Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography A, 1169 (2007) 139–150
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Peak capacity in gradient reversed-phase liquid chromatography of biopolymers Theoretical and practical implications for the separation of oligonucleotides

Martin Gilar*, Uwe D. Neue

Waters Corporation, 34 Maple Street, Milford, MA 01757, USA

Received 15 June 2007; received in revised form 5 September 2007; accepted 6 September 2007
Available online 11 September 2007

Abstract

Reversed-phase ultra-performance liquid chromatography was used for biopolymer separations in isocratic and gradient mode. The gradient elution mode was employed to estimate the optimal mobile phase flow rate to obtain the best column efficiency and the peak capacity for three classes of analytes: peptides, oligonucleotides and proteins. The results indicate that the flow rate of the Van Deemter optimum for 2.1 mm I.D. columns packed with a porous 1.7 μm C_{18} sorbent is below 0.2 mL/min for our analytes. However, the maximum peak capacity is achieved at flow rates between 0.15 and 1.0 mL/min, depending on the molecular weight of the analyte. The isocratic separation mode was utilized to measure the dependence of the retention factor on the mobile phase composition. Constants derived from isocratic experiments were utilized in a mathematical model based on gradient theory. Column peak capacity was predicted as a function of flow rate, gradient slope and column length. Predicted peak capacity trends were compared to experimental results.

Keywords: Peak capacity; Biopolymers; Ultra-performance liquid chromatography; Van Deemter; Gradient theory

1. Introduction

The development and commercialization of novel biopharmaceuticals based on peptides, proteins, and oligonucleotides renewed an interest in liquid chromatography (LC) as a tool for biopolymers analysis. While LC theory and practice are well developed for so-called "small molecules" (a term customarily used for analytes with the molecular weight below ~500 Da), the separation principles of "large molecules" are understood to a lesser degree [1,2]. The reasons are several, among others: (i) The chromatographic behavior of biopolymers is complex, comprising effects such as the changes in molecular conformation [3], incomplete recovery from the column [4], ghosting, peak tailing, etc. This is especially noticeable for protein separations in reversed-phase mode. (ii) Dedicated sorbents for the separation of biopolymers were developed later than the ones for "small molecules". (iii) Biopolymer samples became readily available only after 1990. When the peptide and oligonucleotide synthesis became routine, the analysts gained access to structurally related samples of their choice for systematic analytical studies.

Chromatographic theory developed for small molecules certainly applies to biopolymers as well; however, practitioners are well aware of characteristic differences in their retention behavior. For example, the isocratic retention of biopolymers (or in general all polymers and large molecular weight analytes) is extremely sensitive to the changes in mobile phase elution strength [2,5]. Since the reversed-phase (RP)LC isocratic separation of biopolymers is difficult and impractical, only few such reports can be found in the literature [6–9], and even fewer have attempted to measure Van Deemter curves to establish an optimal mobile phase linear velocity [10].

One can circumvent the practical difficulties of column efficiency estimation by using Van Deemter (or others) chromatographic models, providing that the molecular diffusion coefficient D_m of the analytes is known, and no irregular adsorption–desorption behavior takes place. Unfortunately, the

* Corresponding author. Tel.: +1 508 482 2000; fax: +1 508 482 3625.
E-mail address: Martin_Gilar@waters.com (M. Gilar).

0021-9673/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2007.09.005

D_m values of biopolymers are rarely known [11]. The diffusion coefficients can be approximated using various models [9,12]. Calculations are complicated by the fact that the actual separations are often carried out at elevated temperature [13,14], and that solute diffusion in the stationary phase depends on the sorbent pore size [9,10]. Therefore, the reliable prediction of optimal LC conditions (column efficiency) for the separation of peptides, oligonucleotides and proteins is not a trivial task.

As pointed out above, the separation of biopolymers is typically performed in a gradient mode. Separation performance can be measured as resolution between individual peaks or as peak capacity [9,15,16], which represents the maximum theoretical number of peaks that can be resolved within the gradient time. The peak capacity of RPLC for peptides has been recently evaluated both theoretically and experimentally [17].

Although it is possible to apply the peak capacity theory for the prediction of optimal separation conditions, the challenges of this approach are similar as outlined above. Any prediction requires the knowledge of D_m, retention parameters S or B (defined by Snyder [9] or Neue [15,16] as the slope of $\log k$ or $\ln k$ versus mobile phase composition Φ), and k_0 (extrapolated retention coefficient of the solute in a fully aqueous mobile phase). Unfortunately, none of these parameters are typically known. These constants are characteristic for each solute and must be measured experimentally in isocratic or gradient LC mode.

In this work, we utilized an ultra-performance liquid chromatography (UPLC) system and RP columns packed with 1.7 μm C_{18} sorbent to separate peptides, proteins, and oligonucleotides. We measured retention factors, peak widths, and peak volumes of biopolymers under isocratic and gradient elution conditions with several goals in mind. First, we estimated the optimal mobile phase linear velocity for various large molecules using the peak capacity theory. Second, we intended to illustrate the retention factor dependence on mobile phase elution strengths, represented by the retention parameter B. Third, we investigated the relationship between the retention parameters B and solute molecular weight. Fourth, a variant of the peak capacity model was developed and applied to the separation of oligonucleotides, which represent a class of biopolymers with specific retention behavior.

2. Experimental

2.1. Materials and reagents

Triethylamine (TEA), 99.5%, acetic acid, and hexafluoroisopropanol (HFIP), 99+, were purchased from Sigma (St. Louis, MO, USA). Trifluoroacetic acid (TFA) was purchased from Pierce (Rockford, IL, USA). A 2–2 M solution of TEA and acetic acid (triethylammonium acetate, TEAA) in water was obtained from Fluka (Seelze, Germany). HPLC grade methanol and acetonitrile were purchased from Fisher Scientific (Fair Lawn, NJ, USA). A Milli-Q system (Millipore, Bedford, MA, USA) was used to prepare deionized water (18 MΩ cm) for HPLC mobile phases. Bradykinin, leucine-enkephaline, renin substrate, bovine insulin, bovine ubiquitin, bovine ribonuclease A, bovine serum albumin, bovine β-lactoglobulin A, and phosphodiesterase I were purchased from Sigma. Oligonucleotides were purchased from Integrated DNA technologies (IDT, Coralville, IA, USA). Oligonucleotides were purified by HPLC and digested by phosphodiesterase I [18] to generate oligonucleotide ladders. Digestion was monitored by HPLC and terminated at the desired moment by boiling the reaction mixture for 5 min.

2.2. HPLC instrumentation, columns, and conditions

HPLC experiments were carried out using an ACQUITY UPLC system with an ACQUITY UPLC photodiode array detector (Waters, Milford, MA, USA). The columns used in this study were 30 mm × 2.1 mm, 50 mm × 2.1 mm and 100 mm × 2.1 mm ACQUITY UPLC OST C_{18} columns packed with 1.7 μm sorbent. The UPLC system was operated with 50 μL or 425 μL mixers as indicated in the discussion and in the figure captions.

The mobile phases used for oligonucleotide separation were A 0.1 M TEAA in water, pH ~7.5; B mobile phase A:acetonitrile, 80:20 (v/v). The TEAA ion-pairing mobile phase (200 mL) was prepared by pipetting 10 mL of 2 M TEAA stock and adding 190 g of water; the pH was not adjusted, the final pH was approximately 7.5. The mobile phase B was prepared by mixing 200 mL of 0.1 M TEAA with 39.3 g (50 mL) of acetonitrile. The alternative ion-pairing system for oligonucleotide separation consisted of (A) 15 mM TEA, 400 mM HFIP in water; (B) mobile phase A:methanol, 50:50 (v/v). Mobile phase A was prepared by mixing 191.3 g of water, 13.44 g of HFIP and 0.416 mL of TEA. The mobile phase B was prepared as A with the addition of 158.2 g of methanol.

The separations of peptides and proteins were performed with the mobile phase A 0.1% TFA in water; (B) 0.1% TFA in acetonitrile (A: 200 g of water, add 0.2 mL of TFA; B: 157.6 g of acetonitrile, add 0.2 mL of TFA). All separations were performed at 60 °C; gradient and other separation conditions are specified in the figure captions.

For flow rates between 0.075 and 0.15 mL/min we used 50 mm × 50 μm PEEKsil tubing (Upchurch Scientific, Oak Harbor, WA, USA), connected after the photodiode array detector to maintain the combined pressure on the UPLC pump above 7 MPa.

3. Results and discussion

3.1. Peak capacity model

Expressions useful for the peak capacity prediction in RPLC for random small molecules and biopolymers like peptides were developed by Neue at al. on the bases of the linear-solvent-strength gradient theory (LSS theory) [15,16,19]. These models can be used for assessing the optimal mobile phase flow rate for a separation under a given set of constraints. We will first discuss the model for peptides, and then examine the case for oligonucleotides.

