Journal of Chromatography, 485 (1990) 365–382
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 21 946

REVERSED-PHASE CHROMATOGRAPHIC METHOD DEVELOPMENT FOR PEPTIDE SEPARATIONS USING THE COMPUTER SIMULATION PROGRAM ProDigest-LC

COLIN T. MANT*, T. W. LORNE BURKE, NIAN E. ZHOU, J. M. ROBERT PARKER and ROBERT S. HODGES

Department of Biochemistry and Medical Research Council of Canada Group in Protein Structure and Function, University of Alberta, Edmonton, Alberta T6G 2H7 (Canada)

SUMMARY

A computer program, ProDigest-LC, has been developed that assists scientists in devising methods of size-exclusion, cation-exchange and reversed-phase high-performance liquid chromatography for the analytical separation and purification of biologically active peptides and peptide fragments from enzymatic and chemical digests of proteins. ProDigest-LC accurately predicts the retention behaviour of peptides of known composition, containing 2–50 amino acid residues, and simulates the elution profiles in all three modes of chromatography. In addition, ProDigest-LC is a user-friendly program, designed as a teaching aid for both students and researchers in selecting the correct conditions for chromatography, that is, the mode of chromatography, column selection and mobile-phase selection, and has the ability to examine the effects of gradient-rate, flow-rate and sample size on the separation.

The simulation capabilities of ProDigest-LC as they apply to the reversed-phase chromatography of peptides were examined. The development of the reversed-phase simulation features of the program is described, stressing the importance of peptide standards in the development, testing and practical use of ProDigest-LC. The ease of use of the program is clearly demonstrated by presenting a step-by-step procedure to produce several of the simulations illustrated in the paper.

The predictive accuracy of the program was rigorously tested by its application to retention time prediction, at different gradient-rates and flow-rates, for a sample mixture containing peptides exhibiting a wide range of size (11–50 residues), charge (+1 to +8 net charge), hydrophobicity and conformation (random coil to considerable α-helical structure). The excellent accuracy of these peptide retention time predictions complemented the successful simulation (in terms of peptide retention times, peptide resolution, peak heights and peak widths) of the effects of gradient-rate and flow-rate on the elution profile of a mixture of closely related peptide analogues.

INTRODUCTION

A computer software program, ProDigest-LC, has been developed that assists

0021-9673/89/$03.50

scientists in devising methods of high-performance liquid chromatography (HPLC) for the analytical separation and purification of peptides. ProDigest-LC was developed in response both to the increasing importance of the efficient isolation of peptides for an ever-widening range of research disciplines in recent years, and to the recognition that HPLC has shown immense versatility in the separation and purification of peptides from many sources.

The wide variations in complexity and amounts of peptide mixtures are a reflection of the wide variety of sources from which they are derived:

(a) Peptides obtained from biological tissues are often found only in very small amounts and may require extensive purification. Thus, being able to predict the location of a biologically active peptide from various tissue sources in chromatograms obtained by different HPLC modes would be extremely beneficial.

(b) Separation of peptides from a chemical and/or proteolytic digest of a protein for subsequent peptide characterization is vital in structure–function studies of proteins. The complexity of the resulting peptide mixture will depend on the particular digestion agent in addition to the properties (size, amino acid composition and sequence, etc.) of the protein of interest.

(c) During the biosynthesis of proteins for therapeutic purposes, impurities very similar to the desired protein will be present. Separation systems are required that can detect small changes in the polypeptide chain; thus, peptide mapping, following protein digestion by chemical or proteolytic agents, is one means of verifying the structure of a genetically engineered protein.

(d) The wide use of automated solid-phase peptide synthesis in recent years has also necessitated the efficient isolation of peptides from various impurities, usually closely related to the peptide of interest (deletion, terminated or chemically modified peptides) and perhaps missing only one amino acid residue.

