ELSEVIER

Implications of column peak capacity on the separation of complex peptide mixtures in single- and two-dimensional high-performance liquid chromatography

Martin Gilar*, Amy E. Daly, Marianna Kele, Uwe D. Neue, John C. Gebler

Life Sciences Chemistry R&D, Waters Corporation, 34 Maple St., Milford, MA 01757, USA

Abstract

Column peak capacity was utilized as a measure of column efficiency for gradient elution conditions. Peak capacity was evaluated experimentally for reversed-phase (RP) and cation-exchange high-performance liquid chromatography (HPLC) columns, and compared to the values predicted from RP-HPLC gradient theory. The model was found to be useful for the prediction of peak capacity and productivity in single- and two-dimensional (2D) chromatography. Both theoretical prediction and experimental data suggest that the number of peaks separated in HPLC reaches an upper limit, despite using highly efficient columns or very shallow gradients. The practical peak capacity value is about several hundred for state-of-the-art RP-HPLC columns. Doubling the column length (efficiency) improves the peak capacity by only 40%, and proportionally increases both the separation time and the backpressure. Similarly, extremely shallow gradients have a positive effect on the peak capacity, but analysis becomes unacceptably long. The model predicts that a 2D-HPLC peak capacity of 15,000 can be achieved in 8 h using multiple fraction collection in the first dimension followed by fast RP-HPLC gradients employing short, but efficient columns in the second dimension.

1. Introduction

Modern high-performance liquid chromatography (HPLC) is expected to resolve highly complex samples. One example is proteome research that involves the analysis of protein/peptide samples, consisting of thousands of components at different concentration levels [1,2]. The representative proteome sample may contain 10,000–50,000 proteins, or in cases where proteins were digested prior to analysis, 100,000–500,000 peptides [3]. Despite recent progress in column technology and ultrahigh performance liquid chromatography (UPLC) [4–8], liquid chromatography still lacks the resolving power to separate samples of that complexity within a single analysis [9,10].

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is the standard technique for proteomic analysis (achieving the resolution of thousands or more components) [11]. Two-dimensional high-performance liquid chromatography (2D-HPLC) and multi-dimensional (MD)-HPLC are emerging for complex protein/peptide sample separations [6,8,12–16]. The development of 2D-HPLC techniques was driven by the prospect of achieving high peak capacity, good dynamic range, reproducibility, sensitive detection and quantitation, and automation of analysis. The on-line interface with MS via electrospray ionization (ESI) further improves the 2D-HPLC resolving power, since MS is considered to be an additional separation dimension.

Since the efficient separation of intact proteins in HPLC is difficult, and identification of large proteins by MS is challenging [17–19], 2D-HPLC separation is typically performed after protein digestion into peptides. However, the protein

---

* Corresponding author. Tel.: +1 508 482 2000; fax: +1 508 482 3100.
  E-mail address: Martin_Gilar@waters.com (M. Gilar).

fragmentation multiplies the sample complexity, which increases the demands on 2D-HPLC resolution.

The separation efficiency of HPLC columns under isocratic conditions is measured in theoretical plates. In gradient elution (typically used for separation of proteins and peptides), the separation performance is better described by column peak capacity P. Peak capacity represents the maximum theoretical number of components that can be separated on a column within a given gradient time. Assuming that separation selectivity in multiple dimensions is orthogonal, the peak capacity values can be multiplied [20]. For example, combining orthogonal separation modes with peak capacities of 20 and 100 theoretically provides for a peak capacity of 2000.

Current 2D-HPLC most often utilizes strong cation-exchange (SCX) chromatography in the first dimension, and reversed-phase (RP) HPLC in the second dimension [12–15]. The peak capacity of 2D-HPLC is often less than 5000, falling short of the required degree of separation. As a result, only the peptides (proteins) of greatest abundance are detected by MS. Further improvement in peak capacity and some means of sample pre-fractionation prior to 2D-HPLC is desirable. Although several manufacturers introduced 2D-HPLC systems to the market, it remains to be answered what is the realistic peak capacity that can be achieved with those instruments, and what is the practical level of complexity of proteomic samples that researchers can effectively analyze.

2D-HPLC is always combined with MS for proteomic research, which effectively translates into a 3D separation space with a peak capacity exceeding the chromatographic separation space. Considering the complexity of samples and current HPLC performance, one must assume there is always a certain degree of component overlap eluting from a column into MS. The estimate of a maximum achievable complexity and dynamic range of components that can be simultaneously detected and resolved by state-of-art MS is a non-trivial task, and deserves further investigation. Here we focus our interest on a chromatography and present a judicious look at the resolving power of 2D-HPLC in general, and more specifically for the RP- and SCX-HPLC separation modes. Column peak capacity was experimentally evaluated for a selected set of columns. The data were compared with the peak capacity values predicted from chromatographic theory. The impacts of column efficiency and gradient conditions on the achievable peak capacity were evaluated. We discuss the importance of HPLC peak capacity on the resolution and productivity of a 2D-HPLC system.

2. Experimental

2.1. Materials and reagents

Trifluoroacetic acid (TFA), 99.5%, was purchased from Pierce (Rockford, IL, USA). HPLC grade acetonitrile was purchased from J.T. Baker (Phillipsburg, NJ, USA). A Milli-Q system (Millipore, Bedford, MA, USA) was used to prepare deionized water (18 MΩ cm) for HPLC mobile phases. Peptides L2275, A6677, Bradykinin, Angiotensin II, Angiotensin I, Substance P, Renin Substrate, Insulin beta chain (oxidized), Melittin, sodium chloride, NaH_2PO_4, phosphoric acid, and proteins used for tryptic digestion were purchased from Sigma (St. Louis, MO, USA); sequencing grade Trypsin from Promega (Madison, WI, USA). MassPREP^{TM} MALDI matrix CHCA, and MassPREP^{TM} protein digest standards were obtained from Waters (Milford, MA, USA).

