

## ПРИМЕНЕНИЕ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ В ЭКСПЕРИМЕНТЕ

### ХРОМАТОГРАФИЯ СВЕРХВЫСОКОГО РАЗРЕШЕНИЯ

© 2001 г. Е. Л. Косарев, К. О. Муранов\*

*Институт физических проблем им. П.Л. Капицы РАН  
Россия, 117334, Москва, ул. Косыгина, 2*

Поступила в редакцию 12.02.2001 г.

После доработки 19.02.2001 г.

Описан способ улучшения разрешения хроматографического анализа, основанный на получении аппаратной функции хроматографической колонки, т.е. хроматограммы индивидуального соединения. Совокупность двух функций: хроматограммы анализируемого вещества и аппаратной функции хроматографической колонки вместе со знанием статистики шумов дает возможность применить пакет программ восстановления сигналов RECOVERY для получения хроматограммы со сверхвысоким разрешением. Термин сверхвысокое разрешение означает лучшее, чем определяемое шириной аппаратной функции. Предлагаемый способ опробован при хроматографии бычьего сывороточного альбумина с помощью гель-фильтрации. Получаемое разрешение превышает достигаемое с помощью высокоэффективной жидкостной хроматографии при меньшей в ~15–20 раз стоимости приборного комплекса.

#### ВВЕДЕНИЕ

Хроматография – это физико-химический метод исследования веществ, основанный на разделении компонент при их перемещении через колонку хроматографа [1], представляющую собой калиброванную трубку длиной от 10 см (высокоэффективная жидкостная хроматография – в.э.ж.х.) до 30 м (газовая хроматография).

Ситовая хроматография, или гель-фильтрация, наиболее наглядно демонстрирует процесс разделения смеси веществ. Неподвижная насадка, наполняющая колонку, представляет собой гель, образованный переплетенными молекулами полимера. Полимер (агароза, декстран, полиакриламид и т.д.) сформирован в виде гранул размером ~100 мкм. При движении образца через колонку крупные молекулы, проходя в пространстве между гранулами, движутся вместе с потоком растворителя. Более мелкие молекулы, проникая внутрь гранул, задерживаются и выходят позже. Таким образом, образец разделяется на фракции в зависимости от молекулярного веса. При этом на определенном участке время  $t$  удерживания молекулы массой  $M$  (равное времени прохождения молекулой хроматографической колонки) определяется формулой (см. [2])

$$\log M = a - bt. \quad (1)$$

Постоянные  $a$  и  $b > 0$  определяются калибровочными измерениями образцов с известными молекулярными весами.

\* Институт биохимической физики им. Н.М. Эммануэля РАН, Москва.

Часто хроматограмма представляет собой сумму пиков, многие из которых накладываются друг на друга, что затрудняет точный анализ. Для увеличения разрешающей способности метода, с одной стороны, разрабатываются новые наполнители (сорбенты) для колонок, которые позволяют проводить более тонкое разделение, а с другой – применяют расчетные методы. В основе расчетных методов лежит предположение, что форма пика индивидуального вещества может быть описана аналитически, например, с помощью функции Гаусса [2]. Однако в реальных условиях часто это неверно. Например, насадки обладают некоторой адсорбционной активностью. Поэтому на один вид разделения (например, ситовую хроматографию) накладывается другой вид хроматографии (адсорбционная хроматография), что приводит к искажению формы пика. При этом вследствие необратимой адсорбции свойства неподвижной фазы могут меняться. Все это приводит к тому, что форма пика индивидуального соединения оказывается не только несимметричной, но и может изменяться со временем.

Именно поэтому параметрический подход, применяемый в большинстве стандартных пакетов обработки данных при хроматографии, несмотря на многообразие используемых базисных функций, в принципе не может дать надежных результатов, поскольку наблюдаемая на опыте форма пика часто даже не входит в обширный набор функций стандартного пакета.

Мы полагаем, что корректное решение проблемы разделения перекрывающихся пиков возможно с помощью непараметрического метода. При этом форма пика определяется непосредст-

венно при хроматографии индивидуального соединения, и эта форма пика может быть названа, по аналогии с оптикой, аппаратной функцией хроматографической колонки.

