оказывается различным.

Проанализируем этот вопрос количественно. На рис. 7а показан энергетический профиль “текста” рассматриваемых зеркальных пептидов. Перестановка концевых групп приводит к неболь-

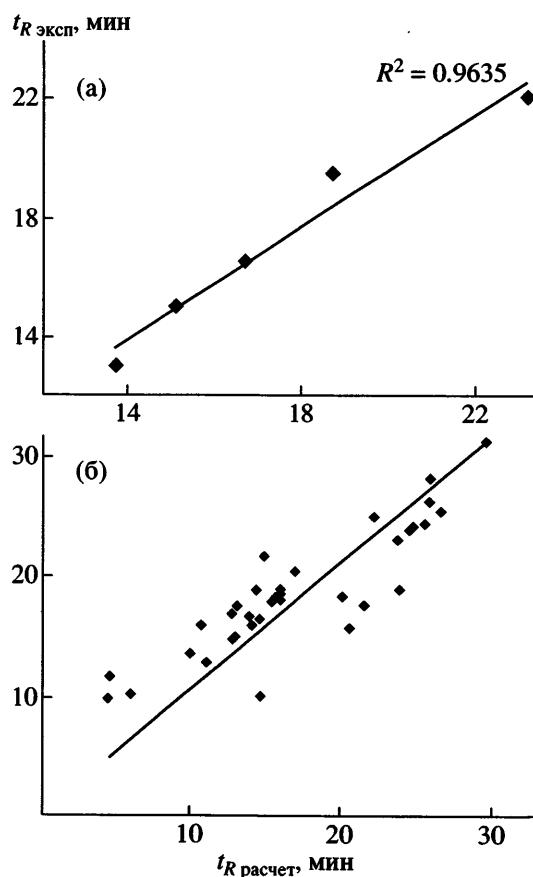


Рис. 6. Корреляция между расчетными и экспериментальными [17] временами удерживания стандартов пептидов S₁–S₅ (а) и пептидов (б) (табл. 3).

шой разности объемов удерживания $\Delta V_R = 0.3$ мл, достаточной для ее экспериментального обнаружения на современных высокоэффективных колонках. Очевидно, что различие в удерживании связано с асимметрией энергетического профиля правой и левой части “текста”. Если переставить аминокислотные остатки и увеличить асимметрию энергетического профиля (рис. 7б), то перестановка концевых групп приведет к заметно большему различию во времени удерживания зеркальных “текстов”, $\Delta V_R = 1.1$ мл. Напротив, симметризация профиля (рис. 7в) делает перестановку концевых групп в зеркальных последовательностях незаметной, $\Delta V_R = 0.08$ мл. В вырожденном случае однородного гомополимера перестановка концевых групп, очевидно, не влияет на объем удерживания.

Рассмотрим теперь пептиды, содержащие остатки лейцина L и изолейцина I. Такие пептиды

