

УДК 541.64:543.544

М.Л. Придатченко¹, И.А. Тарасова^{1,2}, К. Масселон⁴, А.В. Горшков³, М.В. Горшков²

¹ Московский физико-технический институт (государственный университет)

² Институт энергетических проблем химической физики РАН

³ Институт химической физики им. Н.Н.Семенова РАН

⁴ CEA, Université Joseph Fourier, Grenoble, France

Единый подход к созданию универсальных баз данных точных масс и времён удерживания пептидных маркеров белков на основе модели критической хроматографии биомолекул.

Одной из основных задач протеомики является идентификация первичной структуры белка. Как правило, эта задача выполняется хромато-масс-спектрометрическим методом, при этом пептиды идентифицируются на основе масс-спектрометрических, периодически пополняемых баз данных. Хроматограф же выполняет лишь функцию разделителя смеси перед подачей её в масс-спектрометр, что, по меньшей мере, не рационально. Как показала практика не всегда по масс-спектрам можно точно указать первичную структуру пептидов. При этом, в последние годы интенсивно обсуждаются возможности использования наряду с масс-спектральными данными хроматографических времен удерживания, как дополнительного и независимого источника информации о первичной структуре белков [1-3]. Однако, времена удерживания пептидов не универсальны, т.е. для одного и того же вещества при разных параметрах эксперимента (при смене хроматографической колонки, профиля градиента, бинарного растворителя, и т.д.) времена удерживания сильно различаются.

В рамках данного исследования были проведены эксперименты по определению корреляции между временами удерживания, полученными при разных параметрах эксперимента (т.е. наклона градиента, скорости элюирования, состава мобильной фазы, параметры хроматографической колонки и т.д.). В выбранном нами диапазоне экспериментальных параметров (в нём, как правило, проводится большинство протеомных исследований) была выявлена линейная зависимость с высоким коэффициентом корреляции 0,98-0,99. Следует, однако, отметить, что в ряде случаев, для пептидов с близкими временами удерживания, но сильно отличающимися по длине цепи аминокислотной последовательности, наблюдается изменение порядка выхода таких пептидов из хроматографической колонки (рис.1, рис.2). Это существенно осложняет возможность приведения экспериментальных хроматографических данных к

единой шкале нормализованных времён удерживания, а, следовательно, и создание хроматографических баз данных пептидов.

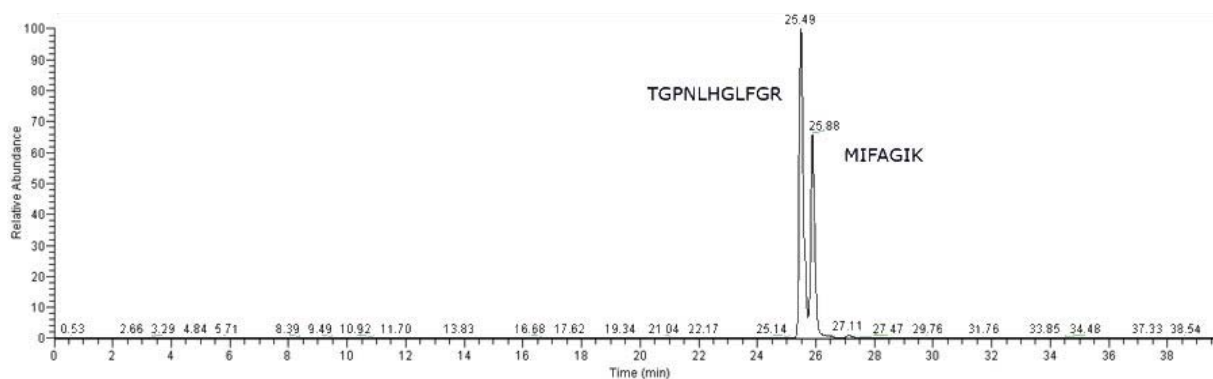


Рис. 1. Хроматографические пики пептидов TGNLHGLFGR и MIFAGIK при градиенте 1,7% В/мин.