2.2. HPLC instrumentation and columns

HPLC experiments were carried out using the following instruments: model 2795 Alliance® HPLC system with a 996 photodiode array detector (Waters, Milford, MA, USA). HPLC columns used in this study were packed in house with Symmetry300^{TM} C_{18} sorbent (Waters, Milford, MA, USA). We used the following column dimensions ($L \times$ i.d.): 50 mm $\times$ 4.6 mm, 150 mm $\times$ 4.6 mm, 250 mm $\times$ 4.6 mm, and 300 mm $\times$ 4.6 mm as well as the following particle sizes: 3.5, 5, and 7 μm. RP-HPLC column configurations are detailed in Table 1. The columns were operated at 40 °C; the temperature was controlled by a built-in column heater. Mobile phases were A: 0.1% TFA in water and B: 0.08% TFA in acetonitrile. Gradients were 0–50% B, however gradient time varied for each column. HPLC conditions are given in the figure captions. A polySULFOETHYL Aspartamide^{TM} SCX 50 $\times$ 4.6 mm, 5 μm column (PolyLC, Columbia, MD, USA) was used for ion-exchange HPLC separations. The column was operated at 25 °C; mobile phases were A: 20 mM NaH_2PO_4, pH 3 with 25% acetonitrile and B: 20 mM NaH_2PO_4, pH 3 with 25% in acetonitrile with addition of 1 M NaCl. Gradients were run from 0 to 33% B in 20, 40 and 80 min. The tryptic digest of selected proteins (MassPREP^{TM} protein digest standards; enolase, phosphorylase, BSA, bovine hemoglobine α, bovine hemoglobine β, and ADH; Swiss-Prot accession numbers P00924, P00489, P02769, P02070, P01966, and P00330, respectively) was separated with a gradient from 0 to 9% B in 60 min. The SCX fractions were collected every 30 s up to 10 min, every 60 s between 10 and 20 min, and every 120 s from 20 to 60 min. Fractions were evaporated to dryness, reconstituted in 0.2 ml of 0.1% TFA in water and desalted by SPE. Oasis HLB μElution plate was first conditioned with 0.5 ml of acetonitrile, and 0.5 ml of 0.1% aqueous solution of TFA. Sample was loaded, washed with additional 0.5 ml of 0.1% aqueous solution of TFA, and eluted with 10 μl of 60% acetonitrile in water. About 0.5 μl of eluted solution was mixed with CHCA MALDI matrix and analyzed by Micromass M@LDI R TOF instrument from Waters (Milford MA, USA). Peptides eluting in each fraction were identified by MALDI-MS comparing the observed accurate mass with the theoretical mass of expected tryptic peptides.

Table 1
Columns and gradients used in study

Column number	Length × i.d. (mm)	Column type	Particle size (μm)	Column efficiency[a]	Void volume (ml)	Gradient time ($s = 0.0123$)[b]	Gradient time ($s = 0.0246$)[b]	Gradient time ($s = 0.0492$)[b]
1	50 × 4.6	Symmetry300 C_{18}	5	2800	0.66	37.5	17.8	8.9
2	150 × 4.6	Symmetry300 C_{18}	7	4930	1.80	100	50	25
3	150 × 4.6	Symmetry300 C_{18}	5	8400	1.86	100	50	25
4	150 × 4.6	Symmetry300 C_{18}	3.5	14230	1.90	100	50	25
5	300 × 4.6	Symmetry300 C_{18}	5	16790	3.46	186	93.1	46.5
							Gradient time	
6	50 × 4.6	PolySULFOETHYL Aspartamide™ SCX	5	–	0.7	20	40	80

[a] Calculated from Eqs. (4)–(6) for, $F = 0.75$ ml/min, and $D_m = 4.5 \times 10^{-10}$ m^2/s.

[b] Slope s was calculated from Eq. (2).

3. Results and discussion

3.1. Experimental peak capacity measurement

Peak capacity is defined as the maximum number of peaks that can be theoretically separated on a column at given chromatographic conditions with $R_s = 1$ [21–24]. Peak capacity can be calculated from the peak width w measured at 4σ (13.4% of peak height) and the gradient (separation) time t_g according to Eq. (1).

$$P = 1 + \frac{t_g}{w} \tag{1}$$

The peak capacity was experimentally estimated for a series of columns and gradient slopes (Table 1) using a peptide mixture containing nine components with a broad range of molecular weights. Selected chromatograms with the example of P calculation are shown in Fig. 1 (P is an average nine

peptides). The normalized gradient slope s, defined by Eq. (2), was an important parameter for the HPLC experiment [23]. Due to the differences in column length, the void time t_0 (measured as the elution time of unretained component) varied. The time of gradient t_g was therefore adjusted (see Table 1) to keep the gradient slope s constant. All peptides eluted in a gradient range from 0 to 50% acetonitrile, hence, the gradient range ΔC value used for calculation was 0.5.

$$s = \Delta C \frac{t_0}{t_g} \tag{2}$$

Fig. 1 shows the separation of the peptide mixture on a 150 mm × 4.6 mm Symmetry300 C_{18} column packed with 5 μm sorbent. The measured peak capacity was over 150 for a 25 min gradient. When a shallow 100 min gradient is used, P exceeds a value of 350. This does not mean that the column would resolve 350 peptides when such a complex mixture is injected on the column. According to statistical theory (as-
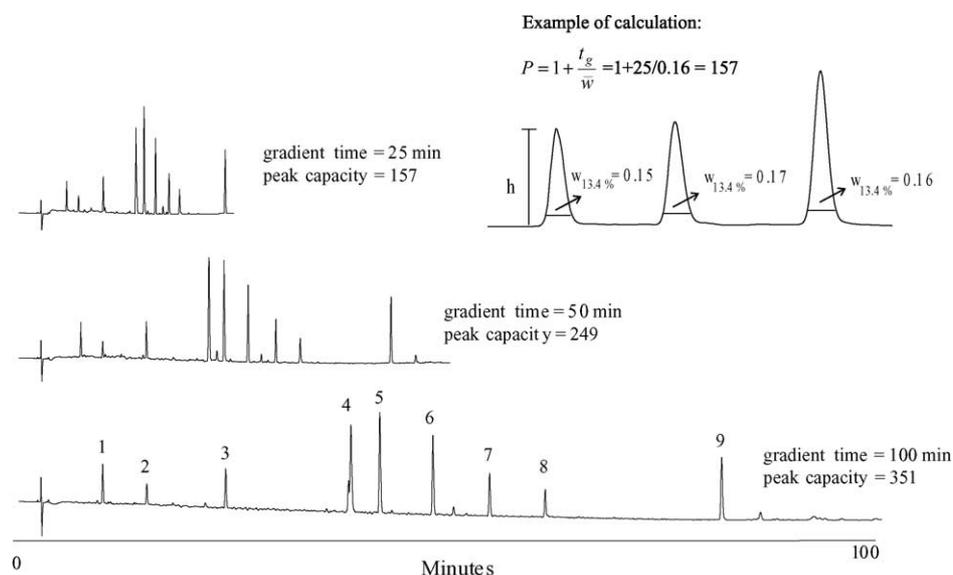

Fig. 1. An example of peak capacity measurement for a 150 mm × 4.6 mm, 5 μm Symmetry300 C_{18} column (column 3 in Table 1) using three different gradient slopes. The peak capacity was calculated from averaged peak widths as shown in the inset. The peak width depends on the nature of peptide; the peak width R.S.D. for the nine peptide standards was 11–28% (within a chromatographic run). The nine peptides used for peak capacity measurements were: (1) L2275; (2) 6677; (3) bradykinin; (4) angiotensin II; (5) angiotensin I; (6) substance P; (7) renin substrate; (8) insulin beta chain oxidized; and (9) melittin.

suming random retention of peptides), one expects to observe ~129 resolved peaks ($R_s = 1$) or less [25–27], with a number of mixture components co-eluting.