The simple model for the peak capacity P for peptides as well as for the fast chromatography of small molecules resulted in the

following relationship:

$$P = 1 + \frac{\sqrt{N}}{4}\frac{B\Delta c}{B\Delta c(t_0/t_g) + 1} \quad (1)$$

The peak capacity P depends on the column efficiency N (valid for isocratic separation), the parameter B, which is defined as the slope of the linear dependency of $\ln k$ versus the volume fraction of the organic solvent in the mobile phase φ. The parameter Δc represents the volume fraction of the gradient span delivered over the gradient time t_g; t_0 is retention time of an unretained marker compound. For example, Δc of 0.1 represents the gradient span of 10% organic modifier (typically acetonitrile or methanol).

It should be noted that this model is derived on the underlying assumption that the analyte properties are reasonably uniform over the elution range. In other words, any distribution of analytes and their molecular weight (which affects the plate count and the parameter B) is random over the chromatogram. This is not an unreasonable assumption for small molecules or peptides.

The model can be expanded, if the plate count N is expressed as a function of the operating parameters, which include column length L and particle size d_p, and the typical diffusion coefficient D_m for the analytes of interest [19]:

$$N = \frac{L}{Ad_p + (\beta D_m t_0/L) + C(d_p^2/D_m)(L/t_0)} \quad (2)$$

A, β, and C are the terms of Van Deemter equation; the term β is typically equal to 1, and can therefore be omitted in further expressions.

We have recently utilized the model for the theoretical evaluation of peak capacity in 2D-LC peptide separation [17]. It was found that with appropriately chosen parameters B and D_m, the predictions and experimental data agreed with acceptable accuracy.

The underlying assumption that the chromatographic properties of the separated entities do not change over the course of the gradient is not true for the separation of oligomers, including oligonucleotides. For these compounds, both the diffusion coefficient, which determines the plate count, and the parameter B vary over the course of the gradient. It can be shown [2,7] that the slope of the function $\ln k$ versus the volume fraction of organic modifier in the mobile phase B becomes steeper with increasing molecular weight, and that longer oligomers are more retained in RPLC. It is also known that the diffusion slows down with increasing molecular weight, which affects the plate count of the analytes over the chromatogram.

From a practical point of view, Eq. (1) is suitable for evaluating the peak capacity for peptide mapping applications, since peptides generally elute randomly between 0 and 50% gradient of organic modifier. Any peptide map generated by the digestion of a protein yields peptides with similar range of lengths and retention behavior. The peak capacity can be readily simulated for different columns, gradient times, and flow rates using uniform value of $\Delta c = 0.5$. On the other hand, oligonucleotides tend to elute within a narrow span of gradient, in most cases within 2–4% of organic modifier. Additionally, the separation selectivity (and the resolution) is lower for larger oligomers. This makes it difficult to predict an "average" peak capacity from Eq. (1) for oligonucleotides in general.

One therefore must use a different approach for the assessment of the performance of a gradient separation of oligonucleotides. We will use the resolution for a peak pair for this calculation and then show that the overall sample peak capacity is nothing but the sum (or the integral) of the resolutions in the chromatogram as shown in Eq. (3):

$$P^{**} = \sum_j R_{s,j} \quad (3)$$

The resolution R_s for every neighboring pair of oligonucleotides is defined as

$$R_s = \frac{t_{r,2} - t_{r,1}}{(w_2 + w_1)/2} \quad (4)$$

Denotations t_r and w stand for the retention time and the peak width, indexes 1 and 2 indicate first or second eluting oligonucleotide. Under LSS conditions, the retention time of a peak is a function of its retention factor at the beginning of the chromatogram k_0 and the generalized gradient slope G (defined by Eq. (10)) [19]:

$$t_{r,i} = t_0\left(\frac{1}{G_i}\ln(G_i k_{0,i} + 1) + 1\right) \quad (5)$$

The peak width (also in time units) is derived from the column efficiency N and the retention factor of the peak at the column outlet k_e:

$$w_i = 4\frac{t_0}{\sqrt{N_i}}(k_{e,i} + 1) \quad (6)$$

The retention factor at the point of elution can also be derived from the LSS theory:

$$k_{e,i} = \frac{k_{0,i}}{k_{0,i}G_i + 1} \quad (7)$$

Therefore the peak width becomes

$$w_i = 4\frac{t_0}{\sqrt{N_i}}\left(\frac{(G_i + 1)k_{0,i} + 1}{G_i k_{0,i} + 1}\right) \quad (8)$$

We can now assemble the equation for the resolution in every segment of our gradient:

$$R_s = \frac{1}{2}\frac{(1/G_2)\ln(G_2k_{0,2} + 1) - (1/G_1)\ln(G_1k_{0,1} + 1)}{\left(1/\sqrt{N_2}\right)(((G_2 + 1)k_{0,2} + 1)/(G_2k_{0,2} + 1)) + \left(1/\sqrt{N_1}\right)(((G_1 + 1)k_{0,1} + 1)/(G_1k_{0,1} + 1))} \quad (9)$$

The gradient slope G contains the gradient parameters and the compound-specific slope B of the log/linear relationship

between the retention factor and the solvent composition:

$$G = B\Delta c\frac{t_0}{t_g} = Bg \tag{10}$$

g is the gradient slope that can be influenced by the investigator $g = \Delta c \times t_0/t_g$. The resolution now is

$$R_s = \frac{1}{2g}\frac{(1/B_2)\ln(gB_2k_{0,2}+1) - (1/B_1)\ln(gB_1k_{0,1}+1)}{\left(1/\sqrt{N_2}\right)((((gB_2+1)k_{0,2}+1))/(gB_2k_{0,2}+1)) + \left(1/\sqrt{N_1}\right)(((gB_1+1)k_{0,1}+1)/(gB_1k_{0,1}+1))} \tag{11}$$

As we will see in Section 3 of this paper, both the retention factor at the beginning of the gradient k_0 and the slope B are the functions of the analyte. Furthermore, there is a linear relationship between this retention factor and the slope B.

Eq. (11) can be further simplified for our purposes. If we assume that the plate counts of the closely eluting neighboring peaks as well as their gradient slopes are very similar, the equation changes into

$$R_s = \frac{\sqrt{N}}{4g}\frac{gBk_0+1}{(gB+1)k_0+1} \times \left(\frac{1}{B_2}\ln(gB_2k_{0,2}+1) - \frac{1}{B_1}\ln(gB_1k_{0,1}+1)\right) \tag{12}$$

The N, B and k_0 here are the average values for closely eluting peaks of selected analytes. Next we assume that B_1 and B_2 change much less with retention than k_0 and that the term in k_0 is much larger than 1:

$$R_s = \frac{\sqrt{N}}{4}\frac{gB}{gB+1}\ln\left(\frac{k_{0,2}}{k_{0,1}}\right) \tag{13a}$$

Both the plate count N and the slope B depend on the MW (size) of analytes that we are considering. Since we are interested in the peak spacing for peak pairs, we can express Eq. (13a) as a function of the specific retention range α_0 of the analytes in which we investigate:

$$R_s = \frac{\sqrt{N_{av}}}{4}\frac{gB_{av}}{gB_{av}+1}\ln(\alpha_0) \tag{13b}$$

The dependence of B, k_0, D_m and N on the molecular weight MW of the analyte can be assessed from Eqs. (14a)–(14d). Parameters a, b, e, and f are numerical constants obtained from regression (see later in this paper); d represents the parameters of the Wilke–Chang equation (333.16 K, viscosity = 2.6 cP, solvent molecular volume = 18 mL, association factor = 0.00468) $d = 0.3604 \times 10^{-3}$. For small molecules, the exponent of Eq. (14c) is −0.6. Similar values have been reported in the literature for polymers [20]. The exact value depends on the environment of the oligomer. A formula for the diffusion coefficient of polymers based on the solvent-dependent Mark Houwink constant can be found in textbook [21], page 88. A commonly found value is −0.6 for a good solvent, although Lukacs et al. [22] reported a value of −0.71 for larger DNA (up to 6000 bp) in a salt solution. Considering the general recommendation for a polymer in a good solvent as well as the fact that reversed-phase chromatography is carried out in the presence of some organic modifier and a low salt concentration, we selected the literature value of −0.6 for the dependence of the diffusion coefficient on the molecular weight of the oligosaccharides. It needs to be stressed that one should not confuse this coefficient with the one obtained for compact molecules such as proteins, where a value of −1/3 is found [23,24]:

$$B = a\ln(\mathrm{MW}) - b \tag{14a}$$

$$\ln k_0 = e\ln(\mathrm{MW}) - f \tag{14b}$$

$$D_m = d\,\mathrm{MW}^{-0.6} \tag{14c}$$

$$N = \frac{L}{H} = \frac{L}{A + C(d_p^2/D_m)u} = \frac{L}{A + C(d_p^2/d)\,\mathrm{MW}^{0.6}u} \tag{14d}$$

This gives the following expression for the resolution:

$$R_s = \frac{1}{4}\sqrt{\frac{L}{A + C(d_p^2/d)\,\mathrm{MW}^{0.6}u}}\frac{1}{g(a\ln(\mathrm{MW}) - b) + 1}\ln(\alpha) \tag{15}$$

The peak capacity P^{**} for any part or the entire chromatogram, can now be calculated from Eq. (3). For multiple analytes we simply sum up the resolution of multiple oligonucleotides for every segment j of the chromatogram.