Хроматограмма смеси веществ представляет собой суперпозицию индивидуальных пиков. Если форма этих пиков одинакова во всем рабочем диапазоне прибора, то наблюдаемая хроматограмма представляет собой свертку распределения по массам исследуемого вещества (в случае гель-фильтрации) с аппаратной функцией хроматографической колонки. Обратная задача разложения сложного спектра на одинаковые составляющие достигается решением интегрального уравнения свертки, в котором входными данными является собственно хроматограмма, а ядром оператора свертки является полученная из измерений аппаратная функция хроматографической колонки. О решении подобных задач см. работы [3, 4].

### ПАКЕТ ПРОГРАММ RECOVERY И СВЕРХРАЗРЕШЕНИЕ

Для решения этой задачи мы используем основанный на методе максимума правдоподобия пакет программ восстановления сигналов из зашумленных данных RECOVERY [5, 6], разработанный специально для восстановления неотрицательных функций, каковыми всегда являются искомые спектры, в данном случае представляющие собой распределение исследуемого вещества по молекулярному весу. В общем же случае спектр будет представлять собой распределение исследуемого соединения по типу межмолекулярных взаимодействий, лежащих в основе разделения.

Для восстановления сигналов в пакете программ RECOVERY используется метод максимума правдоподобия (м.м.п.), поскольку этот метод обеспечивает максимально возможную эффективность восстановления [7]. Согласно м.м.п., сначала нужно определить функцию правдоподобия  $L = \mathcal{P}$ , где  $\mathcal{P} = \mathcal{P}(F | f_0)$  – условная вероятность наблюдения ряда экспериментальных данных,

$$F(x_i), \quad i = 1, 2, \dots, n,$$

совпадающего с реально полученным рядом при условии, что искомое решение равно функции  $f_0$ . Мы можем рассматривать множество неизвестных значений функции  $f_0(t_j), j = 1, 2, \dots, m$ , как вектор в  $m$ -мерном пространстве решений. Каждая точка в этом пространстве соответствует одному из возможных решений, и далее находится максимум функции правдоподобия в этом  $m$ -мерном пространстве решений при условии выполнения некоторых необходимых ограничений на решение.

Функция правдоподобия для экспериментальных данных с гауссовской статистикой сводится к отрицательной величине квадратичной невязки

между экспериментальными данными  $\{F\}$  и их аппроксимацией  $\{\hat{F}_0\}$

$$L = -\frac{1}{2} \|F - \hat{F}_0\|^2.$$

В работе [6] приведены итерационные формулы нахождения максимума функции правдоподобия. Эти итерационные формулы реализуют нелинейный метод восстановления сигнала, который в отличие от линейных методов способен расширять полосу восстановленного сигнала по сравнению с сигналом на входе, т.е. способен определять более тонкие детали в выходном сигнале, даже невидимые во входных данных, и тем самым увеличивать разрешение.

Продолжая аналогию с оптикой, укажем, что естественная мера разрешения хроматографа равна ширине его аппаратной функции  $\Delta$ . Эта ширина определяется процессами диффузии, размывающими пик. При использовании пакета программ RECOVERY возможно получение разрешения  $\delta$  лучшего, чем эта ширина:  $\delta < \Delta$ , и тем самым получать сверхразрешение

$$SR = \frac{\Delta}{\delta} > 1.$$

В работах [3, 4] было показано, что предельная величина достижимого сверхразрешения  $SR_{\max}$  ограничена шумом во входных данных и может быть вычислена по формуле