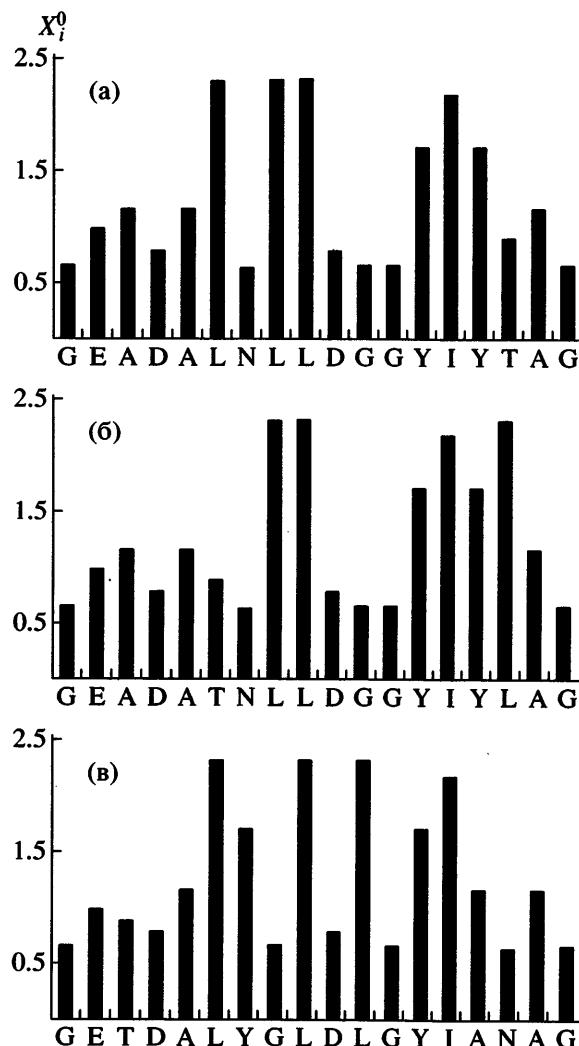


Рис. 7. Энергетический профиль пептида GEADALNLLDGGYIYTAG (а), и результат перестановки аминокислотных остатков, приводящий к асимметричному (б) и симметричному (в) профилю.

характеризуются тем, что современные методы масс-спектрометрии из-за одинаковой массы изомеров не позволяют их идентифицировать ни по макромолекулярному иону, ни по спектрам фрагментации. Вместе с тем хроматография дает возможность разделять такие макромолекулы. В качестве примера рассмотрим стандарт S₅ с текстом Ac–RGVVLGLGK–amide и заменим последовательно все остатки лейцина L на изолейцин I. Теория предсказывает разделение таких последовательностей в градиенте с наклоном 0.5 об. % В/мин (рис. 8).

Исследуем, до какой степени предложенная модель может использоваться для “чтения тек-

Таблица 3. Расчетные и наблюдаемые [17] времена удерживания синтетических пептидов разной длины и последовательности, а также с различными концевыми группами на обращенной фазе $-C_{18}$ ($\alpha = 1.0$)

№ Пептид, №	Строение пептида	$t_{R_{\text{набл}}}$, мин	$t_{R_{\text{теор}}}$, мин
10	H ₂ N-(FFGLM)-amide	20.28	18.20
11	Ac-(KGLGLK)-amide	14.64	9.94
20	Ac-(KGLLLKG)-amide	23.93	18.94
21	Ac-(KKLLLKK)-amide	20.73	15.62
22	Ac-(KLGLGLK)-amide	21.67	17.66
23	H ₂ N-(DYMGMDF)-OH	26.78	25.40
24	Ac-(TDLLAGGK)-amide	13.99	16.68
25	H ₂ N-(TDLLAGGK)-amide	4.68	11.71
26	Ac-(TDGLAGGK)-amide	4.57	9.90
27	Ac-(GAKLEAKG)-amide	6.08	10.03
28	Ac-(TDLLGGGK)-amide	12.87	14.63
30	H ₂ N-(HKTDSFVGLM)-amide	14.91	21.63
31	H ₂ N-(DMHDFFVGLM)-amide	29.75	31.13
32	H ₂ N-(VQAAIDYING)-OH	17.06	20.48
33	H ₂ N-(RPKPQQFFGLM)-amide	26.03	26.20
34	H ₂ N-(DVPKSDQFVGGLM)-amide	22.36	24.90
35	Ac-(GKFKGPPLRRVR)-amide	16.16	18.31
36	H ₂ N-(GGFKRPPLRRVR)-amide	14.18	15.92
37	Ac-(GKFKRPPLRRVR)-amide	15.94	17.90
38	Ac-(GKGKRPPPLRRVR)-amide	10.03	13.53
39	Ac-(GKFGRPPLRRVR)-amide	16.02	19.03
40	Ac-(GKFKRGPLRRVR)-amide	14.63	16.42
41	Ac-(GKFKRPGLRRVR)-amide	14.55	16.33
42	Ac-(GGFKRPPLRRVR)-amide	16.01	18.93
43	Ac-(GKFKRPPLRRVR)-amide	11.12	12.81
44	Ac-(GKFKRPPLRGVR)-amide	15.73	18.32
45	Ac-(GKFKRPPLGRVR)-amide	16.01	18.33
46	Ac-(GKFKRPPLRRVG)-amide	15.59	18.27
47	Ac-(GKFKRPPLRRGR)-amide	12.89	14.94
48	Ac-(VSKTQTSQLVAPA)-amide	12.76	16.92
49	Ac-(VSKTETSQLVAPA)-amide	13.07	17.58
50	Ac-(VSKTATSQLVAPA)-amide	15.51	18.04
51	Ac-(ASKTETSQLVAPA)-amide	10.72	15.98
52	Ac-(DRNAEGYIDAEEL)-amide	25.67	24.35
53	Ac-(NRNANGYIDAEEL)-amide	23.85	23.05
54	Ac-(NRDADGYIDAEEL)-amide	24.72	23.85
55	Ac-(DRDADGYIDAEEL)-amide	24.98	24.16
56	Ac-(SDQEKRKQIASVRGL)-amide	15.47	17.96
57	H ₂ N-(AGCKNFFWKTFTSC)-OH	26.03	28.15
58	Ac-(GAKLEAKG) ₂ -amide	14.40	18.87