Evaluation of the SCX column peak capacity suggests that efficiency of peptide separation is lower than in RP-HPLC mode. A 50 mm × 4.6 mm SCX column provided peak capacities of 63, 85, and 113 using 20, 40, and 80 min gradients, respectively. Both mobile phase A and B contained 25% of acetonitrile. For other conditions, see Section 2 and Table 1. Peak capacity was calculated according Eq. (1) using the average peak width of four peptides (angiotensin I, angiotensin II, bradykinin, and L2275). The silica-based hydrophilic poly-SULFOETHYL Aspartamide^{TM} column was chosen for the experiment because it provided for the best column peak capacity from a series of other SCX columns, we evaluated for this study (silica-based sulfopropyl, sulfated polyDVB or sulfated polymehtacrylate materials; data not shown). However, even the polySULFOETHYL Aspartamide^{TM} columns have significantly lower peak capacity compared to RP-HPLC. The practical peak capacity for tryptic peptides is further reduced by the fact that the tryptic digest generates mostly 2+, and 3+ (4+ to the less extent) peptides. Those peptides therefore elute over the narrower range of the gradient (0–9% B, which represents 20–110 mM of Na$^+$ ions in 60 min). The suitable gradient strength is therefore less than used for SCX experimental peak capacity evaluation (20–350 mM of Na$^+$ ions).

3.2. Peak capacity prediction

For peptides, the peak width throughout the chromatogram is rather uniform. This is explained by the rather steep relationship of the logarithm of the retention factor with solvent composition for large molecular weight analytes. This results in a simplification of the peak capacity equation compared to small molecules and flat gradients [23]. The experimental data were compared with P values predicted from Eq. (3), where B is the slope of the function $\ln k$ (k is retention factor) versus solvent composition C, and N is the number of theoretical plates of the chromatographic column. The other parameters are the same as used in Eqs. (1) and (2).

$$P = 1 + \frac{\sqrt{N}}{4} \frac{B\Delta C}{B\Delta C(t_0/t_g) + 1} \tag{3}$$

Eq. (3) was derived from chromatographic theory for RP-HPLC mode for a linear solvent strength gradient [23,24]. The equation suggests that peak capacity increases proportionally to the square root of column efficiency. This trend is shown in Fig. 2A with other parameters of the equation kept constant. In practice this means that doubling column efficiency will improve P by only 40%. Similarly, when extending the gradient time in an effort to improve separation, the gains in P are diminishing (Fig. 2B).

The column efficiency N needed for P prediction was obtained from the column length and the height equivalent of a theoretical plate H, according to Eq. (4).

$$N = \frac{L}{H} \tag{4}$$

The H value was calculated from the van Deemter equation (Eq. (5)),

$$H = 2\lambda d_p + 2\chi \frac{D_m}{u} + c \frac{d_p^2}{D_m} u \tag{5}$$

where the value of the packed bed structural uniformity factor λ was set to 0.75, the obstruction factor χ to 0.5, and the c value was 0.166. The diffusion coefficient D_m value of 4.5×10^{-10} m^2/s used in this work was calculated from the Wilke–Chang equation [28]. The particle size d_p was substituted according to the column packing size used; the interstitial linear velocity of the mobile phase u was calculated from Eq. (6).

$$u = \frac{F}{\varepsilon_\tau \pi r^2} \tag{6}$$

F is the flow rate, r the radius of the chromatographic column, ε_τ is the fraction of the column occupied by the mobile phase (including the volume inside of porous particles), and π is 3.14. One may notice (Table 1) that the calculated values
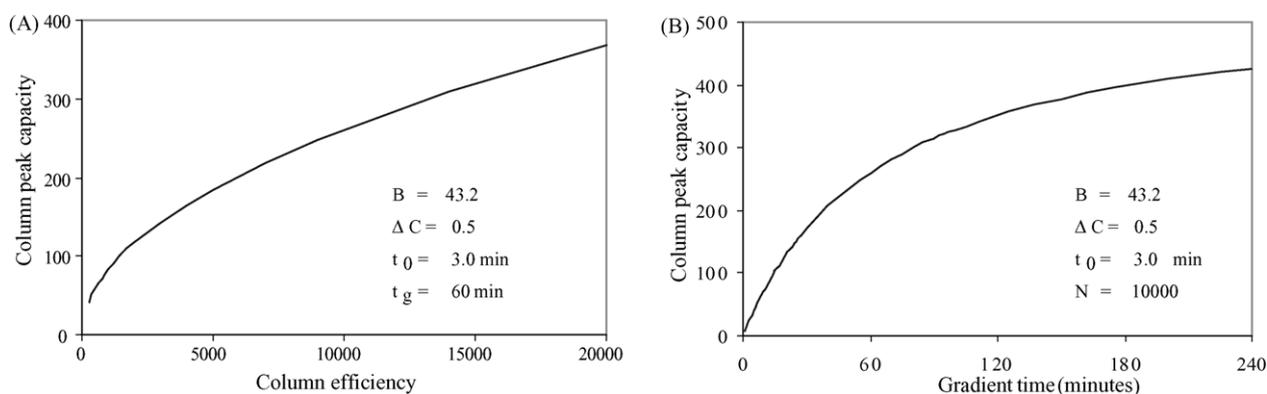

Fig. 2. Column peak capacity calculated from Eq. (3) varying (A) column efficiency or (B) gradient time. Other parameters were kept constant as indicated in the figures.

of theoretical plates N are significantly lower than those expected for small molecules. For example, the experimentally measured column efficiency using acenapthene for column 5 (Table 1) was ∼20,000. This discrepancy is due to the smaller diffusion coefficient D_m of peptides. The optimal flow rate for peptide separations is between 0.1 and 0.3 ml per min (calculated for a 4.6 mm column i.d.) for a chromatographic packing particle size ranging from 7 to 3.5 µm, respectively.

The value of B needed for the peak capacity calculation (Eq. (3)) is not a constant, but varies according to the analyte molecular weight [23,24,29,30]. The slope B of the function $\ln k$ versus c increases with peptide size. Since the experimental measurement of B values of macromolecules under isocratic elution is difficult, few reliable estimates can be found in literature [31–33]. In this work, the B values were calculated from Eq. (7).

$$\ln B = 0.6915 \ln(MW) - 1.49 \tag{7}$$

Assuming that the MW of the average tryptic peptide (and peptides used in this study) is approximately 2000 Da, the B value used for peak capacity calculations was 43.2. This value compares well to other published estimates such as $B = 31$ [24], or $B = 51$–62 [29,30].