$$SR < SR_{\max} = \frac{1}{3} \log_2(1 + E_s/E_n), \quad (2)$$

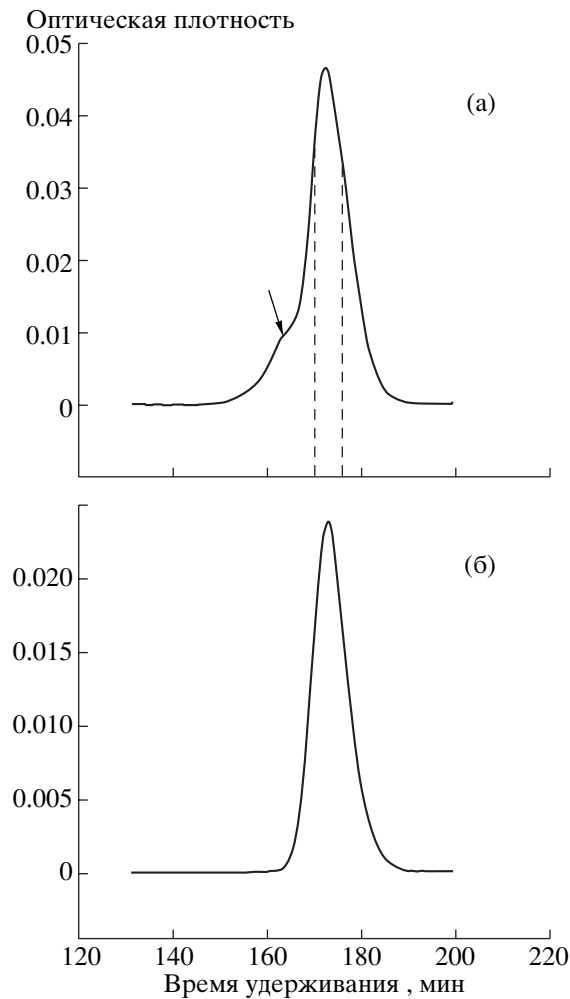
где  $E_s = \int_{t_0}^{t_1} F^2(t) dt$  – энергия сигнала;  $E_n = n\sigma^2$  – энергия шума;  $F(t)$  – значение полученной хроматограммы в момент времени  $t$ ;  $t_1$  и  $t_2$  – начальный и конечный моменты времени на хроматограмме,  $n$  – число данных на хроматограмме,  $\sigma^2$  – дисперсия шума в каждой точке.

Если отношение сигнал/шум в (2) выразить в децибелах

$$1 \text{ дБ} = 10 \lg(E_s/E_n),$$

то можно записать очень удобное для использования приближенное выражение для предельного сверхразрешения:  $SR_{\max} \approx \text{дБ}/10$ . Для большей конкретности укажем, что отношение сигнал/шум для приводимых далее в статье данных хроматографического анализа бычьего сывороточного альбумина (БСА) (кривая А на рис. 2) равно 47 дБ. Это означает, что такое качество данных позволяет получить разрешение в  $\sim 5$  раз лучшее, чем ширина аппаратной функции хроматографа.

Предлагаемый подход ранее показал хорошие результаты в спектроскопии [8–10] и в ядерной



**Рис. 1.** Получение аппаратной функции колонки с TSK-гелем: **а** – хроматограмма исходной смеси БСА (30 мг). Стрелка – неразрешенный пик, пунктирные линии – интервал времени, когда была собрана фракция мономера БСА; **б** – хроматограмма собранной фракции БСА. Следует обратить внимание на отсутствие неразрешенного пика и на небольшую видимую асимметрию.

физике [11]. Здесь мы используем его для задач хроматографии.

#### МАТЕРИАЛЫ И МЕТОДЫ

Методика эксперимента заключалась в следующем: смесь белков известного состава (мономер, димер и тример БСА) разделялась с применением насадки, которая, как было известно заранее, в принципе не способна разделить по массам указанную смесь, т.е. в результате разделения получается хроматограмма с сильно перекрывающимися пиками. Данные обрабатывались с помощью программного комплекса RECOVERY, и результат сравнивался с данными более тонкого разделения, полученными с помощью высокоэффективной жидкостной хроматографии.

БСА (фракция № 5, Sigma, USA) растворялся в фосфатном буфере ( $\text{NaCl} - 100 \text{ mM}$ ,  $\text{NaH}_2\text{PO}_4 - 40 \text{ mM}$  этилендиаминтетрауксусная кислота –  $1 \text{ mM}$ ,  $\text{pH} = 6.8$ ) и после фильтрации через ультрафильтр (Millipore,  $0.2 \text{ мкм}$ ) наносился на колонку ( $25 \times 900 \text{ мм}$ ) с гелем TSK HW 55 (TOYO SODA, Japan). Разделение проводилось при скорости тока растворителя  $1.5 \text{ мл/мин}$ . Детектирование осуществлялось измерением оптической плотности выходящей из хроматографической колонки жидкости при помощи у.ф.-детектора на длине волны  $\lambda = 280 \text{ нм}$  (Uvicord S 2138, LKB, Sweden). Выходной сигнал с детектора передавался в компьютер с помощью интерфейса [12]. Чистота белка контролировалась с помощью электрофореза [13] в полиакриламидном геле в присутствии додецилсульфата натрия и меркаптоэтанола (SDS-PAAG) как в денатурирующих (кипячение  $5 \text{ мин}$ ), так и нативных условиях. В.э.ж.х. проводилась на колонке TSK G2000 SW Spherogel  $10 \text{ мкм}$ ,  $7.5 \times 600 \text{ мм}$  (LKB, Sweden) при скорости тока растворителя  $0.5 \text{ мл/мин}$  на хроматографе Beckman-Altex 334 (USA). Калибровка колонок проводилась с помощью набора стандартных белков. Площадь пиков оценивалась расчетным способом с помощью пакета программ MicroCal Origin, 4.1. Вычисления производились на IBM-совместимых персональных э.в.м., операционная система MS-DOS и Windows NT.