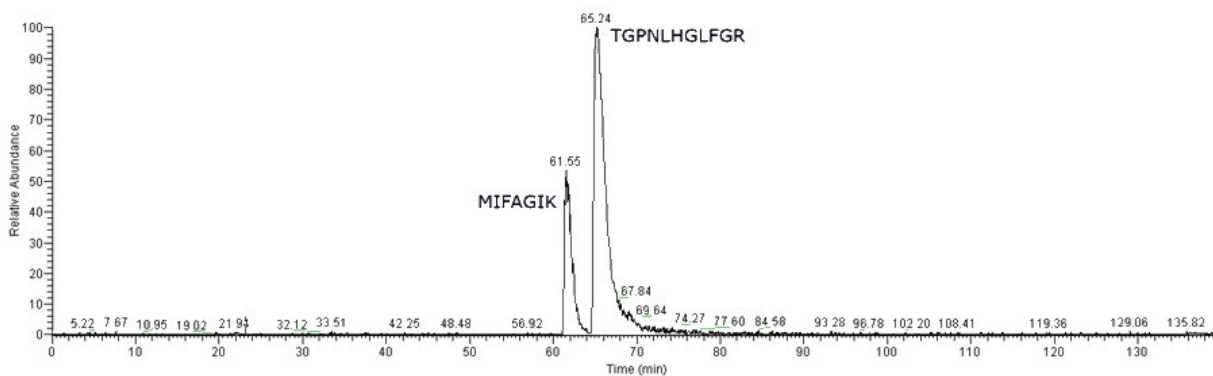


Рис. 2. Хроматографические пики пептидов TGNLHGLFGR и MIFAGIK при градиенте 0,3% В/мин.

В данной работе предложена процедура калибровки на основе линейного уравнения, которая позволит стандартизовать хроматографические данные и сделать время удерживания независимым от условий проведения хроматографического эксперимента для одного и того же образца. Для определения коэффициентов в уравнении калибровки в качестве "реперных" точек были выбраны рассчитанные с помощью модели BioLCCC⁴ теоретические времена удерживания для пептидов из дайджеста белка Cytochrome C (Dionex/LCPacking, USA). Также с помощью программы «Теоретический хроматограф», разработанной на основе модели BioLCCC решается проблема инверсии времён выхода пептидов (таблица 1). Предлагаемый подход продемонстрировал возможность получения нормализованных времён удерживания пептидов с точностью порядка 3%.

| Пептид | Градиент с 0 до 50%В за 30мин | | Градиент с 0 до 35%В за 120 мин | |
|-----------------------|-------------------------------|----------------|---------------------------------|----------------|
| | <i>TGNLHGLFGR</i> | <i>MIFAGIK</i> | <i>TGNLHGLFGR</i> | <i>MIFAGIK</i> |
| RT _{эксп} | 20.31 | 20.64 | 64.89 | 61.45 |
| RT _{BioLCCC} | 20.37 | 21.31 | 64.8 | 60.45 |
| RT _{SSRCalc} | 20.0 | 21.3 | 55.5 | 62.3 |

Таблица 1. Сравнение экспериментальных времен удерживания пептидов TGNLHGLFGR и MIFAGIK с предсказанными на основе разработанной авторами модели BioLCCC и альтернативного алгоритма предсказания времен удерживания SSRCalc, разработанного в Центре протеомики Манитобы (Канада), для различных наклонах градиента.

Эксперименты были выполнены на нанохроматографе Ultimate3000 (Dionex, USA), работающим в комбинации с масс-спектрометром LTQ-FT (ThermoFisher, Germany).

Работа была выполнена при финансовой поддержке Российского фонда фундаментальных исследований (РФФИ, гранты 06-04-49632 и 06-08-08085-офи), Международной ассоциации (INTAS, гранты № 04-83-2643 и Genomics 05-1000004-7759), Американского фонда гражданских исследований и развития (АФГИР, грант RUE1-000588-МО-05), и Отдела химических наук и материалов (программа ОХНМ 4.2 "Создание эффективных методов химического анализа и исследования структуры веществ и материалов").

Литература

1. *Smith R.D. et al.* Rapid quantitative measurements of proteomes by Fourier transform ion cyclotron resonance mass spectrometry // *Electrophoresis*. -2001. -N.22. - P.1652–1668.
2. *Lipton M.S. et al.*, Global analysis of the *Deinococcus radiodurans* proteome by using accurate mass tags // *Proc. Natl. Acad Sci. U S A*. -2002. –N. 99. –P.11049-54.
3. *Kieffer-Jaquinod S. et al.* Proteomic biomarkers discovery using the Accurate Mass and Time Tags approach // *Proceedings of the 23th Montreux LC-MS Symposium*. -2006.
4. *Gorshkov A.V. et al.* Liquid Chromatography at Critical Conditions: Towards a Comprehensive Approach to Sequence Dependent Retention Time Prediction // *Anal. Chem.* -2006. –N.78. –P.7770-7778.