There are several advantages expressing the peak capacity this way. Contrary to Eq. (1), Eq. (3) does not require the knowledge of the gradient span Δc. If a wider than necessary gradient span is used in Eq. (1), the peak capacity is overestimated. Eq. (3) also correctly predicts the decline of the peak capacity for larger oligonucleotides, while Eq. (1) yields an opposite trend (Eq. (1) relies on peak widths, larger oligonucleotides elute as narrower peaks; see further discussion). Finally, the goal of a typical oligonucleotide separation is to resolve the parent compound from several truncated oligomers called $n-1$, $n-2$, etc. [25,26]. In such a case the analysts are interested in the resolution (peak capacity) of adjacent peaks, which can be readily calculated from their molecular weight, provided that their retention parameters B and k_0 are known.

3.2. Estimating optimal mobile phase flow rate for biopolymer separation

For columns packed with conventional chromatographic sorbents (3–5 μm) the separation of biopolymers such as peptides, proteins and oligonucleotides is generally performed with a larger than optimal flow rates. While the <2 μm chromatographic sorbents provide a higher separation efficiency and peak capacity, it remains to be investigated what is the optimal mobile phase velocity for biopolymers separation, and how it depends on the molecular weight of the analytes.

Because the isocratic measurement of the theoretical plate height for the biopolymers using Van Deemter curves is difficult (the minimum occurs at very low flow rate; some proteins do not

elute as well defined peaks under isocratic conditions), we have investigated the application of the peak capacity theory to the same goals, employing a gradient elution mode.

If a constant gradient volume is maintained throughout the experiment (by varying the experimental flow rate and the time of gradient proportionally), the plot of $1/(R_s)^2$ results in a pattern that is related to the theoretical plate height H of the analytes:

$$\frac{1}{R_s^2} = \text{const}\left(A + C\left(\frac{d_p^2}{d}\right)\text{MW}^{0.6}u\right) = \text{const}\,H \tag{16}$$

The resolution is inversely proportional to the peak volume V_r:

$$\frac{1}{R_s^2} = 4\frac{p_v^2}{(V_{r,2} - V_{r,1})^2} \tag{17}$$

p_v is a peak volume. And therefore

$$H = \text{const}\,p_v^2 \tag{18}$$

The following experiment was executed with the goal to evaluate the validity of the relationship proposed in Eq. (18). The plot of squared peak volumes should permit an assessment of the pattern of the theoretical plate height H as a function of the flow rate.

The experiment depicted in Fig. 1A and B was carried out by adjusting the flow rate proportionally to t_g, so the gradient volume in kept constant. Fig. 1A illustrates the plots of the square of the peak volume p_v^2 versus the mobile phase flow rate for 15, 20, and 40mer oligodeoxythymidine peaks; the MWs of the analytes were 4501, 6022, and 12105.9 Da; respectively. The shape of the curves indicated that the optimal flow rate is below the flow rates used in this experiment.

Fig. 1A shows that the measured peak volumes decrease in the oligonucleotide series from 15 to 20 to 40mer. This is the opposite behavior one intuitively expects. However, this behavior does not imply a greater column efficiency for larger oligonucleotides; instead it is related to the fact that the retention factor at the column outlet decreases with the increasing molecular weight of the analyte. The retention factor changes faster from the column inlet to the outlet for the analytes with a larger molecular weight (larger B). This is captured in Eq. (9) in conjunction with Eq. (13a). The reduced retention factor compensates for the increase in peak width caused by the lower plate count for large molecules. As a consequence, the peak volume becomes smaller for larger oligonucleotides.

Similar p_v^2 versus flow rate curves were measured for peptides and small proteins as well (Fig. 1B). The analytes were: Leucine-enkephaline, 555.6 Da; bradykinin, 1060 Da; renin substrate, 1758 Da; bovine insulin, 5733.5 Da, bovine ribonuclease A, 13,685 Da; and bovine β-lactoglobuline A, 18,352 Da. Only selected plots are shown in Fig. 1B. As expected, the curves for smaller peptides such as leucine-enkephaline or bradykinin reach the minimum at a larger mobile phase linear velocity than proteins and oligonucleotides. The plots in Fig. 1 resemble Van Deemter curves. The optimal flow rate values for leucine-enkephaline and bradykinin were in agreement with conventional Van Deemter plots measured at isocratic elution conditions (the position of maximum efficiency closely matched the minimum on the curve p_v^2 versus flow rate; data not shown).

The results suggest that the proposed gradient method can be used to evaluate the optimal chromatographic conditions for the analysis of biopolymers. Nevertheless, the analysis of Fig. 1 reveals several features of the method that deserve additional discussion.

3.3. Experimental evaluation of peak capacity: gradient time

In Section 3.2, we assessed the flow rates that achieve the maximum plate count for the various biopolymers for columns packed with a 1.7 μm C_{18} sorbent. Despite using a small particle size sorbent and elevated temperatures, these flow rates lie below the operational values of analytical LC instruments. Does that mean that even state-of-the-art columns are unsuitable for the separation of biopolymers?

The answer to this question is *certainly not*. Columns packed with a larger sorbent particle size have been successfully applied for the analysis of biopolymers, despite the fact that they are typically operated at flow rates greatly exceeding the flow rate range where the minimum plate height occurs. In addition, the peak capacity theory presented in Eqs. (1), (13a) and (13b) implies

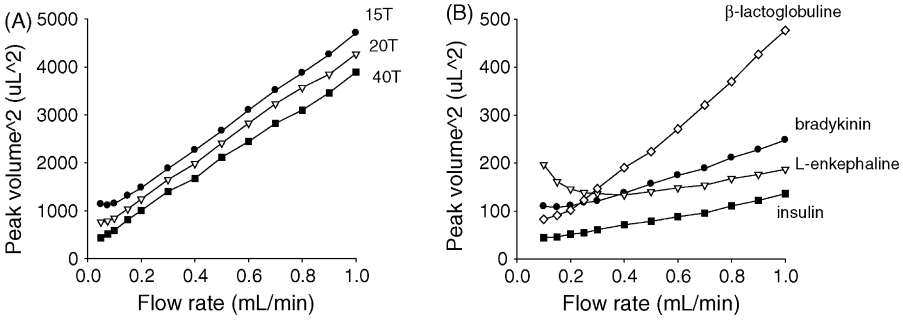

Fig. 1. Pseudo Van Deemter curves measured in gradient RPLC mode represented as a plot of the squared peak volume vs. mobile phase flow rate. RPLC (A) oligonucleotides, (B) peptides and proteins. Conditions: 50 mm × 2.1 mm, 1.7 μm ACQUTIY C_{18} column, 60 °C, 425 μL mixer. Mobile phases for oligonucleotides were: A, 15 mM TEA, 400 mM HFIP; B, 50:50 mobile phase A:methanol. Oligonucleotide gradient: 30–54% B, proportionally varying flow rate and gradient time; gradient slope $\Delta c \times (t_0/t_g)$ was kept constant. Example of gradient: FR = 0.2 mL/min, t_g = 40 min. Mobile phases for protein/peptide separation: A, 0.1% TFA in water; B, 0.1% TFA in 50:50 water:acetonitrile. Protein/peptide gradient: 20–100% B, proportionally varying flow rate and gradient time; gradient slope was kept constant. Example of gradient: FR = 0.2 mL/min, t_g = 30 min.