Although a desired peptide separation may be obtained by trial and error, this may take many attempts, with subsequent loss of time and final peptide yield. This could be a particularly serious problem if only limited amounts of sample are available. Therefore, any methodology that can aid the researcher in selecting a purification protocol without using precious sample or requiring an extensive method development time would be invaluable. ProDigest-LC has been designed to simulate peptide elution profiles for the three major modes of HPLC employed for peptide separations: size-exclusion HPLC (SEC), cation-exchange HPLC (CEC) and reversed-phase HPLC (RPC). The experiments simulated on the computer eliminate the time-consuming trial-and-error approach to peptide purification. In addition, ProDigest-LC is also a teaching aid for chromatographers, designed to help the student or researcher to select the correct conditions for chromatography (HPLC mode, column and mobile phase) and allowing him or her the option of examining the effect of varying flow-rate, gradient-rate and sample size on the separation.

The value of ProDigest-LC to the chromatographer can, of course, only be assessed by rigorous testing of its simulation capabilities. In this study we examined the capability of ProDigest-LC in simulating reversed-phase elution profiles of peptide mixtures. RPC is by far the most widely used mode of HPLC at present. The ability of this technique to separate peptides with closely related structures has made it an extremely powerful analytical and preparative tool. Whatever the source of a particular peptide sample, the resolving capability of RPC makes it the obvious

choice for the initial HPLC run to gauge the complexity of the peptide mixture and help design the best approach for its resolution.

This paper reviews the development of the RPC simulation capabilities of ProDigest-LC and examines the utility of the program by comparing simulated and observed reversed-phase elution profiles of peptide mixtures containing peptides varying widely in size, hydrophobicity and α-helical content. The validity of the program was rigorously tested with regard to its ability to predict the effect of varying gradient-rate and flow-rate on peptide retention behaviour in RPC.

EXPERIMENTAL

Materials

Water was purified by passage through a Hewlett-Packard (Avondale, PA, U.S.A.) HP 661A water purifier. Acetonitrile was obtained from Fisher Scientific (Fairlawn, NJ, U.S.A.). HPLC-grade trifluoroacetic acid (TFA) was obtained from Pierce (Rockford, IL, U.S.A.).

The peptides described (Table I) were synthesized either on a Beckman (Beckman Instruments, Berkeley, CA, U.S.A.) Model 990 peptide synthesizer or an Applied Biosystems (Foster City, CA, U.S.A.) Model 430A, using the general procedure for solid-phase synthesis described by Hodges and co-workers[1,2].

One set of five reversed-phase peptide standards (denoted 10X–50X in Table I)

TABLE I

PEPTIDE SEQUENCES

Peptide	*Peptide sequence*[a]
Peptide polymers[b]:	
10X–50X	Ac–(Gly–Leu–Gly–Ala–Lys–Gly–Ala–Gly–Val–Gly)$_n$–amide, where n = 1, 2, 3, 4 and 5 (10X, 20X, 30X, 40X and 50X, respectively)
5G–50G	Ac–(Gly–Lys–Gly–Leu–Gly)$_n$–amide, where n = 1, 2, 4, 6, 8, 10 (5G, 10G, 20G, 30G, 40G and 50G, respectively)
5A–40A	Ac–(Leu–Gly–Leu–Lys–Ala)$_n$–amide, where n = 1, 2, 4, 6, 8 (5A, 10A, 20A, 30A and 40A, respectively)
5L–30L	Ac–(Leu–Gly–Leu–Lys–Leu)$_n$–amide, where n = 1, 2, 4, 6 (5L, 10L, 20L and 30L, respectively)
7T–35T	Ac–Lys–Cys–Ala–Glu–Gly–Glu–Leu–(Lys–Leu–Glu–Ala–Gly–Glu–Leu)$_n$–amide, where n = 0, 1, 2, 3, 4 (7T, 14T, 21T, 28T and 35T, respectively)
Other peptides:	
I1	Ac–Arg–Gly–Gly–Gly–Gly–Ile–Gly–Ile–Gly–Lys–amide
I2	Ac–Arg–Gly–Gly–Gly–Gly–Ile–Gly–Leu–Gly–Lys–amide
S2	Ac–Arg–Gly–Gly–Gly–Gly–Leu–Gly–Leu–Gly–Lys–amide
S3	Ac–Arg–Gly–Ala–Gly–Gly–Leu–Gly–Leu–Gly–Lys–amide
S4	Ac–Arg–Gly–Val–Gly–Gly–Leu–Gly–Leu–Gly–Lys–amide
C1	Ac–Gly–Gly–Gly–Leu–Gly–Gly–Ala–Gly–Gly–Leu–Lys–amide
C2	Ac–Lys–Tyr–Gly–Leu–Gly–Gly–Ala–Gly–Gly–Leu–Lys–amide

[a] Ac = N$^{\alpha}$-acetyl; amide = C$^{\alpha}$-amide.