Примечание. Номера пептидов взяты из работы [17].

“стата” пептида. Определим, как объем удерживания зависит от всевозможных перестановок аминокислотных остатков в цепи полипептида. Снова возьмем стандарт S_5 и найдем, насколько плотно разные последовательности цепей с одним и тем же аминокислотным составом распределены по объему удерживания.

Стандарт S_5 состоит из $N = 10$ аминокислотных остатков общей формулы $G_4L_2V_2KR$. Такая структура может образовать следующее количество разных последовательностей:

$$\frac{n!}{\prod_i n_i!} = 37800$$

Применяя уравнения (13) и (14) ко всем “текстам”, можно определить для каждого из них объем удерживания. При стандартном градиенте 1 об. % В/мин объем удерживания распределен в интервале от $V_{R\min} = 8.56$ мл до $V_{R\max} = 9.88$ мл. Максимальной величине соответствует структура RKGGGGVLLV, а минимальной – LRGVGGVGKL.

Плотность распределения последовательностей стандарта S_5 представлена на рис. 9. Как видно, это распределение в интервале от $V_{R\min}$ до $V_{R\max}$ очень неравномерно, что очевидно. Редко встречающиеся последовательности, в которых “сильные” остатки $X_i^0 > 1$ (правая половина диаграммы на рис. 3) образуют блоки, имеют заметно большие времена удерживания, чем последовательности, в которых “сильные” и “слабые” $X_i^0 < 1$ остатки (левая половина диаграммы на рис. 3) перемешаны по цепи. В целом область распределения последовательностей по объему удерживания заметно больше, чем типичная ширина хроматографического пика для колонки с эффективностью $N_{\text{eff}} = 5000$. Таким образом, точное определение времени удерживания значительно сокращает число возможных последовательностей-“кандидатов”. Однако количество таких последовательностей все-таки остается достаточно большим, что не дает возможности “прочитать текст” и однозначно определить последовательность искомой цепи S_5 . Для этого необходимо дополнять хроматографические изменения данными масс-спектрометрии.

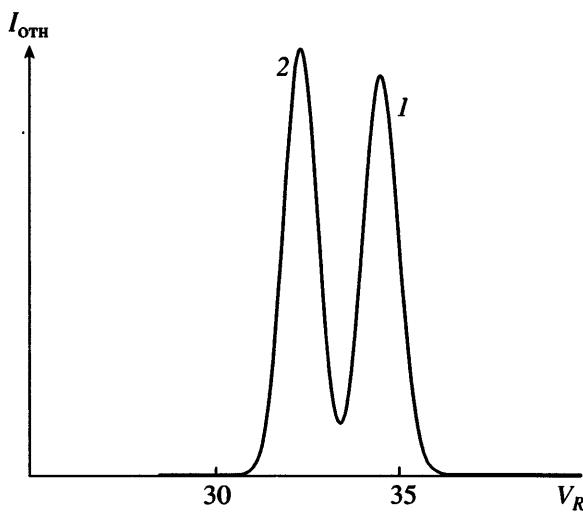


Рис. 8. Теоретические хроматограммы пептида S_5 , Ac-RGVVGLGLK-amide (1) и пептида, в котором лейцин L заменен на изолейцин I (2), демонстрирующие возможность их разделения в градиенте с наклоном 0.5 об. % В/мин.