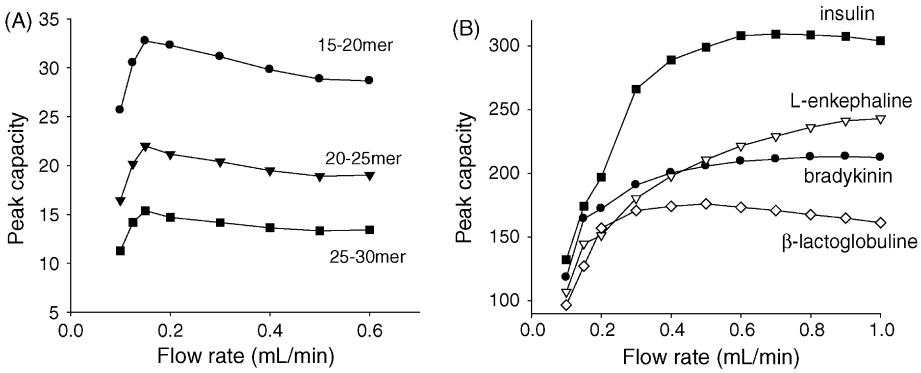

Fig. 2. Estimation of optimal peak capacity for (A) oligonucleotides, and (B) peptides/proteins using gradient RPLC. Column, LC conditions, and mobile phases were the same as in Fig. 1. Oligonucleotide gradient: 30–54% (B) in 26.67 min. Protein/peptide gradient: 20–100% (B) in 10 min. The gradient time was kept constant, the flow rate (hence gradient slope) varied.

that the maximum peak capacity for a given analysis time is achieved at flow rates much faster than those required for the plate count maximum. These equations suggest that not only the column efficiency N, but also the gradient slope g has an important impact on the overall peak capacity. The optimal combination of both for a given analysis time determines the best flow rate.

The peak capacity was experimentally evaluated for selected biopolymers as shown in Fig. 2. The gradient time was fixed, while the flow rate was varied. Please note that this is a different experimental design than in Fig. 1, where the gradient slope was maintained constant by adjusting the gradient time and the flow rate proportionally. The approach used in Fig. 2 expands the gradient volume and flattens the gradient slope. At the same time, the plate count deteriorates with increasing flow rate.

The peak capacity in Fig. 2A was calculated directly from chromatograms for 15–20mer, 20–25mer, and 25–30mer oligonucleotide pairs (Fig. 2A). The retention times t_1 and t_2 represent the first and second eluting peaks; similarly the peak widths w_1 and w_2 (measured at 4σ, 13.4% of the peak height):

$$P = 1 + \frac{t_2 - t_1}{(w_2 + w_1)/2} \tag{19}$$

The peak capacity in Fig. 2B was calculated from Eq. (20); the gradient time t_g was 10 min. The use of different methods for the calculation of the peak capacity for peptides/proteins was forced by the fact that their selectivity (peak spacing in the chromatogram) varied with the effective gradient slope:

$$P = 1 + \frac{t_g}{w} \tag{20}$$

Fig. 2A and B illustrates that the overall peak capacity of the column has indeed shifted towards faster than optimal flow rates observed in Fig. 1A and B. Even though the column efficiency decreases at a higher mobile phase velocity, the peak capacity loss is offset by the positive contribution of the shallower gradient (Eqs. (13a) and (13b)), which – on average – expands the distance between the peaks. The shallower gradient is caused by the increased flow rate at the fixed gradient run time. As one would expect, the maximum peak capacity for peptides was obtained at a faster flow rates than for proteins (Fig. 2B). Interestingly, the absolute peak capacity values were not in direct correlation with the analyte molecular weight. Peak compression [27] as well as other contributions to peak width play a distinct role in protein/peptides chromatographic behavior. For example insulin is formed by two chains cross-linked via disulphide bonds, which forms a highly compact molecule. The penetration into the pores of a packing becomes hindered for larger molecules, which in turn impedes performance.

Oligonucleotides appear to have substantially different patterns compared to peptides. The optimal peak capacity for oligonucleotides is observed at about the same flow rate. This is due to the difference in the influence of the molecular weight of the analyte on the two parts of Eq. (15). As the flow rate (or better the linear velocity u) increases, the value of the square root decreases, while the gradient steepness parameter g becomes smaller, which increases the value of the gradient contribution to the resolution. Consequently, the position of the peak capacity maximum in Fig. 2A remains rather insensitive to the molecular weight of the oligonucleotide.

The comparison of Figs. 1 and 2 reveals information that should not be overlooked. The optimal peak capacity does not coincide with the smallest peak volume. Therefore, operating at flow rates providing the best peak capacity will result in a moderate (in this case two to three times) loss of detection sensitivity. This is acceptable for many applications, where the resolution is the most desirable goal of analysis.

3.4. Measurement of B parameters

The proposed peak capacity models in Eqs. (1), (13a) and (13b) rely on the knowledge or reliable estimates of the retention factors such as B. Therefore, we investigated these parameters experimentally.

The constant B for oligonucleotides was adapted from Ref. [7] (Fig. 2 in [7]). Apparently, the B values increase with the size of the molecule. This correlates to observation made earlier by Snyder [1,2,9]. We used the log–log fit [9] that seems to be universal for oligonucleotides, peptides and proteins (Fig. 3 and Table 1).

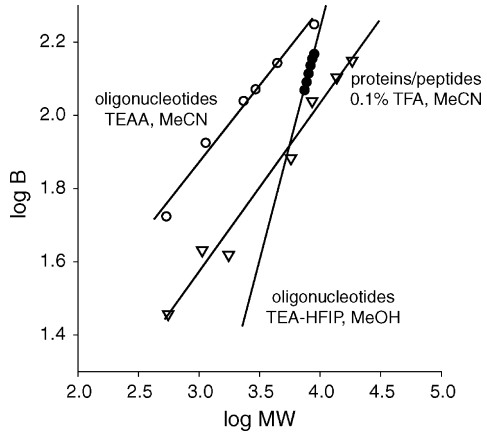

Fig. 3. Factor B was measured at isocratic conditions for oligonucleotides and peptides/proteins, and plotted as function log B vs. logarithm of analyte molecular weight. B values are listed in Table 1. The column and the conditions for proteins/peptides and oligonucleotides (TEA-HFIP, MeOH experiment) were the same as in Fig. 1. Data for TEAA, MeCN system for oligonucleotides were adopted from Gilar et al. [7]. Circles: experimental data for oligonucleotides; triangles: experimental data for proteins/peptides.

Eqs. (21) and (22) describe log B as a function of log MW or the log length of oligonucleotides (n; number of nucleotides in sequence). The linear fit for log B versus log (MW) and log (n) gave a correlation coefficient of $R^2 = 0.983$ and 0.979, respectively, for the TEAA ion-pairing system:

$$\log\ B = 0.42\ \log(\mathrm{MW}) + 0.60 \quad (21)$$

$$\log\ B = 0.43\ \log(n) + 1.63 \quad (22)$$

In Ref. [7] we used Eq. (21) for the prediction of the B factor for peptides and proteins for data taken from the literature [8]. Although this equation predicts acceptable B values for tryptic peptides in the MW range 1000–2000 Da, we believe this equation is incorrect. In light of recent experiments we propose Eq. (23). The fit of log B versus log MW for 7 peptides/proteins was linear with a correlation coefficient of $R^2 = 0.978$.

$$\log\ B = 0.46\ \log(\mathrm{MW}) + 0.18 \quad (23)$$

Eqs. (21)–(23) are valid for the given chromatographic conditions used for the separation. Understandably, the intercept and the slope of the equations may change with temperature, the type of stationary phase, organic modifier (acetonitrile, methanol, isopropanol, ...), and the nature of the ion-pairing system. We investigated the impact of the last two parameters on oligonucleotide separations.

Fig. 3 compares data for two different ion-pairing systems and organic modifiers. TEAA was used in combination with acetonitrile, while the TEA-HFIP ion-pairing system employed methanol as the organic modifier [28–32]. Eqs. (24) and (25) significantly differ from Eqs. (21)–(23) (see also Fig. 3 and Table 1). Correlation for both equations was $R^2 = 0.995$:

$$\log\ B = 1.27\ \log(\mathrm{MW}) - 2.86 \quad (24)$$

$$\log\ B = 1.28\ \log(n) + 0.28 \quad (25)$$

The predicted B values for the TEA-HFIP/methanol system (using the same column and temperature) are lower than the ones derived from Eqs. (21)–(23), at least for the shorter oligonucleotides. However, because the change of the slope B with the volume fraction of the organic modifier is steeper, the two lines cross at the position corresponding to a 36mer oligonucleotide. The slope of the log B − log (MW) relationship is much steeper for TEA-HFIP ion-pairing system. The reasons for that are not clear at this point, but the observation correlates with the fact that TEA-HFIP seems to be a more efficient system for the separation of oligonucleotides than TEAA [7,18].