Fig. 3 shows the peak capacity as a function of both column efficiency and gradient time. The graph suggests that the practical number of peaks separated on columns reaches an upper limit, despite using highly efficient columns and shal-

Fig. 4. Comparison of predicted (lines) and experimentally measured (isolated data points) column peak capacity values at three different gradient slopes. The dotted line represents slope $s = 0.0492$, solid line slope $s = 0.0246$, and scattered line slope $s = 0.0123$. The experimentally measured peak capacity values are indicated in the graph; triangles represent column 1, squares column 2, crosses column 3, diamonds column 4, and circles column 5. For column assignment see Table 1.

low gradients. It should also be pointed out that the gain in P is achieved at the expense of time and backpressure. The surface of the peak capacity function reaches a plateau close to the value of 1400–1600 for B values ranging from 40 to 50, respectively.

The experimentally acquired data were compared to the predicted P in Fig. 4. Using the values of $B = 43.2$ and $D_m = 4.5 \times 10^{-10}$ m^2/s, a good fit between predicted and measured peak capacity was observed. The experimentally measured P values are listed in Fig. 4. This confirms that Eq. (3) describes the chromatographic behavior of peptides with good accuracy.

3.3. 1D and 2D-HPLC peak capacity and throughput

The separation of complex mixtures necessitates the use of highly efficient columns. Two different approaches can be applied to fabricate columns with high efficiency: (i) pack extensively long columns; or (ii) pack the column with a small particle size sorbent while keeping the column length constant. Employing long columns for the separation (connecting columns in series) has the apparent limitations of increased backpressure and increased analysis time [34]. As shown in Table 1 and Eq. (2), the gradient time increases proportionally with the t_0 (V_0) value. As the void volume of the column increases, t_g soon becomes unacceptably long. It has to be noted that using longer columns without adjusting t_g results in marginal gains in peak capacity. For example, increasing the column length from a 50 to 150 mm, while keeping the gradient time constant causes the normalized gradient slope s to be effectively three times greater (V_0 of 150 mm column is three times bigger). The experimentally measured P for 50
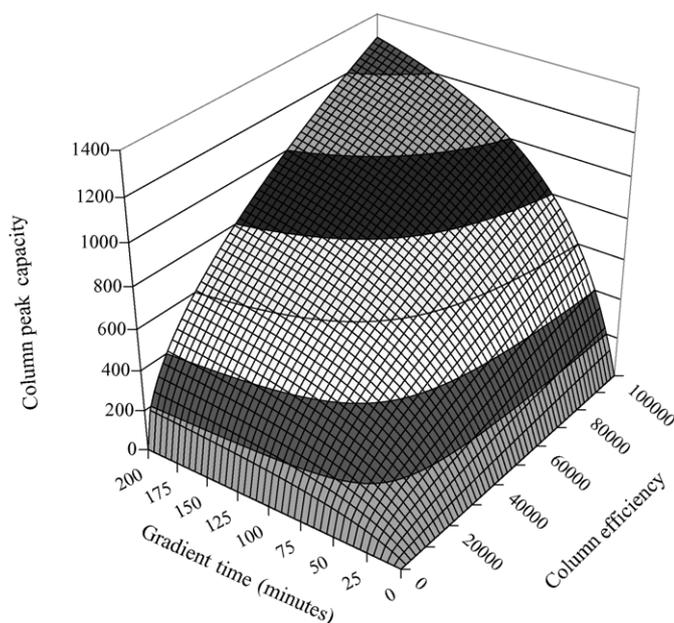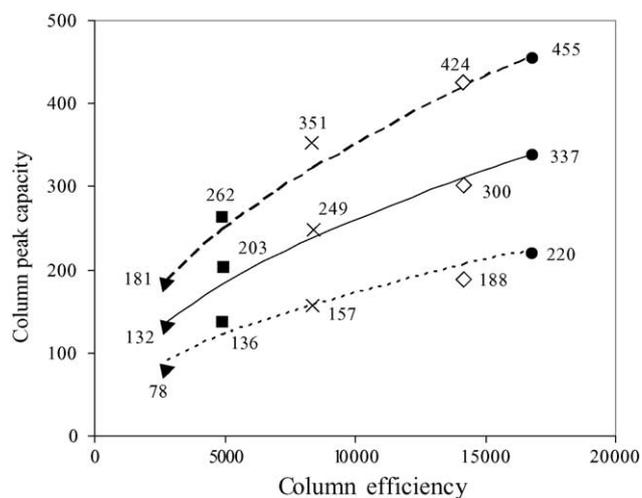

Fig. 3. Calculated column peak capacity (Eq. (3)) as a function of both column efficiency and gradient time. The graph of peak capacity plateaus at long gradient times and highly efficient columns. The graph was calculated using $B = 43.2$ and $\Delta C = 0.5$. The t_0 was set to 3 min, which approximately represents the void time of a 250 mm × 4.6 mm long column at 1 ml/min flow rate. The efficiency of a 250 mm column packed with 5, 3.5, 1.7 and 1 µm sorbent calculated for peptide separation is 14,000, 24,000, 61,000, and 108,000, respectively.
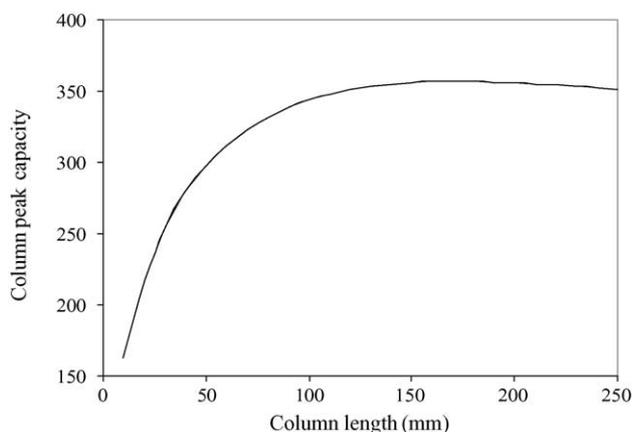
Fig. 5. The column peak capacity calculated (Eq. (3)) for various column lengths. Gradient time and other parameters were kept constant: $t_g = 60$ min, $B = 43.2$, $\Delta C = 0.5$, $t_0 = 0.16$ ml/cm, and $N = 1000$/cm (column packed with 3.5 μm sorbent).

and 150 mm columns was 197 and 249, respectively, which represents only a 26 % increase in peak capacity (both separations were performed with a 50 min gradient; for other conditions see Fig. 1).