Экспериментальная проверка следствий предложенной модели, а также ее интеграция с массспектрометрическими базами данных макромолекулярных ионов и их фрагментов для “чтения” последовательностей будут представлены в следующей части работы.

ОБСУЖДЕНИЕ МОДЕЛИ

Кратко обсудим некоторые особенности предложенной модели, границы применимости и возможные пути ее усовершенствования.

Простота модели и ее практической реализации для предсказания объемов удерживания связаны с пренебрежением взаимодействия звеньев друг с другом и с молекулами растворителя, т.е. с идеальностью модели адсорбирующейся цепи. Ясно, что такое упрощенное представление предназначено в первую очередь для примерной оценки эффектов связывания аминокислотных остатков в цепь. Вместе с тем это приближение достаточно хорошо согласуется и с корреляционной теорией определения эффективной энергии, в которой также пренебрегается взаимодействиями в растворе. Можно сказать, что обе части модели оказываются приближениями одного порядка. Усложнение какой-то одной из этих частей без

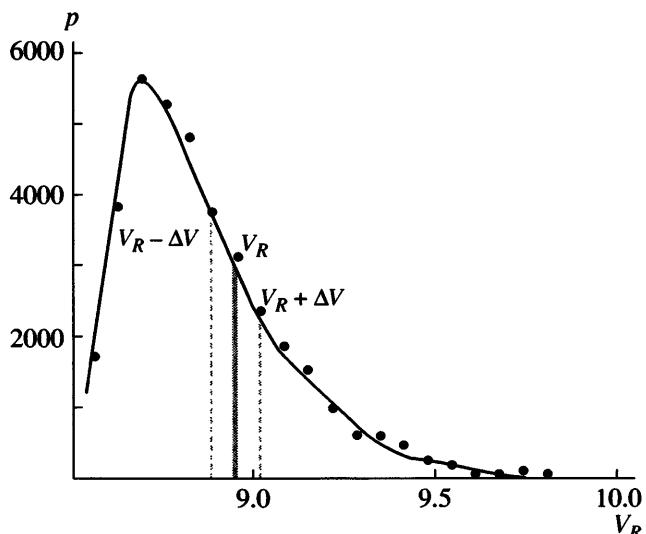


Рис. 9. Плотность распределения последовательностей p по объему удерживания. V_R — удерживаемый объем “искомой” первичной последовательности, ΔV_R — ширина пика на полуысоте; в области $2\Delta V_R$ содержится порядка 800 вариантов последовательностей.

соответствующего усложнения другой вряд ли оправдано, а последовательное введение взаимодействий в растворе приведет (даже в рамках численного моделирования) к заметному усложнению расчетов и (что более важно) к усложнению экспериментального определения возникающих дополнительно феноменологических параметров.

Сказанное выше относится и к макромолекулам синтетических полимеров, однако, как показывает накопленный экспериментальный опыт, излишняя детализация и усложнение модели разделения за счет взаимодействий в растворителе пока не оправданы. Введение критической точки адсорбции как стандартного состояния корреляционной теории и перенормировка эффективной энергии взаимодействия расширяют возможности модели адсорбции идеальной цепи и позволяют осуществить достаточно точное описание закономерностей разделения сложных макромолекул гомополимеров и сополимеров.