Table 1
Experimentally measured B and ln k_0 parameters for oligonucleotides (deoxythymidines), peptides and proteins

Solute	MW (Da)	B[a]	ln k_0[a]	Conditions
2T	546.4	52.6	3.7	0.1 M TEAA, acetonitrile, 60 °C
4T	1154.8	83.6	7.3	0.1 M TEAA, acetonitrile, 60 °C
8T	2371.6	108.7	10.7	0.1 M TEAA, acetonitrile, 60 °C
10T	2980	117.0	12.1	0.1 M TEAA, acetonitrile, 60 °C
15T	4501	138.0	15.5	0.1 M TEAA, acetonitrile, 60 °C
30T	9063.9	176.0	20.8	0.1 M TEAA, acetonitrile, 60 °C
25T	7542.9	116.3	26.78	15 mM TEA, 400 mM HFIP, methanol, 60 °C
26T	7848.1	122.5	28.31	15 mM TEA, 400 mM HFIP, methanol, 60 °C
27T	8151.3	129.2	29.96	15 mM TEA, 400 mM HFIP, methanol, 60 °C
28T	8455.5	135.9	31.62	15 mM TEA, 400 mM HFIP, methanol, 60 °C
29T	8759.7	141.7	33.06	15 mM TEA, 400 mM HFIP, methanol, 60 °C
30T	9063.9	146.1	34.16	15 mM TEA, 400 mM HFIP, methanol, 60 °C
Leucine-enkephaline	555.6	28.65	7.087	0.1% TFA, acetonitrile, 60 °C
Bradykinin	1060	42.75	9.196	0.1% TFA, acetonitrile, 60 °C
Renin substrate	1758	41.57	9.551	0.1% TFA, acetonitrile, 60 °C
Insulin	5733.5	76.37	24.81	0.1% TFA, acetonitrile, 60 °C
Ubiquitin	8565	109.5	33.39	0.1% TFA, acetonitrile, 60 °C
Ribonuclease A	13,685	127	32.73	0.1% TFA, acetonitrile, 60 °C
β-Lactoglobulin A	18,352	141.2	59.33	0.1% TFA, acetonitrile, 60 °C

[a] From function $\ln k = \ln k_0 - B \times \Phi$ [15,16]. Parameters S and log k_0 defined by Snyder [1,9] can be calculated as $S = B/2.303$ and $\log k_0 = \ln k_0/2.303$.

Table 2
Comparison of predicted and experimental peak capacity and retention times

	15–20mer	20–25mer	25–30mer	30–40mer	40–50mer	50–60mer	Sum 15–60mer[b]
Experimental data							
t_1 (min)[a]	12.89	16.49	19.07	21.11	24.18	26.35	12.89
t_2 (min)[a]	16.49	19.07	21.11	24.18	26.35	28.05	28.05
$t_2 - t_1$ (min)	3.60	2.58	2.04	3.07	2.17	1.70	15.16
P_{exp}[c]	19.8	13.1	10.1	15.2	11.0	8.5[d]	77.8
Predicted data							
t_1 (min)[e]	11.20	13.90	15.79	17.20	19.20	20.58	11.20
t_2 (min)[e]	13.9	15.79	17.20	19.20	20.58	21.60	21.60
$t_2 - t_1$ (min)	2.70	1.89	1.41	2.00	1.38	1.02	10.40
P_{calc}	16.2	11.6	8.8	12.7	8.9	6.8	64.9

P and t_r were calculated for six selected segments, and for the entire 15–60mer range.
[a] Retention time t_g values were obtained from Fig. 5A.
[b] The sum 15–60mer peak capacity is calculated by summing the contributions from the selected segments.
[c] Calculated from Eq. (19), see also captions in Fig. 5.
[d] Only 50mer peak width was used for the calculation.
[e] Predicted from the model, Eq. (11).

3.5. Peak capacity prediction for oligonucleotides and comparison with experimental results

The peak capacity model developed here is a useful tool for choosing the optimal experimental parameters for biopolymer separation. We have investigated the impact of column length L, gradient time t_g, and sorbent particle size d_p *in silico* and compared the finding to experimental data. In this case study, we utilized oligonucleotides as a sample. Unlike proteins and peptides, oligonucleotides have a predictable chromatographic behavior, unaffected by secondary structure.

The peak capacity in the following examples was predicted from Eq. (11). Because the separation selectivity (hence the peak capacity) is greater for shorter oligonucleotides, we divided 15–60mer oligonucleotides into six subgroups and summed the results into an overall peak capacity. The comparison of experimental and predicted P values are shown in Table 2. The data are discussed in greater depth in Section 3.7. The simplified Eq. (12) provided comparable results, while the very simple Eqs. (13a) and (13b) predicted slightly higher peak capacity values. However, the trends observed with the simple Eqs. (13a) and (13b) were the same as for the more complex equations.

Fig. 4A illustrates the impact of the particle size of the chromatographic packing on the average peak capacity of 15–60mer oligonucleotides. The peak capacity was calculated for 50 mm × 2.1 mm columns packed with 1.7 μm, 3 μm, and

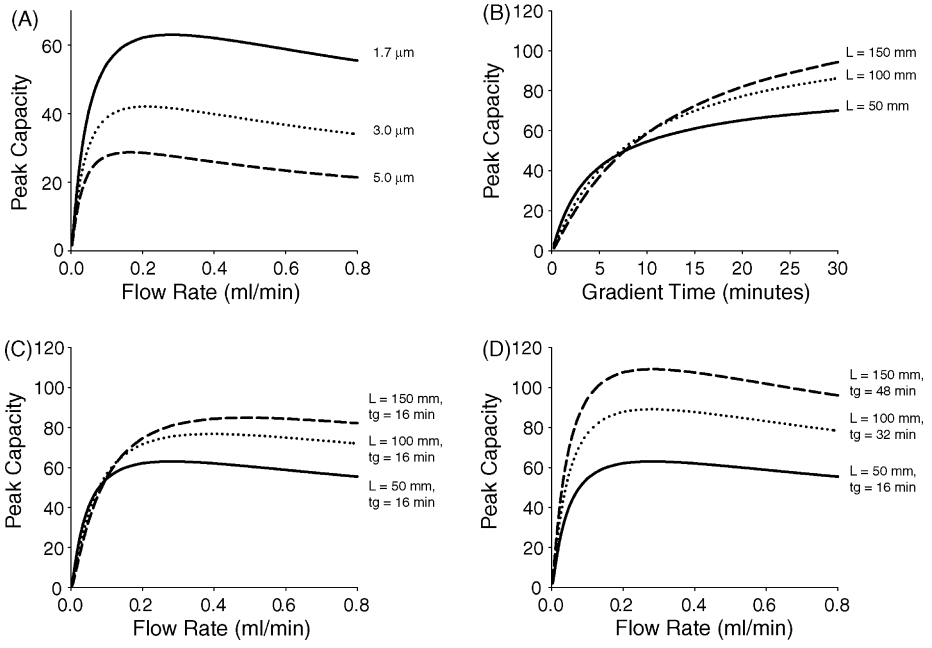

Fig. 4. Averaged peak capacity for 15–60mer oligonucleotides calculated for 2.1 mm column I.D., 1.7 μm sorbent, $\Delta c = 0.024$, initial gradient strength 9% MeCN, $t_g = 16$ min, and Van Deemter C term = 0.05, unless indicated otherwise. (A) P as a function of sorbent particle size; (B) P as a function of gradient run time at fixed flow rate 0.2 mL/min; (C) P as a function of flow rate at a fixed gradient run time of 16 min; (D) P as a function of flow rate, the gradient time was scaled to column length (16, 32, and 48 min, for 50, 100, and 150 mm columns, respectively).