The impact of column length on the peak capacity, while keeping the particle size and gradient time constant ($d_p = 3.5$ μm, $t_g = 60$), is illustrated in Fig. 5. The P (calculated from equation 3) reaches a maximum for ~150 mm column length, but decreases for longer columns. Apparently, the increase in the normalized gradient slope value abolishes the contribution of the greater N value of longer columns.

The strategy of improving column peak capacity by using a small particle size chromatographic packing seems to be more promising than using longer columns. The limitation of this approach, however, is a sharp increase in backpressure. For isocratic chromatography, this has previously been studied in detail by Guiochon [35]. One of us has previously dealt with similar complications under gradient conditions

in a theoretical study focusing on small molecules [36]. In a gradient of a fixed run time, the column characteristics such as column length and particle size interact with the operational parameters such as the linear velocity and the gradient span in a complex fashion. Note that in Eq. (3) the plate count and the gradient span (Eq. (2)) change as we change the void time t_0 (and thus the flow rate). In the discussion here, we will deal with a few specific issues relevant for the separation of peptides.

In Fig. 6A, the peak capacity for three columns packed with 3.5 μm particles is plotted versus the flow rate for a 60 min gradient separation of peptides. The column lengths were 250, 150 and 50 mm, column i.d. was 4.6 mm. It is interesting to see that the performance of all three columns is very similar at a flow rate of about 0.5 ml/min. This is in agreement with the practical experience of many chromatographers who find that for peptide separations under these and similar circumstances the column length plays only a subordinate role. Only at a flow rate of about 1 ml/min do the longer columns outperform the 5 cm column, but at the same time, the peak capacity of the 250 mm column and the 150 mm column are indistinguishable. These interesting effects are due to the fact that at a fixed gradient run time, the longer columns have an improved plate count, but this effect is counterbalanced by the smaller gradient span.

On the other hand, significant gains in separation power can be made by keeping the column length constant and reducing the particle size, as it is shown in Fig. 6B. In this case, we have compared 50 mm columns packed with 5, 3.5 and 1.8 μm particles for a 30 min gradient. For the 5 μm column, the maximum peak capacity of 185 for this gradient is reached at a flow rate of 0.95 ml/min. The performance improves for the 3.5 μm column: the maximum peak capacity is now 245 at about 1.1 ml/min. The by far better results though are obtained with the 1.8 μm column: a peak capacity of 403 at a flow rate of 1.5 ml/min. The price of such a performance improvement is the column backpressure, as shown
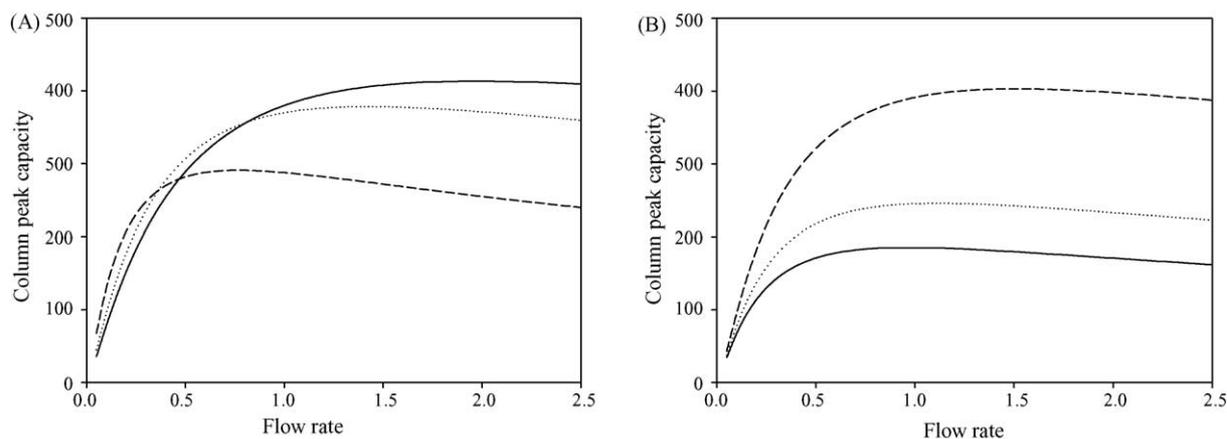

Fig. 6. The column peak capacity expressed as a function of the flow rate for various column lengths (A) or columns packed with different particle size sorbent (B). The solid line represents 250 mm column, dotted line 150 mm column, and dashed line 50 mm column (A); the gradient time: $t_g = 60$ min. In case (B), the dimension of all three columns were 50 mm × 4.6 mm; solid line represents column packed with 5 μm sorbent, dotted line column with 3.5 μm sorbent, and dashed line column packed with 1.8 μm sorbent; the gradient time used for calculation was $t_g = 30$ min.

Table 2
Impact of sorbent particle size on the column efficiency, peak capacity and operating pressure

Flow rate (ml/min)	Column $L \times$ i.d. (mm × mm)	Particle size (μm)	Column efficiency	Peak capacity[a]	Pressure (MPa)[b]	Pressure (psi)[b]
0.95	50 × 4.6	5.0	2340	185	2.3	343
1.10	50 × 4.6	3.5	3780	246	5.5	811
1.50	50 × 4.6	1.8	8820	403	28.4	4181

[a] Calculated using $t_g = 30$ min, $\Delta C = 0.5$, $B = 43.2$, $D_m = 4.5 \times 10^{-10}$ m^2/s, $V_0 = 0.58$ ml, $\varepsilon_i = 0.35$, and viscosity = 0.7 cP.
[b] Calculated from Kozeny–Carman equation.

in Table 2. The efficiency, peak capacity, and backpressure is calculated for 50×4.6 mm columns packed with 5, 3.5, and 1.8 μm porous sorbents.

The main benefit of using short, but efficient columns is the productivity of separation, defined as the number of separated components (P) per unit of time. The 3D plot of the separation productivity (Fig. 7) shows that although the longest column and gradient is expected to provide for the highest P (see e.g. Fig. 1), the productivity is low. The best productivity is achieved using fast gradients and relatively short columns. This has direct implication on the separation throughput in 2D-HPLC.

The overall 2D-HPLC separation throughput depends strongly on the peak capacity and separation speed in both dimensions. A current approach for 2D-HPLC separation employs a fast step gradient fractionation by SCX-HPLC, followed by long RP-HPLC separations. The second dimension is, therefore, the throughput limiting step. The step gradient fractionation in the first dimension (∼10 fractions) often under-utilizes the full separation potential of SCX-HPLC. It has been suggested that employing a long linear gradient with multiple collected fractions in the first dimension combined with fast, but efficient separation in the second may provide better overall productivity for 2D-HPLC [6,17,37,38]. Table 3 outlines the impact of various 2D-HPLC setups on the separation throughput (productivity), defined as the total achieved peak capacity per unit of time.