Другим упрощающим обстоятельством является представление о точечном характере взаимодействия аминокислотных остатков с поверхностью. Для преимущественно гидрофобных взаимодействий остатков с поверхностью $-C_{18}$ такое приближение, по-видимому, оправдано. Вместе с

тем многие остатки содержат ионогенные группы, способные ионизоваться в смеси воды и ацетонитрила. В частности, в условиях pH 2.0 группы $-NH_2$ существуют в форме $-NH_3^+$, так что взаимодействие этих групп с поверхностью не носит характера гидрофобного и вряд ли локально. Кроме того, в цепи может быть несколько таких групп, и из-за их электростатического отталкивания часть цепи между ними никак не может быть гауссовой и соответствовать цепи случайных блужданий. Цепь между двумя одноименно заряженными остатками скорее должна рассматриваться как полностью вытянутая палочка, адсорбирующаяся на поверхности целиком. Это "улучшает" условия применимости принципа аддитивности, который, как следует из работ [16–19], во многих случаях дает хорошее предсказание для времени удерживания. Такие усложнения необходимо учитывать при рассмотрении реальных пептидов и интерпретации наблюдаемых экспериментально отклонений от предсказаний в рамках предложенной модели. Однако даже в таком случае мы считаем продуктивным использовать модель случайных блужданий в качестве нулевого приближения и рассматривать отклонения на ее фоне.

Что касается других объемных взаимодействий удаленных по цепи звеньев, сблизившихся в результате изгиба цепи и имеющих преимущественно характер отталкивания, то их влиянием, по-видимому, можно пренебречь. Как показывает опыт хроматографического исследования олигомеров [10, 11], для цепей со степенью полимеризации менее 100 влияние объемных взаимодействий в хороших растворителях не слишком заметно.

Вместе с тем в пептидах могут существовать и другие объемные взаимодействия, энергия которых по порядку величины близка к химической связи, например, сульфидные мостики $\sim S-S \sim$. Тогда такая связь в процессе взаимодействия с поверхностью не разрушается и накладывает дополнительные ограничения на реализацию равновесных состояний цепи. Если в цепи имеется один мостик, то ее структура имеет вид цикла размером N , к которому "ковалентно" привязаны два хвоста длиной N_1 и N_2 . Предложенную модель можно обобщить и на подобные структуры. В

этом случае статистическая сумма макромолекулы в поре может быть записана в виде

$$Z \sim \sum_{k=1}^D U^T \prod_{i=1}^{N_1} W(\varepsilon_i) |k\rangle\langle k| \prod_{i=1}^N W(\varepsilon_i) |k\rangle\langle k| \prod_{i=1}^{N_2} W(\varepsilon_i) P_0$$

($|k\rangle$, $\langle k|$ соответственно вектор-столбец и вектор-строка, все элементы которых равны нулю за исключением k -го, равного единице).

В рамках модели никак не учитывается возможное влияние на взаимодействие аминокислотного остатка его ближайших соседей. Например, взаимодействие остатка А "в окружении" остатков Н, $\sim H-A-H \sim$, может отличаться от того, которым он обладает в окружении L, $\sim L-A-L \sim$. Такой учет потребует введения дополнительных поправок к эффективной энергии. Хотя технически это несложно сделать, однако нет оснований вводить поправки без более подробного экспериментального исследования предсказаний модели и ее возможностей для расчета времени удерживания. Возможно, что при более точном сопоставлении теории и эксперимента учет данных поправок будет необходим.

Возможность "чтения текста" аминокислотной последовательности в хроматографии пептидов связана с тонкими эффектами компенсации вблизи поверхности потерь энтропии цепи энергией притяжения. В этом смысле предложенная модель разделения идеально гораздо ближе к критической хроматографии LCCC, нежели к традиционной адсорбционной хроматографии низкомолекулярных соединений. Поэтому естественно называть предложенную модель критической хроматографией биомакромолекул, или BioLCCC. Такое сокращение предполагается использовать в дальнейшем.

ЗАКЛЮЧЕНИЕ

Принципиальным отличием предложенной модели хроматографии пептидов является то, что в ней явным образом учитывается связанность аминокислотных остатков в цепь, т.е. последовательность. Такой подход дает больше возможностей для понимания характера разделения биомакромолекул. Он значительно глубже аддитивного, применяемого для интерпретации хроматограмм пептидов и предсказания их объема удерживания. Эта модель может быть положена в основу хромато-

графии пептидов, а также применена для совместного использования с масс-спектрометрическими базами данных. Модель дает правильное описание разделения пептидов с разными последовательностями в целом, однако в частностях ее предсказания могут отличаться от экспериментально наблюдаемых объемов удерживания. Это связано с идеализацией как модели цепи пептида и ее адсорбции, так и самой хроматографической системы.