[b] For the purposes of this study, each peptide in the mixtures of peptide polymers is referred to by a number and letter which denote, respectively, the number of residues it contains and to which polymer series it belongs. Thus, 5G refers to the five-residue G series peptide, 28T to the 28-residue T series peptide, etc.

was obtained from Synthetic Peptides (Department of Biochemistry, University of Alberta, Edmonton, Canada).

Apparatus

The HPLC instrument consisted of a Hewlett-Packard HP 1090 liquid chromatograph coupled to a HP 1040A detection system, HP 9000 Series 300 computer, HP 9133 disc drive, HP 2225A Thinkjet printer and HP 7440A plotter.

Peptide mixtures were separated on a SynChropak RP-P C_{18} column, 250 × 4.6 mm I.D., 6.5-μm particle size, 300-Å pore size (SynChrom, Linden, IN, U.S.A.) and an Aquapore RP300 C_8 column, 220 × 4.6 mm I.D., 7-μm particle size, 300-Å pore size (Brownlee Labs., Santa Clara, CA, U.S.A.).

Computer program

ProDigest-LC is available from Synthetic Peptides. A series of peptide standards are supplied with ProDigest-LC for each mode of HPLC simulated by the program (SEC, CEC or RPC).

The minimum requirements for operating ProDigest-LC is an IBM-AT or compatible computer with 256K memory, equipped with two floppy disk drives and a monitor with graphics capability. The use of a math coprocessor is strongly recommended, as calculations that take *ca.* 2 min on the IBM-AT can be accomplished in *ca.* 15 s with the math coprocessor.

Features of the program

General applications. Fig. 1, screen A, shows the Main Menu screen of ProDigest-LC and illustrates the various options open to the researcher, which are as follows.

(a) MAIN A section of screen A: this section of the program is designed so that the researcher can enter peptide sequence(s) for separation by a particular mode of chromatography (SEC, CEC or RPC). The amount of the peptide(s) entered can be varied; in addition, after the mode of chromatography has been selected, the flow-rate, gradient-rate or sample volume (in SEC) desired is entered, making the program completely versatile. Fig. 1 demonstrates program manipulation for simulation of RPC of a peptide mixture.

(b) MAIN B section of screen A: in the protein digest menu, the operator has the option of entering a protein sequence and performing an enzymatic or chemical digestion of a particular protein sequence (*e.g.*, cyanogen bromide cleavage, tryptic or other proteolytic cleavage). The computer performs the digestion and the operator is given the option of the mode of HPLC by which the digest is to be separated (SEC, CEC or RPC). This section also contains a peptide mapping option, whereby the operator has the ability to add the particular digest of a mutant protein to the same digest of the native protein sequence. Following CEC or RPC, peptides found in the mutant that are different from those in the native sequence are marked on the elution profile.

(c) SUPPLEMENTAL section of screen A (Standards): ProDigest-LC is an interactive program, in that researchers can run peptide standards on their particular columns (SEC, CEC or RPC) and enter the peak width at half-height, peak height, retention time and amount injected (Fig. 1, screen D) to adjust the simulation to their particular column.