The first three examples (Table 3) consider a step gradient elution from the first dimension (e.g. SCX 10 fractions, peak capacity of fist LC dimension is therefore considered to be 10). Extending the length of the column (and the gradient time) in the RP-HPLC dimension, it is possible to achieve an overall peak capacity of several thousand. Combining two 250 mm long columns, it is theoretically possible to exceed a peak capacity of 10,000. However, productivity is low, which results in a very long total analysis time. The other three scenarios assume a linear gradient in the first dimension with 20–80 collected fractions. Using a smaller particle size sor-
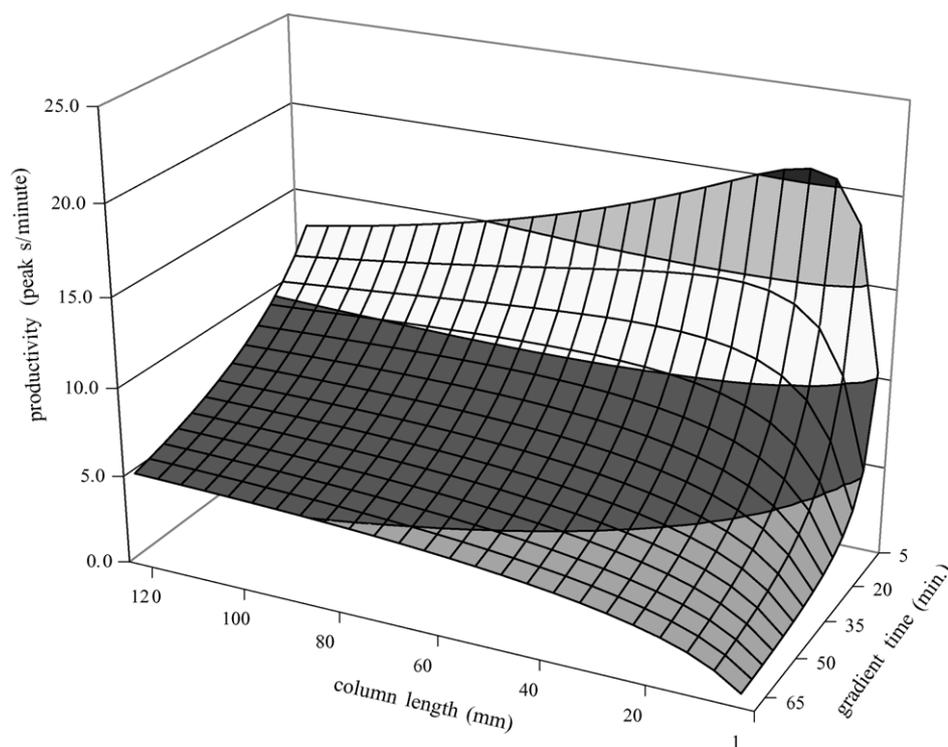

Fig. 7. Productivity of peptide separation defined as peak capacity/minute. The maximum productivity is achieved for rather short columns and fast gradients. Parameters for calculation were similar to those in Fig. 5.

Table 3
Analysis of 2D-HPLC peak capacity and productivity

First LC dimension no. of fractions	RP-HPLC gradients (min)	Column L (mm)	Particle size (µm)	RP-HPLC peak capacity[a]	Total 2D peak capacity[b]	Total run time (h)	Productivity peaks/min
10	60	150	5	276	2764	10	4.6
10	120	250	5	382	3849	20	3.2
10	240	500	5	540	5400	40	2.3
20	240	250	3.5	612	12241	80	2.6
40	60	150	3.5	365	14614	40	6.1
80	15	50	3.5	186	14920	20	12.4
80	6	20	1.8	188	15065	8	31.4

A different number of fractions collected in the first dimension and, different columns and gradient times in the second dimension are considered.

[a] Calculated using $F = 1.0$ ml/min, $\Delta C = 0.5$, $B = 43.2$, $D_m = 4.5 \times 10^{-10}$ m^2/s, and $t_0 = 0.12$ min/cm of column.

[b] Calculated as number of first dimension LC fractions $\times P$ in RP-HPLC.

bent and shorter columns in the RP-HPLC dimension, greater productivity is achieved. Table 3 predicts that a peak capacity of about 14,000 can be achieved using an on-line 2D-HPLC system using various experimental setups. The use of short columns and gradients in the second dimension provides for greater productivity of the separation. Since the first and second dimension are operated concurrently or on-line in a comprehensive mode, the first dimension adds only minimal time to the overall separation scheme and can be neglected in the productivity calculation in Table 3.

Since mass spectrometry is used as the final detection technique after 2D-HPLC separation, the duty cycle of MS instruments for MS/MS peptide identification has to be taken into a consideration. State-of-art mass spectrometers have an MS/MS duty cycle of 1–2 s, which corresponds well to the expected widths of chromatographic peaks when using short, efficient columns (Table 3). However, if a significantly more complex sample containing more than 14,000 components is injected onto the 2D-HPLC systems proposed in Table 3, multiple sample components will overlap and obscure the identification of some peaks of interest. In such a case, only the most abundant peptides will be selected and identified for data dependent MS/MS analysis. In fact, this is a drawback of 2D-HPLC/MS analysis as practiced today for the separation of highly complex samples in separation systems with insufficient peak capacity. The pre-fractionation of complex samples (on the intact protein level) prior to 2D-HPLC should alleviate this problem, improve the protein sequence coverage, and, consequently, the identification reliability.

3.4. Impact of separation selectivity on 1D- and 2D-HPLC

As discussed earlier, an improvement in peak capacity in 1D- and 2D-HPLC is achieved at the expense of increased column pressure and separation time. Employing more than two separation dimensions holds some promise for further improvement of total peak capacity. However, the separation selectivity has to be orthogonal in all dimensions, which may be difficult to achieve. Published reports suggest that even the most popular combination of SCX and RP-HPLC has limited orthogonality [1,15]. Proteomic samples prepared by tryptic

digestion of proteins are comprised mainly of peptides that are 10–20 amino acids in length. Those peptides are not distributed randomly in the RP-HPLC separation space; most of the peptides elute clustered in the middle of the analysis, eluting between 15 and 35% of acetonitrile content in the mobile phase. Also, the separation in SCX-HPLC is based on charge, which means that 2+, 3+ and 4+ charged peptides (most common peptide charges) effectively elute in three clusters [1]. The observed separation between peptides of the same charge is presumably due to the secondary hydrophobic/hydrophilic interaction with the SCX sorbent. The most common tryptic peptides are doubly charged [1]. They elute within a rather narrow retention time window, which makes collection of multiple fractions from SCX difficult. The representative example of the SCX selectivity for moderately complex mixture of peptides generated by tryptic digestion of proteins (see Section 2) is shown in Fig. 8. Note that unlike in RP-HPLC, the peak width increases with the elution time, which complicates the peak capacity estimate. In agreement with the reports [1], the most numerous group of peptides (in this example ∼70% of all peptides in the sample) has 2+ charge and elutes between 2 and 10 min. This essentially means that although the multiple fraction collection from SCX is possible, the separation of complex peptide samples will be less successful than predicted by the model. To fully utilize the potential of 2D-HPLC outlined in the Table 3, it is desirable to employ the first dimension LC technique with a high peak capacity and orthogonal selectivity with RP-HPLC.