В данной работе из всего комплекса программ RECOVERY использовались программы Dconv\_ и Dconv2\_, предназначенные для решения интегрального уравнения свертки с гауссовским распределением шумов с увеличенным числом входных данных на хроматограмме: до 4096 и 2048 соответственно для указанных выше программ. Полные листинги стандартного варианта пакета программ RECOVERY на Фортран-77 представлены в [5], а также находятся в библиотеке компьютерных программ <http://cpc.cs.qub.ac.uk/> под именем ACLJ [6]. Дополнительная информация об этом пакете программ может быть получена по адресу <http://kapitza.ras.ru/people/kosarev/recovery.html>

#### ХРОМАТОГРАФИЯ БСА И ПОЛУЧЕНИЕ АППАРАТНОЙ ФУНКЦИИ

На рис. 1а представлена хроматограмма БСА. Видно, что вещество состоит, по крайней мере, из двух компонент (вторая компонента на рисунке обозначена стрелкой). Разрешающей способности используемой колонки недостаточно для разделения этих компонент. Это не удивительно, так как эта колонка предназначена для разделения в широком интервале молекулярных масс  $10-700 \text{ kD}$  [14], она не способна разделить мономер и димер БСА, имеющие массы  $67$  и  $134 \text{ kD}$  соответственно.

Пунктирными линиями обозначен интервал времени, когда была собрана фракция, содержащая мономер БСА. Полученная фракция была

повторно хроматографирована. Форма хроматограммы, представленная на рис. 1б, свидетельствует об относительной чистоте (отсутствии димера или тримера БСА) в полученной фракции. Полученная фракция была также исследована с помощью электрофореза в SDS-ПААГ. Результаты электрофореза свидетельствует о том, что в полученной фракции присутствует мономер БСА, ~15% низкомолекулярных примесей, а также следовые количества димера и тримера БСА.

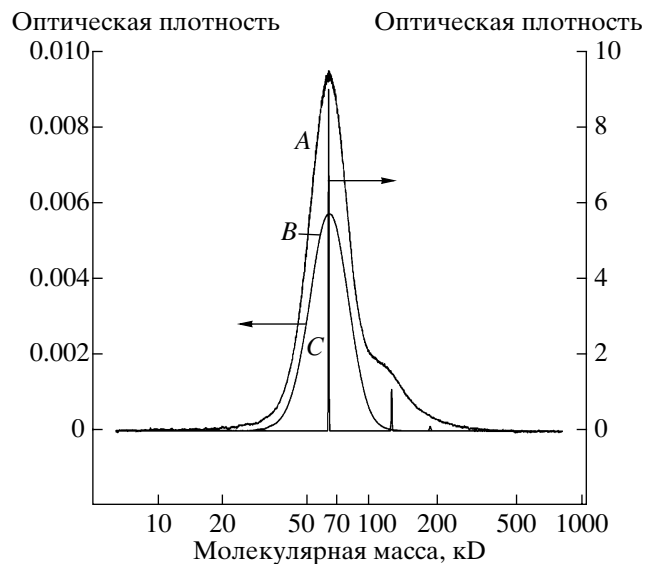
Наблюдаемая небольшая асимметрия пика может быть вызвана примесями низкомолекулярных соединений (продукты гидролитического расщепления БСА), и уже упомянутыми посторонними причинами – необратимой адсорбцией, гидрофобными взаимодействиями и т.д.

Эта асимметрия не мешает использованию предлагаемого метода получения сверхразрешения при хроматографическом разделении. Кривая, изображенная на рис. 1б, использовалась в данной работе в качестве аппаратной функции.