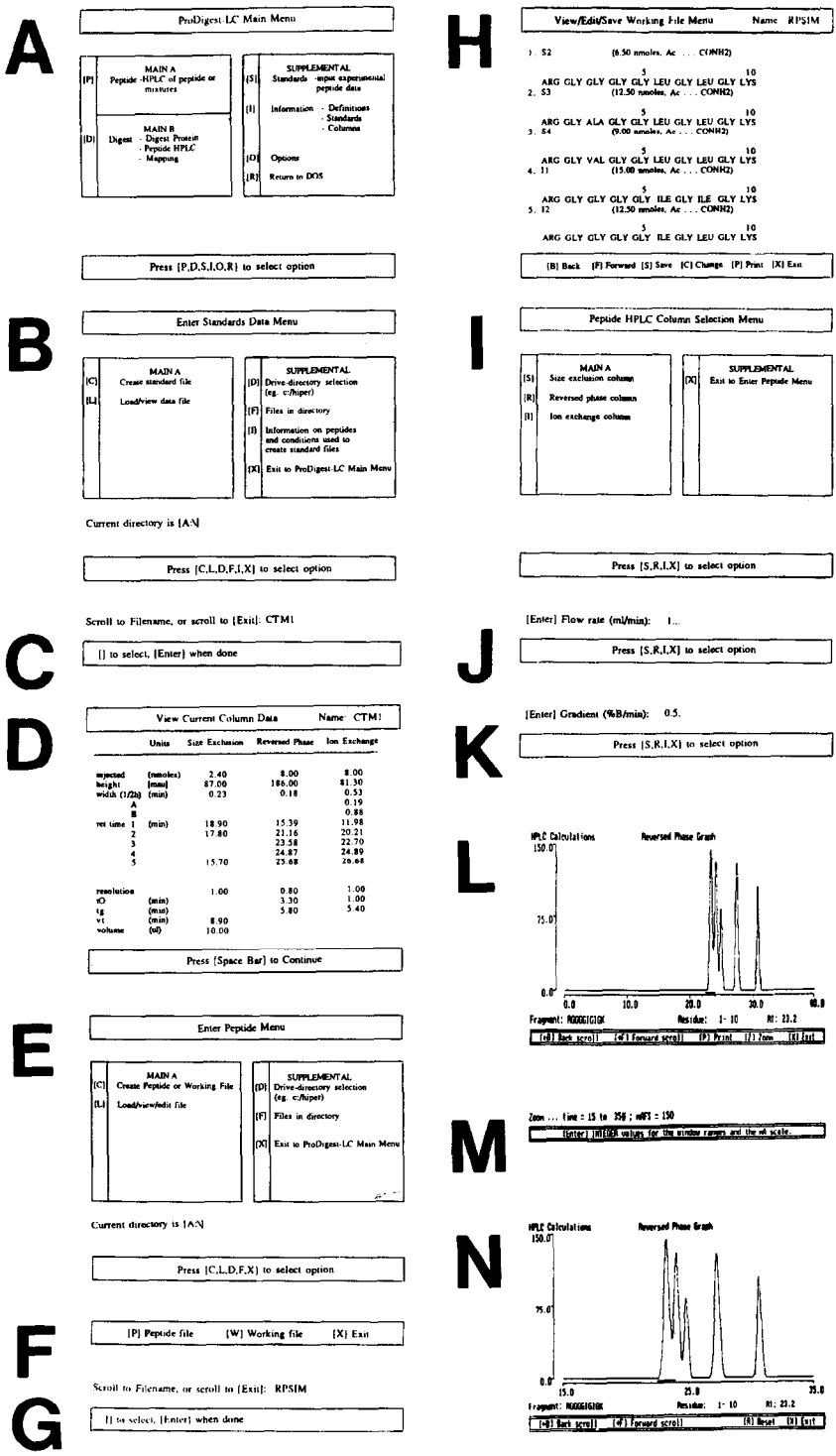

Fig. 1. Example of application of ProDigest-LC to the simulation of a reversed-phase elution profile of a mixture of five peptide standards. Details are described in the text under *Features of the program.*

(d) SUPPLEMENTAL section of screen A (Information): this section demonstrates the utility of peptide standards for monitoring the resolving power of SEC, CEC and RPC columns, in addition to describing peptide standards that are available for these modes of HPLC. Relevant references are listed, so that the researcher can quickly access more detailed information, if required. In addition, this section provides definitions of HPLC parameters used in the program.

Program manipulation. As the present study is concerned only with demonstrating the reversed-phase simulation capabilities of ProDigest-LC, we have chosen to illustrate an example of the clarity of the program presentation as it applies to RPC of a mixture of standard peptides (Fig. 1).