Other separation modes have been proposed for 1D-, 2D- or multi-dimensional separations, such as size-exclusion chromatography (SEC) [6,17], hydrophobic interaction chromatography (HIC) [39], and hydrophilic interaction chromatography (HILIC) [40,41]. However, those modes are only partially orthogonal to RP-HPLC (SEC separates peptides according to length; peptide lengths correlate with their hydrophobicity). RP-HPLC is currently the most efficient technique, and due to the compatibility with MS, it remains indispensable as a last separation step prior to MS analysis.

A promising approach to enhance the peak capacity of the separation system is a fractionation of intact proteins (SEC, ion-exchange HPLC, chromatofocusing, or 1D PAGE). Isolated fractions can be digested and submitted for further anal-
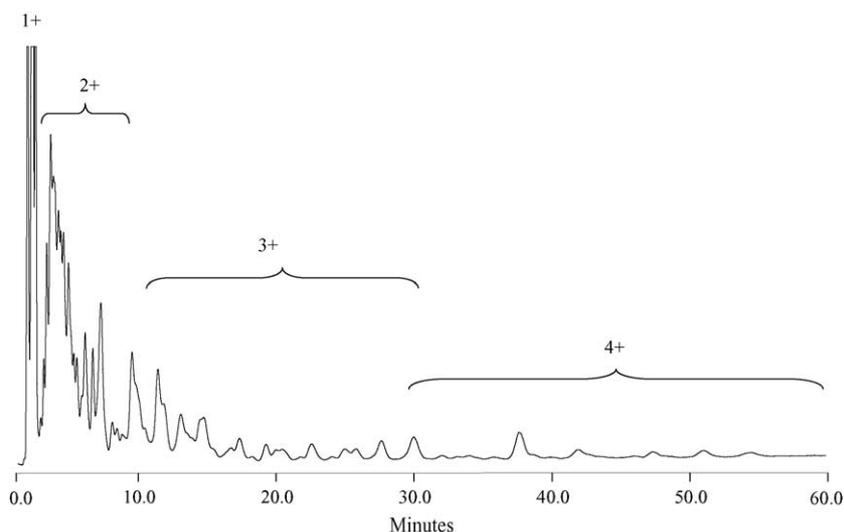
Fig. 8. SCX-HPLC separation of tryptic peptides using 50 mm × 4.6 mm PolySULFOETHYL Aspartamide™ column packed with 5 μm sorbent. Peptides elute according to their charge; a charge and the approximate elution range of peptides are indicated in the figure. Tryptic peptides were prepared by digestion of six proteins; for details see Section 2.

ysis. Using this approach, the sample complexity is reduced, however, the number of fractions to be processed by 2D-HPLC increases. Therefore, the productivity of 2D-HPLC becomes greatly important for the ultimate throughput of proteome analysis.

3.5. Limitations of the peak capacity prediction model

The peak capacity predicted from Eq. (3) describes the experimentally measured RP-HPLC data with good accuracy. It has to be noted, however, that the peptide D_m and B values may vary (peptides with different molecular weights). In this work, we averaged the P values obtained for nine peptides. However, the choice of the peptide test set may result in some differences in experimentally measured peak capacity. Additionally, lower P values were found when using aqueous formic acid instead of TFA as the mobile phase modifier.

It may be difficult to apply the model directly to capillary and nano-HPLC columns. Since their efficiency usually does not match that of 4.6 mm i.d. columns, and extra-column contributions to the peak width become more prominent, the peak capacity of capillary/nano-HPLC systems is expected to be lower than predicted.

The predictions made in the figures and tables assume that columns are packed with porous particles. When non-porous chromatographic sorbents are used, the slow peptide diffusion (mass transfer in stationary phase) is no longer a factor limiting the achievable peak capacity. Therefore, it becomes possible to perform faster separations and achieve higher P overall. The drawback of using non-porous particles is the low mass load capacity of such columns. Column overloading typically results in deterioration of peak shape and P. Most of the currently used systems employ porous particles, at least in the second separation dimension.

It is important to notice that the predicted P is the maximum number of peaks that can be theoretically separated on a column (not the actual number of peaks observed). In practice, there are several reasons why that ideal separation cannot be achieved. Some peptides are retained with similar selectivity and may be incompletely resolved, while some separation space in the chromatogram remains unpopulated by peaks. Giddings et al. [20,27] predicted that when injecting a random sample of equal complexity to the peak capacity of column, only 37% of peaks will be resolved in the chromatogram with $R_s = 1$. This means that at least two, three or four peptides are likely to co-elute under an observed peak. In our experience, peptide samples (protein digests) can be considered as random samples, and their separation behavior correlates well with the proposed Giddings model. This separation behavior in 1D-HPLC obviously translates also into 2D-HPLC, resulting in a lower number of ideally separated peptides than predicted by the P value of the 2D separation system. The situation is further complicated by the fact that subsequent fractions collected in the first dimension will contain some of the same components, as the fractionation splits some of the peaks into the neighboring fractions. The resulting complexity of the fractions submitted to the second separation dimension is therefore greater than considered by the model.

4. Conclusions

A mathematical model evaluated for column peak capacity prediction in RP-HPLC is in good agreement with the experimentally measured data. The model was used to investigate peak capacity for various setups of 1D- and 2D-HPLC, and to evaluate the productivity of separation of complex peptide samples. Results suggest that current 2D-HPLC systems us-

ing step gradient elution are capable of separating moderately complex samples, achieving a peak capacity of about 5000. The achieved peak capacity has a direct impact on the number of peptides that can be successfully identified by on-line MS/MS. The described model predicts that a peak capacity of approximately 15,000 can be achieved in 8 h of separation time. More advanced separations of greatly complex proteomic samples may require pre-fractionation (on either the peptide or protein level) prior to 2D-HPLC/MS analysis. Such multi-dimensional schemes demand a fast and efficient separation in the last separation step, which has to be capable of processing multiple fractions generated in previous (orthogonal) separation dimensions.