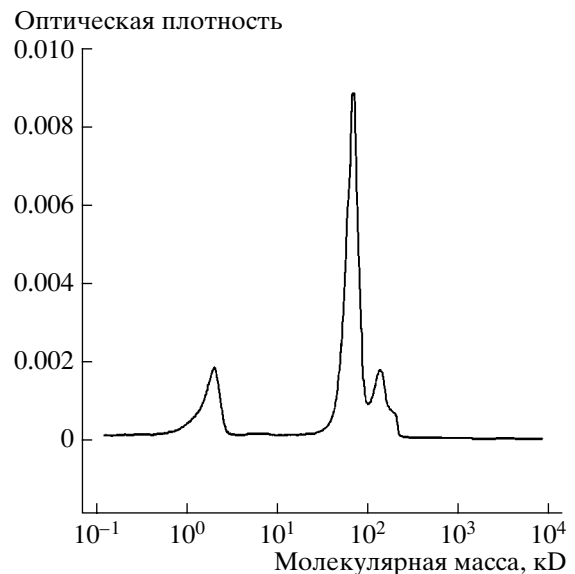
На рис. 2 показаны хроматограмма БСА (кривая А) и фракция мономера БСА, используемая в качестве аппаратной функции (кривая В). Узкие пики представляют результат восстановления данных хроматографического разделения с помощью программы Dconv2\_. По оси абсцисс отложено не время удерживания вещества, как на рис. 1, а вычисленные по формуле (1) значения молекулярных масс. Результат восстановления представляет собой совокупность трех пиков, которые соответствуют мономеру, димеру и тримеру БСА.

На рис. 3 представлено разделение того же самого образца БСА с помощью в.э.ж.х. Как известно, высокоэффективное разделение этим методом достигается за счет использования мелкодисперсных насадок. В использованной нами колонке в.э.ж.х. размер гранул составлял 10 мкм. Такой малый размер гранул позволяет существенно увеличить эффективность разделения, что однако приводит к значительному увеличению рабочего давления, необходимого для прокачивания растворителя через колонку, а значит, и к использованию специализированных насосов, высококачественных соединений и специальных материалов. Все это заметно увеличивает стоимость анализа.

Колонка в.э.ж.х. разделяет смесь гораздо эффективнее, чем TSK-гель – на рисунке видно, что препарат содержит четыре компонента: низкомолекулярную примесь (массой 2 кД), мономер и димер с неразрешенным тримером БСА. Приведенные на этом рисунке результаты совпадают с литературными данными [15], в которых показано, что фракция № 5 БСА, использованная в нашей работе, содержит мономер, димер и тример молекулы, некоторое количество низкомолекулярных примесей, а также следовые количества тетрамера БСА.

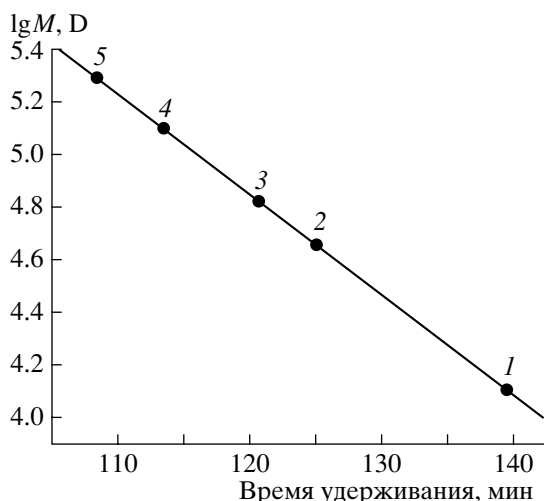


**Рис. 2.** Применение комплекса RECOVERY к исходным данным, представленным на рис. 1. А – хроматограмма БСА, полученная на колонке с TSK-гелем; В – аппаратная функция (хроматограмма мономера БСА); С – результат восстановления данных хроматографического разделения с помощью программы Dconv2\_ из пакета RECOVERY.



**Рис. 3.** Высокоэффективная жидкостная хроматограмма препарата БСА. Колонка TSK G2000 SW, Spherogel 10 мкм, 7.5 × 600 мм (LKB), скорость тока растворителя 0.5 мл/мин.