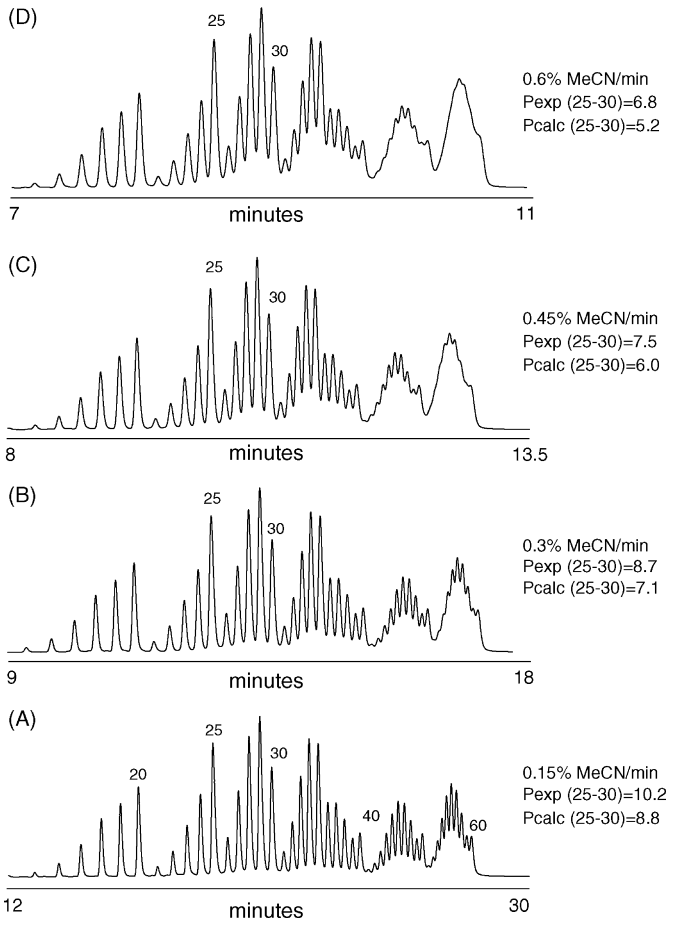

Fig. 5. 15–60mer oligodeoxythymidine separation using different gradient slopes. Conditions: 50 mm × 2.1 mm, 1.7 μm ACQUTIY OST C_{18} column, 60 °C, 425 μL mixer, flow rate 0.2 mL/min. Mobile phases—A, 100 mM TEAA; B, 80:20 mobile phase A:acetonitrile. The gradient was 40–70% B (8–14% MeCN). Gradient slopes are indicated in the figure. The programmed LC gradient was 40, 20, 13.3, and 10 min for (A), (B), (C) and (D) chromatograms, respectively. The experimental peak capacity P_{exp} was estimated from the chromatograms for 25–30mer using Eq. (19); peak widths measured at 50% peak height were multiplied by 1.7 to obtain peak width at 4σ. Peak capacity P_{calc} 25–30mer was calculated from Eq. (11), for 2.1 mm × 50 mm, 1.7 μm column, gradient slope 0.6, 0.45, 0.3, and 0.15% MeCN/min; Van Deemter C term = 0.05.

5 μm sorbents, respectively. The gradient time was 16 min, and the overall gradient span Δc was 0.024 (slope 0.15% MeCN/min, similar to conditions in Fig. 5A). Not surprisingly, the best peak capacity was obtained for the column packed with the smallest particle size sorbent. The experimental data illustrating this trend were published elsewhere [7].

Fig. 4B compares the peak capacity for three different column lengths. In this prediction, the gradient span Δc was kept constant, while the gradient time was varied. The flow rate was 0.2 mL/min. Fig. 4B illustrates a scenario intuitively understood by practitioners: extending the gradient time (using shallow gradients) is expected to improve the prospects for a successful separation. The gradient span of 2.4% acetonitrile per analysis may seem narrow; however, the separation of oligonucleotides indeed occurs within such a limited range of mobile phase elution strength (at least for the TEAA ion-pairing system with acetonitrile as the eluent).

Fig. 4C (as well as Fig. 4A and D for 1.7 μm, 50 mm long column) shows the peak capacity as a function of flow rate for a fixed 16 min long gradient ($\Delta c = 0.024$). The predicted P curve for a 50 mm long column resembles the shape of experimental curves shown in Fig. 2A, although the predicted peak capacity maximum is shifted to a slightly larger flow rate. The absolute P values are not directly comparable, since the data in Fig. 2A were acquired with a TEA-HFIP ion paring system, rather than with TEAA, which was used for the prediction in Fig. 4.

Intriguingly, the gains in resolution for longer columns shown in both Fig. 4B and C are not as pronounced as one may expect. The gains in P are certainly not proportional to column length. Furthermore, at short gradient times and slow flow rates the peak capacity predicted for longer columns is in fact less than for shorter columns. This seemingly counterintuitive prediction correlates well with the practical experience of many chromatographers. It has been reported that for gradient RPLC separation of proteins the resolution is practically independent of the column length [1,2]. This behavior is explained by the impact of the gradient slope g defined as $\Delta c \times (t_0/t_g)$. Longer columns have both greater efficiency N and volume (t_0). At constant gradient run time t_g the gradient slope is proportionally shallower for shorter columns. In other words, the peak capacity of longer columns is reduced by proportionally sharper gradient, which tends to reduce or eliminate the positive impact of higher column efficiency.

The full benefits of longer columns in a gradient separation are realized when changing the gradient duration in proportion with the column volume (length). This scenario is depicted in Fig. 4D. All conditions are the same as in Fig. 4C, except for extending the gradient time to 32 min and 48 min for 100 and 150 mm columns, respectively. In this scenario, the gradient slope $\Delta c \times (t_0/t_g)$ is the same for all three columns. In such a case, the peak capacity calculated from Eq. (15) depends solely on the square root of the column efficiency N, which is proportional to the column length L. The expected gain for the scale up from a 50 mm column to a 100 mm column is $(100/50)\hat{0}.5 = 1.41$. Similarly, the peak capacity gain for the scale up from a 50 to 150 mm column is 1.73. The 41% or 73% increase in peak capacity is achieved at the expense of 2× or 3× longer analysis times.

The impact of the gradient slope predicted in Fig. 4B was evaluated experimentally using a 15–60mer ladder of oligonucleotides. The results are shown in Fig. 5. Oligonucleotides were separated on a 50 mm × 2.1 mm, 1.7 μm Acquity OST C_{18} column; the gradient time (slope) was varied. Calculated peak capacities P_{calc} for 25–30mer peaks can be compared to P_{exp} obtained from chromatograms. The details are given in the caption to Fig. 5. The experimental P_{exp} values clearly followed the trends predicted from the model (see Figs. 5–7). The correlation coefficient between predicted and experimental results was 0.998.

Fig. 5 suggests that shortening the analysis time by using sharp gradients is detrimental for resolution. Fig. 6 illustrates an alternative approach. Oligonucleotide separations were performed at different flow rates, reducing the gradient time proportionally to the flow rate. The gradient slope g was therefore constant throughout this experiment.

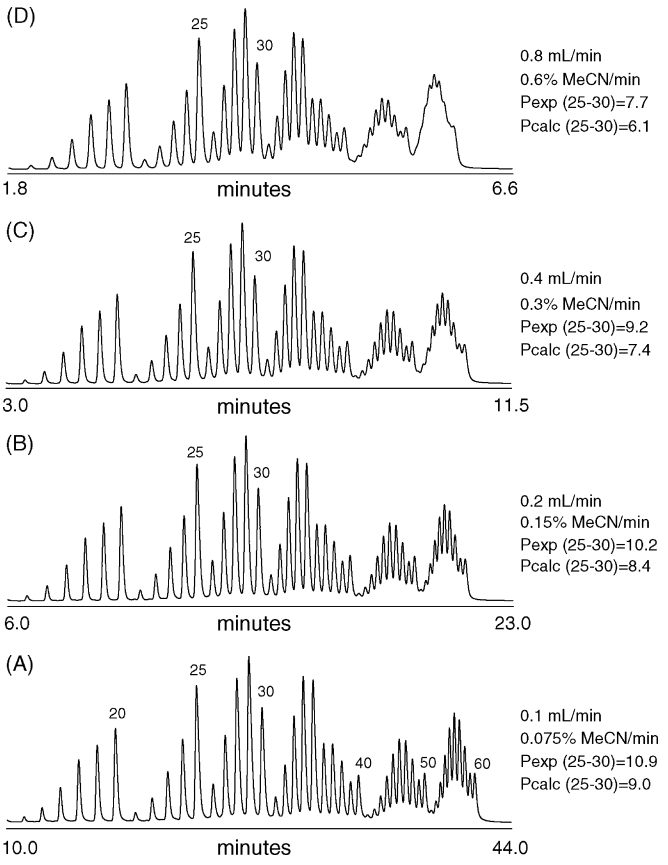

Fig. 6. 15–60mer oligodeoxythymidine separation at various flow rates. Gradient slope $\Delta c \times (t_0/t_g)$ was kept constant by adjusting the gradient and the flow rate proportionally. Conditions: 50 mm × 2.1 mm, 1.7 μm ACQUITY OST C_{18} column, 60 °C, 425 μL mixer. Mobile phases—A, 100 mM TEAA; B, 80:20 mobile phase A:acetonitrile. The gradient was 45–64.5% B (9–12.9% MeCN). Flow rates are indicated in figure. The programmed LC gradient was 52, 26, 13, and 6.5 min for (A), (B), (C) and (D) chromatograms, respectively. The P_{exp} values for 25–30mer were obtained as described in Fig. 5. Peak capacity P_{calc} 25–30mer was calculated as in Fig. 5, gradient slope 0.6, 0.3, 0.15, and 0.075% MeCN/min; Van Deemter C term = 0.05.