Although advances in MS and MS/MS instrumentation have dramatically improved scientists abilities to analyze complex samples, the separation of complex samples remains a bottleneck for proteome analysis. It remains to be seen whether the peak capacity achievable with commercially available 2D-HPLC systems is sufficient and practical for proteome studies.

Acknowledgement

References

[1] J. Peng, J.E. Elias, C.C. Thoreen, L.J. Licklider, S.P. Gygi, J. Proteome Res. 2 (2003) 43.
[2] J.R. Yates 3rd, Electrophoresis 19 (1998) 893.
[3] T. Wehr, LC–GC North Am. 20 (2002) 954.
[4] J.E. MacNair, K.D. Patel, J.W. Jorgenson, Anal. Chem. 71 (1999) 700.
[5] L. Tolley, J.W. Jorgenson, M.A. Moseley, Anal. Chem. 73 (2001) 2985.
[6] G.J. Opiteck, J.W. Jorgenson, J.E. MacNair, M.A. Moseley 3rd, Anal. Chem. 69 (1997) 2283.
[7] Y. Shen, R.J. Moore, R. Zhao, J. Blonder, D.L. Auberry, C. Masselon, L. Pasa-Tolic, K.K. Hixson, K.J. Auberry, R.D. Smith, Anal. Chem. 75 (2003) 3264.
[8] Y. Shen, J.M. Jacobs, D.G. Camp 2nd, R. Fang, R.J. Moore, R.D. Smith, W. Xiao, R.W. Davis, R.G. Tompkins, Anal. Chem. 76 (2004) 1134.
[9] Y. Shen, J.M. Jacobs, D.G. Camp 2nd, R. Fang, R.J. Moore, R.D. Smith, W. Xiao, R.W. Davis, R.G. Tompkins, R. Zhao, S.J. Berger, G.A. Anderson, N. Rodriguez, Anal. Chem. 76 (2004) 1134.
[10] Y. Shen, R. Zhao, S.J. Berger, G.A. Anderson, N. Rodriguez, R.D. Smith, Anal. Chem. 74 (2002) 4235.
[11] S.P. Gygi, G.L. Corthals, Y. Zhang, Y. Rochon, R. Aebersold, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 9390.
[12] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.
[13] D.A. Wolters, M.P. Washburn, J.R. Yates 3rd, D. Wolters, Anal. Chem. 73 (2001) 5683.
[14] M.P. Washburn, D. Wolters, J.R. Yates 3rd, Nat. Biotechnol. 19 (2001) 242.
[15] K. Wagner, T. Miliotis, G. Marko-Varga, R. Bischoff, K.K. Unger, Anal. Chem. 74 (2002) 809.
[16] B.J. Cargile, J.L. Bundy, T.W. Freeman, J.L. Stephenson Jr., J. Proteome Res. 3 (2004) 112.
[17] G.J. Opiteck, S.M. Ramirez, J.W. Jorgenson, M.A. Moseley 3rd, Anal. Biochem. 258 (1998) 349.
[18] H. Liu, S.J. Berger, A.B. Chakraborty, R.S. Plumb, S.A. Cohen, J. Chromatogr. B 782 (2002) 267.
[19] D.B. Wall, M.T. Kachman, S. Gong, R. Hinderer, S. Parus, D.E. Misek, S.M. Hanash, D.M. Lubman, Anal. Chem. 72 (2000) 1099.
[20] J.C. Giddings, J. High. Res. Chromatogr. 10 (1987) 319.
[21] M.A. Stadalius, B.F. Ghrist, L.R. Snyder, J. Chromatogr. 387 (1987) 21.
[22] B.F. Ghrist, M.A. Stadalius, L.R. Snyder, J. Chromatogr. 387 (1987) 1.
[23] U.D. Neue, J.L. Carmody, Y.F. Cheng, Z. Lu, C.H. Phoebe, T.E. Wheat, in: P. Brown, E. Grushka (Eds.), Design of Rapid Gradient Methods for the Analysis of Combinatorial Chemistry Libraries and the Preparation of Pure Compound, Marcel Dekker, New York, 2001, p. 93.
[24] L.R. Snyder, M.A. Stadalius, in: C. Horvath (Ed.), High-Performance Liquid Chromatography Separations of Large Molecules: A General Model, Academic Press, Orlando, FL, 1986, p. 195.
[25] J.C. Giddings, Anal. Chem. 39 (1967) 1027.
[26] J.M. Davis, J.C. Giddings, Anal. Chem. 57 (1985) 2178.
[27] J.M. Davis, J.C. Giddings, Anal. Chem. 55 (1983) 418.
[28] U.D. Neue, HPLC Columns: Theory, Technology and Practice, Wiley/VCH, New York, 1997.
[29] W.S. Hancock, R.C. Chloupek, J.J. Kirkland, L.R. Snyder, J. Chromatogr. A 686 (1994) 31.
[30] R.C. Chloupek, W.S. Hancock, B.A. Marchylo, J.J. Kirkland, B.E. Boyes, L.R. Snyder, J. Chromatogr. A 686 (1994) 45.
[31] S. Terabe, H. Nishi, T. Ando, J. Chromatogr. 212 (1981) 295.
[32] M. Gilar, K.J. Fountain, Y. Budman, U.D. Neue, K.R. Yardley, P.D. Rainville, R.J. Russell 2nd, J.C. Gebler, J. Chromatogr. A 958 (2002) 167.
[33] B.G. Belenkii, A.M. Podkladenko, O.I. Kurenbin, V.G. Mal'tsev, D.G. Nasledov, S.A. Trushin, J. Chromatogr. 645 (1993) 1.
[34] J.W. Dolan, L.R. Snyder, N.M. Djordjevic, D.W. Hill, T.J. Waeghe, J. Chromatogr. A 857 (1999) 1.
[35] G. Guiochon, in: C. Horvath (Ed.), High-Performance Liquid Chromatography, Advances and Perspectives, 1980, p. 1.
[36] U.D. Neue, J.R. Mazzeo, J. Sep. Sci. 24 (2001) 921.
[37] N. Tanaka, H. Kimura, D. Tokuda, K. Hosoya, T. Ikegami, N. Ishizuka, H. Minakuchi, K. Nakanishi, Y. Shintani, M. Furuno, K. Cabrera, Anal. Chem. 76 (2004) 1273.
[38] C.J. Venkatramani, Y. Zelechonok, Anal. Chem. 75 (2003) 3152.
[39] S.A. Berkowitz, Adv. Chromatogr. 29 (1989) 175.
[40] T. Yoshida, Anal. Chem. 69 (1997) 3038.
[41] A.J. Alpert, J. Chromatogr. 499 (1990) 177.