Пик с малой массой, видимый на рис. 3 и отсутствующий на рис. 2, лежит за границей диапазона масс используемой нами колонки. В то же время диапазон масс 67–201 кД, не полностью разделенный с помощью в.э.ж.х., прекрасно разрешен нашим методом. Если оценить сверхразрешение,



**Рис. 4.** Калибровочный график для определения постоянных  $a$  и  $b$  в формуле (1), полученный при разделении стандартных белков: 1 – цитохрома С (12.5 кД), 2 – овальбумина (45 кД) и 3 – БСА (67 кД). По оси ординат – десятичный логарифм молекулярной массы белка. Аппроксимирующая прямая  $y(x) = -0.0385x + 9.47$  проведена методом наименьших квадратов по точкам 1, 2, 3. Точки 4 и 5 – димер ( $127 \pm 1.62$  кД) и тример ( $200 \pm 5.78$  кД) БСА.

полученное с помощью пакета программ RECOVERY по данным, приведенным на рис. 2, как отношение ширины аппаратной функции к расстоянию между димером и тримером БСА, то получим величину 1.1, что уже оправдывает использование термина “сверхразрешение” в названии статьи. Как отмечено выше, качество данных позволяет улучшить эту величину еще в  $\sim 5$  раз.

Ниже представлены молекулярные массы соединений (кД), полученные с помощью стандартной TSK-гель-хроматографии, в.э.ж.х. и описываемым в статье способом в сравнении с данными из каталога Sigma-Aldrich [16]:

Номер пика	0	1	2	3
Табличное значение		$67 \pm 1$	$134 \pm 1.4$	$201 \pm 17$
TSK-гель	–	67	–	–
В.э.ж.х.	2	68	135	190
RECOVERY	–	$66.4 \pm 0.53$	$127 \pm 1.62$	$200 \pm 5.78$

В строчке, обозначенной RECOVERY, приведены данные, усредненные по пяти независимым измерениям с вероятной ошибкой  $\pm 1\sigma$ . Нумерация пиков идет в порядке увеличения молекулярного веса соединения. Пик с номером 0 соответствует низкомолекулярной примеси, а с номерами 1, 2, 3 – мономеру, димеру и тримеру БСА.

Калибровочный график для определения постоянных  $a$  и  $b$  в формуле (1) приведен на рис. 4.

Предлагаемый метод демонстрирует хорошее согласие с опубликованными данными по определяемым массам – лучше 5%, и демонстрирует гораздо лучшее разрешение, чем в.э.ж.х. Следует отметить, что с помощью стандартных методов оказалось возможным достичь погрешности 6% лишь при использовании одновременно трех детекторов: рассеянного лазерного излучения, визкозиметрического и рефрактивного [17], в то время как в данной работе используется лишь один детектор.

Причины остающегося небольшого расхождения в определяемых массах можно разделить на две группы: первая возникает при определении молекулярного веса с помощью калибровочного графика, вторая связана с использованием непараметрического метода. Рассмотрим эти причины последовательно. Как известно, молекула каждого индивидуального белка имеет свою специфическую форму, которая часто бывает сферической или эллипсоидальной. Молекула мономера альбумина имеет форму, приближающуюся к эллипсоиду вращения с размерами главной оси и оси в поперечном сечении  $140 \times 40 \text{ \AA}$ , в то время как димер является комплексом двух таких молекул с перекрытием в 50% [18]. Гидродинамические свойства димера и, естественно, его способность проникать в разделяющий гель существенно отличаются от свойств мономера. Возможно, что именно поэтому молекула димера, имеющая форму ближе к сферической, чем молекула мономера, дает уменьшенное значение молекулярного веса.

Вторая причина – особенности используемого метода получения распределения по молекулярному весу исследуемого сложного вещества. Как указано выше, мы применяем непараметрический метод, не фиксируя заранее форму аппаратной функции хроматографической колонки, а определяя ее из опыта, а также не предполагая заранее форму спектра масс определяемого вещества в виде набора дискретных линий. За такое смягчение требований приходится платить уменьшением точности при восстановлении спектра масс по сравнению с параметрическим методом. Как известно, наименьшая возможная погрешность определения параметров при использовании м.м.п. дается неравенством Рао–Крамера [7], и эта погрешность степенным образом зависит от погрешности входных данных. В то же время, предельное возможное разрешение при использовании непараметрического метода логарифмически зависит от погрешности входных данных [3, 4], и это разрешение, вообще говоря, хуже, чем при использовании параметрического метода.

Оптимальным представляется последовательное использование сначала непараметрического метода, и лишь после получения достаточно на-

дежной информации о форме аппаратной функции и восстанавливаемого спектра следует переходить к использованию параметрического метода. Пример успешного использования такого подхода можно найти в работе [11].