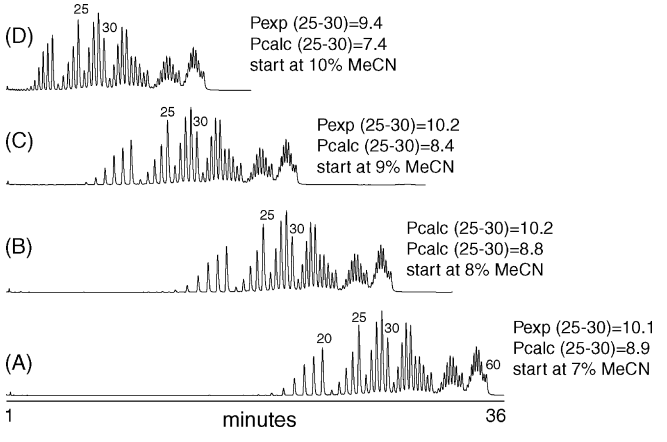

Fig. 7. Strategy for reducing analysis time without compromising peak capacity. 15–60mer oligonucleotide ladder was separated with constant gradient slope (0.15% MeCN/min), initial gradient mobile phase strength was changed as indicated in figure. Conditions: 50 mm × 2.1 mm, 1.7 μm ACQUITY OST C_{18} column, 60 °C, 425 μL mixer. Mobile phases—A, 100 mM TEAA; B, 80:20 mobile phase A:acetonitrile. The flow rate was 0.2 mL/min. The P_{calc} and P_{exp} values for 25–30mer were obtained as described in Fig. 5.

When increasing the flow rate, the column efficiency decreases as shown earlier in Fig. 1A. However, Fig. 6 indicates that the loss in peak capacity is less detrimental compared to using sharp gradients (compare Figs. 5 and 6). Therefore, it is more practical to maintain relatively shallow gradient and reduce the analysis time by increasing the flow rate and gradient time proportionally (maintaining constant gradient slope).

When applying Eq. (1), one can easily arrive to misleading results by choosing a larger t_g and Δc than justified by the retention behavior of the chosen analytes. For example, while peptides typically elute in a composition range between 0 and 50% acetonitrile ($\Delta c = 0.5$), the effective elution window for oligonucleotides is only several percent of organic modifier. This is illustrated in Fig. 7. The gradient slope $\Delta c \times (t_0/t_g)$ was kept constant for all experiments, while the initial mobile phase elution strength varied. Because oligonucleotides eluted within the same mobile phase elution strength span, the remaining space in the chromatogram is not useful for the separation. Predicting the peak capacity from unrealistically large t_g and Δc, the separation capabilities of the column will be overestimated. Therefore, we prefer to calculate the P for experimental data from the newly developed models (Eqs. (11), (12), (13a), (13b) and (15)). P can be realistically predicted from Eq. (1) when using rational t_g and Δc values, if possible estimated from experiment.

The peak capacity model predicts the scenario experimentally confirmed in Fig. 7. The figure illustrates a suitable approach for reducing the analysis time without undesirable sacrifice in peak capacity. The initial gradient strength may be optimized, while keeping the gradient slope constant, and by doing so, the resolution of oligonucleotides is virtually unaffected. A small loss in the peak capacity P is observed only in Fig. 7D. Due to the gradient delay of the LC system, the early eluting peaks eluted under isocratic conditions (a 425 μL mixer was used).

3.6. Fast analysis of oligonucleotides

Equipped with the knowledge of the optimal mobile phase flow rate and column length/gradient time scale-up relationships, one can develop fast methods for the biopolymer separation. Fig. 8A and B illustrate the method development for a 15–35mer ladder. Flow rate, gradient slope, and initial mobile phase elution strength were selected appropriately, guided by data in Figs. 2A, 5 and 7. Fig. 8B shows the first result, using the mobile phase flow rate of 0.2 mL/min, a gradient slope 0.25% MeOH/min and initial mobile phase strength of 19% MeOH. The separation achieved for the 15–35mer ladder was more than adequate, therefore it was possible to further shorten the analysis time. This was achieved by doubling the flow rate and reducing the gradient time to half of the initial value (therefore, the gradient slope was unchanged). The initial gradient elution strength was also adjusted to further reduce the analysis time. The separation at 0.4 mL/min was accomplished in less than 4 min (Fig. 8A). As expected, the peak capacity somewhat decreased, but this is acceptable as long as the resolution meets the requirements defined by the analyst.

Fig. 8C shows the separation of a 30–60mer oligonucleotide ladder at the flow rate of 0.2 mL/min. In this case, the peaks

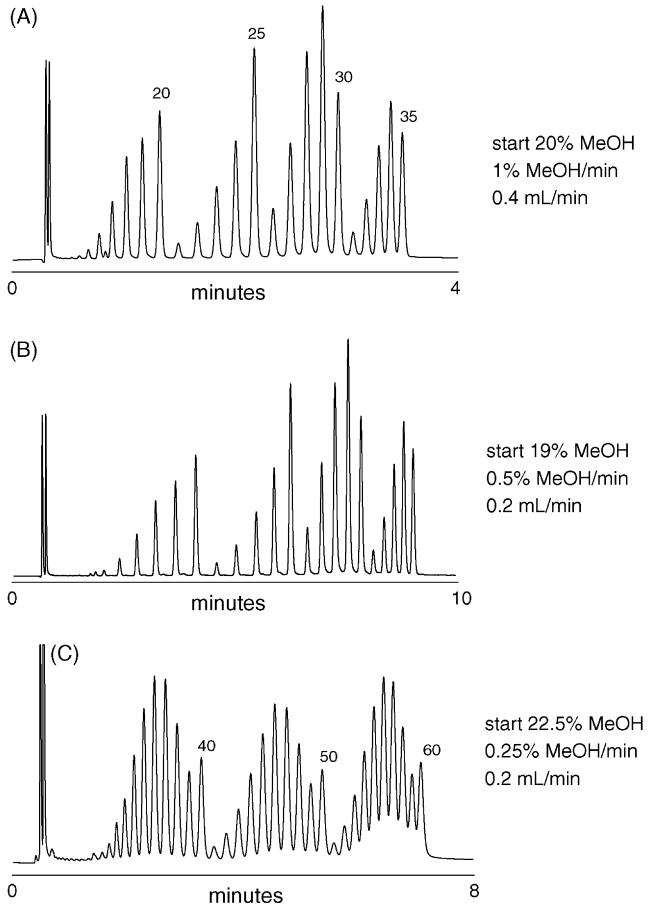

Fig. 8. Fast oligonucleotide separations in UPLC. (A) and (B) 15–35mer oligodeoxythymidine ladder. (C) 30–60mer oligodeoxythymidine ladder. Conditions: 50 mm × 2.1 mm, 1.7 μm ACQUTIY C_{18} column, 60 °C, 50 μL mixer. Mobile phases—A, 15 mM TEA, 400 mM HFIP; B, 50:50 mobile phase A:methanol. Initial gradient strength, gradient slope and flow rate are indicated next to chromatograms.

are not completely resolved. Therefore, it is not advisable to further increase the flow rate and decrease the gradient time. The achieved separation is a compromise between speed and resolution.

One may consider using longer columns to improve the separation. It is important to understand that this is possible at the expense of the separation time. Fig. 4C shows the best strategy for scaling up column length. As discussed above, the peak capacities achieved with 50–100 and 150 mm columns are in ratios of 1:1.41:1.73, when the gradient time is scaled to the column volume (Fig. 4C). One may wonder whether it is possible to avoid extending the analysis time for longer columns by adjusting the flow rate as suggested in Fig. 6 (keeping the gradient slope constant). The scenario can be as follows: 50, 100, and 150 mm columns can be operated at 0.2, 0.4, and 0.6 mL/min, respectively, while maintaining a constant gradient time. One apparent problem represents the high operational pressure of long columns at high flow rates. More importantly, this strategy fails to provide a higher peak capacity. Calculated P ratios for 50–100 and 150 mm long columns are 1:1.07:1.10.

For fast separations, we prefer to use a 50 μL mixer over the 425 μL one. The smaller mixer adds less gradient delay and reduces the overall analysis time. Precautions have to be made for the separation of biopolymers to avoid imperfect mobile phase mixing, especially for shallow gradients. The premixed mobile phases used here resolved any issues with artifacts caused by insufficient mobile phase mixing.

3.7. Benefits and limitations of the peak capacity model

As discussed above, the peak capacity is a useful tool for prediction of separation quality at various chromatographic conditions. Some limitations of the model defined by Eq. (1) were noted [17], such as the necessity to define an appropriate gradient span and B constant. The model in Eq. (1) has been successfully used to estimate the peak capacity for RPLC separation of peptides, because the Δc is typically 0.5, and the elution pattern of analytes in the chromatogram is random [17].