Each menu is self-explanatory in ProDigest-LC, providing simple instructions and one-letter keying to access any particular section of the program. For example, Fig. 1 shows the various screens in a stepwise fashion (A–N) required to perform a series of simulations. The first step is loading a standards file (either the default file or standards file created from data obtained by running ProDigest-LC standards on the column; screens A–C). Having selected the standards file, it is viewed on the screen (screen D). One now loads a peptide file (a single peptide sequence) or a working file (containing a peptide mixture of interest) as shown in screen H. Then one selects the mode of chromatography, flow-rate and gradient-rate desired (*e.g.*, RPC, 1 ml/min, 0.5 B/min; screens I–K, respectively) and the simulation is displayed as shown in screen L. The zoom option can be used to enlarge a particular section of the chromatogram, changing the abscissa (elution time in minutes) and the ordinate [milliabsorbance (mAU)] to any desired value. In this case, a time scale of 15–35 min has been chosen (screen M) and the resulting chromatogram is shown in screen N.

Further details of the various program menus and examples of the simulation capabilities of ProDigest-LC have been reported previously by Hodges *et al.*[3].

RESULTS AND DISCUSSION

Importance of peptide standards

The use of peptide standards has played a key role in the evolution of ProDigest-LC as a viable analytical tool and chromatography teaching aid. Their importance is highlighted in four major areas, as follows.

(a) Reversed-phase silica-based packings may contain surface silanols which act as weak acids and are ionized above pH 3.5–4.0[4]. These negatively charged silanols may interact with the basic residues of peptides and have an adverse effect on resolution, characteristically producing long retention times and peak broadening. Predictable reversed-phase retention behaviour of peptides requires that the mechanism by which they interact with the reversed-phase sorbent is based solely on peptide hydrophobicity. Any non-ideal, ionic interactions must be identified and suppressed by manipulation of the mobile phase. Peptide standards are commercially available for identifying non-specific interactions and are supplied with ProDigest-LC[5,6].

(b) Peptide standards were required during the development of the simulation capabilities of ProDigest-LC in terms of both prediction of peptide retention time and prediction of the effects of varying chromatographic parameters (sample load, gradient-rate, flow-rate) on peptide retention time, resolution, peak height and peak width.

(c) Standards permit calibration of a researcher's HPLC column and instrumentation. After chromatographing a set of standards (supplied with ProDigest-LC) on a particular column, the amount injected, peak heights, peak widths, retention times and other parameters necessary for the program to adjust the predicted elution profiles to the researcher's particular column can be entered. The data generated by this single standard run calibrate the researcher's column to allow for column-to-column differences in peptide retention behaviour resulting from column ageing, different column dimensions (length and diameter), differences in ligand density and *n*-alkyl chain length of the bonded phase, and instrumentation variations which affect the gradient delay time, etc. This standard run also allows ProDigest-LC to use retention coefficients derived from a different column to that used by the researcher.

(d) Model peptides have proved vital during assessment of the simulation capabilities of ProDigest-LC. The most logical initial approach to testing the accuracy of the program is to chromatograph, under varying conditions, well defined mixtures of peptides differing in sequence and conformation from those used to calibrate the column [see point (c) above] and comparing observed and simulated results.

Development of RPC simulation capabilities of ProDigest-LC

Prediction of peptide retention time. For ProDigest-LC to be of widespread practical use, it was necessary to base the development of the reversed-phase simulation capabilities of the program on chromatographic conditions favoured by as many researchers as possible in the peptide and protein field. Although excellent resolution of peptide mixtures may be obtained at both acidic and neutral pH, most researchers carry out RPC at pH <3.0, using volatile mobile phases and linear AB gradients, where A = 0.1% aqueous TFA and B = 0.1% TFA in acetonitrile[7,8]. Apart from the suppression of silanol ionization, low pH conditions are more favourable for silica-based column stability than pH values above neutrality, where the silica matrix is prone to dissolution.

Development of the ability of ProDigest-LC to predict accurately peptide retention times in RPC required a reliable means of assigning a value to the hydrophobicity of a peptide. Specifically, a means of expressing peptide hydrophobicity in terms of HPLC-derived parameters was required. Several research groups have determined sets of amino acid residue hydrophobicity coefficients for predicting peptide retention times in RPC, on the assumption that the chromatographic behaviour of a peptide is mainly or solely dependent on amino acid composition, and this assumption holds well enough for small peptides (up to *ca.* fifteen residues)[9–18]. Retention values have generally been obtained by computer-calculated regression analyses of the retention times of a wide range of peptides of varied composition[9–15]. These methods were not particularly successful for two reasons: first, the sampling of peptides was not large enough to ensure a high frequency of occurrence of all amino acid residues; second, any polypeptide chain-length effect on retention behaviour would be averaged into values obtained for the retention coefficients, resulting in substantial errors[18,19].