Процентное содержание компонент разделяемой смеси

Номер пика	0	1	2	3
TSK-гель	–	100	–	–
В.э.ж.х.	17.3	66.7	13.7	2.3
RECOVERY	–	83 ± 6	14.7 ± 6	2.3 ± 1

Эти данные демонстрируют хорошее соответствие процентного содержания компонент, определяемого известными ранее и новым методом. При этом следует обратить внимание на то, что хотя при используемой в данной работе колонке TSK-геля пик низкомолекулярной примеси не виден, его присутствие отражается в величине процентного содержания основного пика номер 1, вклад которого в общий баланс с хорошей точностью совпадает с суммой нулевого и первого пиков, определяемых с помощью в.э.ж.х.

Таким образом, мы показали, что предложенный нами метод существенно улучшает качество хроматографического разделения, что позволяет количественно измерить содержание отдельных компонент смеси. Следует отметить, что эти новые возможности достигаются за счет рациональной обработки измеренных данных, без всякого усложнения измерительной аппаратуры.

Применение программного комплекса RECOVERY к данным гель-фильтрации существенно повысило разрешающую способность этого метода и превысило качество разделения, достигаемое с помощью в.э.ж.х. Следует отметить, что стоимость использованного приборного комплекса для гель-фильтрации (~US\$ 1000,-) примерно в 15-20 раз меньше, чем стоимость установки для в.э.ж.х. (~US\$ 20000,-).

В заключение авторы выражают благодарность С.Н.Бубенчиковой из НИИ физико-химической биологии им.А.Н.Белозерского МГУ им. М.В. Ломоносова за интерес к работе и помощь

в проведении в.э.ж.х. Авторы также благодарят Российский фонд фундаментальных исследований и Научную программу Министерства образования Российской Федерации и МГУ им. М.В. Ломоносова “Фундаментальные исследования высшей школы в области естественных и гуманитарных наук. Университеты России” за финансовую поддержку по грантам 00-04-48364 и 992582 соответственно.

#### СПИСОК ЛИТЕРАТУРЫ

1. Сакодьинский К.И., Бражников В.В., Волков С.А. и др. Аналитическая хроматография. М.: Химия, 1993. С. 16–21.
2. Фрайфелдер Д. Физическая биохимия. М.: Мир, 1980. С. 200–201
3. Косарев Е.Л. // ПТЭ. 1989. № 4. С. 84.
4. Косарев Е.Л. // Радиотехника и электроника. 1990. Т. 35. № 1. С. 68.
5. Гельфгат В.И., Косарев Е.Л., Подольяк Е.Р. // ПТЭ. 1991. № 5. С. 86 (Деп. рук. № 2635–В91. М.: ВИНТИ, 1991).
6. Gelfgat V.I., Kosarev E.L., Podolyak E.R. // Computer Physics Communications. 1993. V. 74. № 3. P. 335.
7. Линник Ю.В. Метод наименьших квадратов и основы теории обработки наблюдений. М.: Физматгиз, 1962. Гл.3. §3.
8. Косарев Е.Л., Песков В.Д., Подольяк Е.Р. // ЖТФ. 1983. Т. 53. Вып. 6. С. 1101.
9. Коротких В.Л., Косарев Е.Л., Ормонт А.Б., Коротких А.В. // ПТЭ. 1994. № 6. С. 88.
10. Каминский А.С., Косарев Е.Л., Лавров Э.В. // ПТЭ. 1997. № 3. С. 57.
11. Kolganova E.A., Kosarev E.L., Ososkov G.A. // Nucl. Instrum. and Methods in Phys. Res. 2000. V. A443. P. 464.
12. Кириллов В.И. // ПТЭ. 1981. № 4. С. 89.
13. Laemmli U.K. // Nature. 1970. V. 227. № 259. P. 680.
14. TSK-Gel Toyopearl. TOYO SODA manufacturing Co. LTD.
15. Arakawa T., Kita Y. // J. Pharmaceutical Sci. 2000. V. 89. P. 646.
16. Sigma-Aldrich 2000-2001, корпорация Sigma-Aldrich.
17. Ross S. // LabPlus International. 2000. V. 14. № 4. P. 26.
18. Squire P.G., Moser P.O., Kanski C.T. // Biochemistry. 1968. V. 7. № 12. P. 4261.