Notwithstanding its utility, this simple model is not applicable for oligonucleotides or oligomers with a specific elution behavior. Most importantly, their retention increases, and the separation selectivity decreases with the oligomer chain length. Also, they tend to elute in an extremely narrow Δc range. Hence, a different mathematical model must be used.

The model presented in Eq. (3) addresses this problem by calculating the peak capacity as a sum of resolution values for neighboring peaks (segments) across the chromatogram. The definition of resolution (Eqs. (11), (12), (13a), (13b), (14a), (14b), (14c), (14d), (15)) encompasses the selectivity term $\ln(k_{0,2}/k_{0,1})$, as well as the steepness of B of the retention pattern. Consequently, the peak capacity prediction accurately captures the experimentally observed trends of diminishing separation prospects for longer oligonucleotides, as illustrated in Table 2. The modeling can be performed without prior knowledge of Δc. Since the model contains predetermined k_0 and B values, it can be also used to predict oligonucleotide retention times (Table 2).

Although the new proposed model is suitable for the oligonucleotides peak capacity prediction, some limitations should be mentioned: (i) The k_0 and B values have to be obtained experimentally for the desirable chromatographic system. The model reliability depends on the accuracy of the k_0 and B estimates. The values need to be adjusted when different ion-pairing system, organic modifier, or the stationary phases are considered. (ii) The oligonucleotide diffusion coefficients D_m were calculated from a theoretical model. Inaccurate D_m estimates will introduce some bias into P prediction. (iii) The peak capacity model was developed for homooligonucleotides. Separation of heterooligonucleotides partially depends on their sequence [7,18]; the spacing between peaks is more irregular and resolution varies. However, when using experimentally measured k_0 values for heterooligonucleotides of interest, Eq. (11) should predict the correct P values. (iv) Peak capacity cannot be reliable calculated for oligonucleotides with strong secondary structure. In such case the k_0 and B (hence the separation selectivity) will change unpredictably. (v) The model does not include the impact of the gradient delay of the LC instrument. It is assumed that the gradient starts immediately after the sample is injected onto the column. Table 2 lists the predicted P for six segments of the chromatogram, 15–20, 20–25, 25–30, 30–40, 40–50, and

50–60mer; and the total 15–60mer peak capacity is defined as the sum of all segments. The predicted P values were compared to the experimental peak capacity obtained from Fig. 5A; separation conditions were the same as used for the theoretical prediction. Apparently, the model underestimates P; this was also observed in Figs. 5–7 for the 25–30mer span.

Experimental and predicted retention times are listed in Table 2. A systematic shift towards shorter times can be seen for the experimental data. This could be partially explained by the gradient delay of the LC system (in this case 0.5 mL), which is not considered in the prediction.

In spite of the limitations, the developed model is a suitable tool for method development and *in silico* experiments, allowing the users to semi-quantitatively evaluate the impact of chromatographic conditions on the oligonucleotide separation performance.

4. Conclusions

The peak capacity theory was applied to solve several practical problems in RP separation of biopolymers. First, the peak capacity model was used to evaluate an optimal mobile phase linear velocity for the separation of peptides, proteins and oligonucleotides. The results obtained in gradient elution mode were in agreement with the values derived from Van Deemter equation. The proposed gradient method is simple and applicable to a range of analytes from moderate to large molecular weight.

The retention data of biopolymers were measured at isocratic conditions and used to define the slope of the $\ln k$ function versus the fraction of organic modifier Φ in the mobile phase (factor B). The relationship of B to the molecular weight of the analyte was defined for oligonucleotides, peptides and proteins. A rational prediction of B allows for more accurate peak capacity modeling in RPLC.

The peak capacity theory was validated by experimentally measured trends using UPLC columns packed with a 1.7 μm sorbent. The flow rate for the maximum peak capacity for 50 mm × 2.1 mm column was found to be 0.15 mL/min for oligonucleotides and 0.4–0.8 mL for proteins and peptides (column temperature 60 °C). The impact of gradient slope, initial gradient strength and flow rate on the separation was evaluated. The data indicate that the proposed peak capacity model realistically describes the separation behavior of oligonucleotides in RPLC.

Acknowledgement

The authors thank Joomi Ahn for her helpful suggestions to the manuscript.

References

[1] L.R. Snyder, M.A. Stadalius, in: C. Horvath (Ed.), High-performance Liquid Chromatography. Advances and Perspectives, Academic Press, Inc., Orlando, 1986, p. 195.
[2] L.R. Snyder, M.A. Stadalius, M.A. Quarry, Anal. Chem. 55 (1983) A1412.
[3] X.M. Lu, K. Benedek, B.L. Karger, J. Chromatogr. 359 (1986) 19.
[4] J.W. Eschelbach, J.W. Jorgenson, Anal. Chem. 78 (2006) 1697.
[5] S. Yamamoto, M. Nakamura, C. Tarmann, A. Jungbauer, J. Chromatogr. A. 1144 (2007) 155.
[6] B.G. Belenkii, A.M. Podkladenko, O.I. Kurenbin, V.G. Mal'tsev, D.G. Nasledov, S.A. Trushin, J. Chromatogr. 645 (1993) 1.
[7] M. Gilar, K.J. Fountain, Y. Budman, U.D. Neue, K.R. Yardley, P.D. Rainville, R.J. Russell II, J.C. Gebler, J. Chromatogr. A 958 (2002) 167.
[8] S. Terabe, H. Nishi, T. Ando, J. Chromatogr. 212 (1981) 295.
[9] L.R. Snyder, J.W. Dolan, High Performance Gradient Elution. The Practical Application of the Linear-Solvent-Strength Model, Wiley, New York, 2007.
[10] R. Skudas, B.A. Grimes, E. Machtejevas, V. Kudirkaite, O. Kornysova, T.P. Hennessy, D. Lubda, K.K. Unger, J. Chromatogr. A 1144 (2007) 72.
[11] H.F. Yin, M.H. Kleemis, J.A. Lux, G. Chomburg, J. Microcol. Sep. 3 (1991) 331.
[12] C.F. Poole, The Essence of Chromatography, Elsevier, Amsterdam, 2003.
[13] P.J. Oefner, J. Chromatogr. B 739 (2000) 345.
[14] J.J. Kirkland, F.A. Truszkowski, C.H. Dilks Jr., G.S. Engel, J. Chromatogr. A 890 (2000) 3.
[15] U.D. Neue, J.L. Carmody, Y.F. Cheng, Z. Lu, C.H. Phoebe, T.E. Wheat, in: P. Brown, E. Grushka (Eds.), Advances in Chromatography, vol. 41, Marcel Dekker, New York, 2001, p. 93.
[16] U.D. Neue, J.R. Mazzeo, J. Sep. Sci. 24 (2001) 921.
[17] M. Gilar, A.E. Daly, M. Kele, U.D. Neue, J.C. Gebler, J. Chromatogr. A 1061 (2004) 183.
[18] M. Gilar, Anal. Biochem. 298 (2001) 196.
[19] U.D. Neue, J. Chromatogr. A 1079 (2005) 153.
[20] R. Groh, I. Halasz, Anal. Chem. 53 (1981) 1325.
[21] W.W. Yau, J.J. Kirkland, D.D. Bly, Modern Size Exclusion Liquid Chromatography, Willey–Interscience, New York, 1979.
[22] G.L. Lukacs, P. Haggie, O. Seksek, D. Lechardeur, N. Freedman, A.S. Verkman, J. Biol. Chem. 275 (2000) 1625.
[23] M.T. Tyn, T.W. Gusek, Biotech. Bioeng. (1990) 327.
[24] R.R. Walters, J.F. Graham, R.M. Moore, D.J. Anderson, Anal. Biochem. 140 (1984) 190.
[25] M. Gilar, A. Belenky, Y. Budman, D.L. Smisek, A.S. Cohen, J. Chromatogr. B 714 (1998) 13.
[26] M. Gilar, A. Belenky, D.L. Smisek, A. Bourque, A.S. Cohen, Nucleic Acids Res. 25 (1997) 3615.
[27] U.D. Neue, D.H. Marchand, L.R. Snyder, J. Chromatogr. A 1111 (2006) 32.
[28] A. Apffel, J.A. Chakel, S. Fischer, K. Lichtenwalter, W.S. Hancock, Anal. Chem. 69 (1997) 1320.
[29] K.J. Fountain, M. Gilar, J.C. Gebler, Rapid Commun. Mass Spectrom. 17 (2003) 646.
[30] H.J. Gaus, S.R. Owens, M. Winniman, S. Cooper, L.L. Cummins, Anal. Chem. 69 (1997) 313.
[31] M. Gilar, E.S.P. Bouvier, J. Chromatogr. A 890 (2000) 167.
[32] C.G. Huber, H. Oberacher, Mass Spectrom. Rev. 20 (2001) 310.