СПИСОК ЛИТЕРАТУРЫ

- Лифшиц И.М. // Журн. эксперим. и теорет. физики. 1968. Т. 55. № 6. С. 2408.
- Lifshits I.M., Grosberg A.Yu., Khokhlov A.R. // Rev. Mod. Phys. 1978. V. 50. № 3. P. 684.
- Де Жен П. Идеи скейлинга в физике полимеров. М.: Мир, 1982.
- Гросберг А.Ю., Хохлов А.Р. Статистическая физика макромолекул. М.: Наука, 1989.
- DiMarzio E.A., Rubin R.J. // J. Chem. Phys. 1971. V. 55. № 9. P. 4318.
- Скворцов А.М., Беленький Б.Г., Ганкина Э.С., Теников М.Б. // Высокомолек. соед. А. 1978. Т. 20. № 3. С. 678.
- Горбунов А.А., Скворцов А.М. // Высокомолек. соед. А. 1980. Т. 22. № 5. С. 1137.
- Горшков А.В., Евреинов В.В., Энтелис С.Г. // Высокомолек. соед. А. 1982. Т. 24. № 3. С. 524.
- Горшков А.В., Евреинов В.В., Энтелис С.Г. // Докл. АН СССР. 1983. Т. 272. № 3. С. 632.
- Entelis S.G., Evreinov V.V., Gorshkov A.V. // Adv. Polym. Sci. 1986. V. 76. P. 129.
- Энтелис С.Г., Евреинов В.В., Кузав А.И. Реакционноспособные олигомеры. М.: Химия, 1985.
- Pasch H., Trathnigg B. HPLC of Polymers. Berlin; Heidelberg; New York: Springer-Verlag, 1999.
- Горшков А.В., Евреинов В.В. // 100 лет хроматографии. М.: Наука, 2003. С. 136.
- Eisenriegler E., Kremer K., Binder K. // J. Chem. Phys. 1982. V. 77. № 12. P. 6296.
- Snyder L.R. Principles of Adsorption Chromatography. New York: Marcel Dekker, 1968.
- Guo D., Mant C.T., Taneja A.K., Parker J.M.R., Hodges R. // J. Chromatogr. 1986. V. 359. P. 499.
- Guo D., Mant C.T., Taneja A.K., Hodges R. // J. Chromatogr. 1986. V. 359. P. 519.
- Mant C.T., Burke L.T.W., Black J.A., Hodges R. // J. Chromatogr. 1988. V. 458. P. 193.
- Mant C.T., Burke L.T.W., Zhou N.E., Parker R.J.M., Hodges R. // J. Chromatogr. 1990. V. 485. P. 365.

Applicability of the Critical Chromatography Concept to Proteomics Problems: Dependence of Retention Time on the Sequence of Amino Acids

A. V. Gorshkov^a, V. V. Evreinov^a, I. A. Tarasova^b, and M. V. Gorshkov^b

^a Semenov Institute of Chemical Physics, Russian Academy of Sciences,
ul. Kosyginia 4, Moscow, 119991 Russia

^b Institute of Energy Problems of Chemical Physics, Russian Academy of Sciences,
Leninskii pr. 38/2, Moscow, 119334 Russia

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

Abstract—A peptide separation model based on the technique of liquid chromatography of macromolecules at the critical condition was proposed. In terms of this model, the array of experimental data on the separation of peptides is considered. The main phenomenological parameters of the model—effective adsorption energies of amino acid residues—were determined, thus allowing the influence of character of their alternation in the chain on retention times to be predicted. The model is applicable to investigation into the feasibility of separation in different chromatographic modes of peptides with the same amino acid composition and different sequences of units in the chain, as well as peptides containing amino acid isomers and mirror sequences with different terminal groups.