The approach of Guo *et al.*[16] of using model synthetic peptides overcame the problems associated with the computer-calculated regression analysis approach. They measured the contributions of individual amino acid residues to the retention time of a synthetic model peptide at a fixed chain length. The eight-residue octapeptide sequence, Ac–Gly–X–X–$(Leu)_3$–$(Lys)_2$–amide, was substituted at position X by

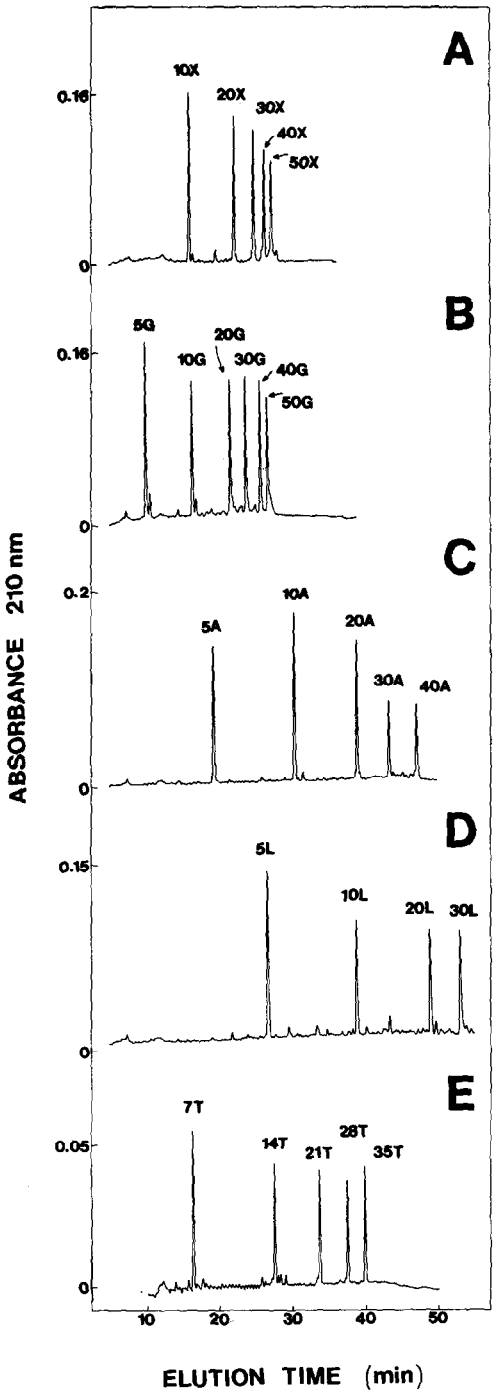

Fig. 2. RPC of mixtures of synthetic peptide polymers. Column: SynChropak RP-P C_{18} (250 × 4.6 mm I.D.). Conditions: linear AB gradient (1% B/min), where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0); flow-rate, 1 ml/min; temperature, 26°C. Panels A–E show the elution profiles of the X, G, A, L and T series of peptide polymers, respectively (see Table I for peptide sequences).

all twenty naturally occurring amino acids. The coefficients reported by Guo *et al.*[16], the most precise set of retention coefficients currently available, formed the basis of peptide retention prediction in RPC by ProDigest-LC.

Several researchers have noted that peptides larger than 15–20 residues tended to be eluted more rapidly than predicted from hydrophobic considerations alone[11–13,18–23]. Lau *et al.*[24] reported a linear relationship between log MW and peptide retention time in RPC for a series of five peptide polymers with 8–36 residues. Hodges and co-workers[3,8] demonstrated a similar exponential relationship for a series of five peptide polymers of 10–50 residues (10X–50X in Table I). The effect on peptide retention of increasing peptide length is clearly illustrated in the RPC profiles (Fig. 2) of five series of synthetic peptide polymers (X, G, A, L and T series; Table I)

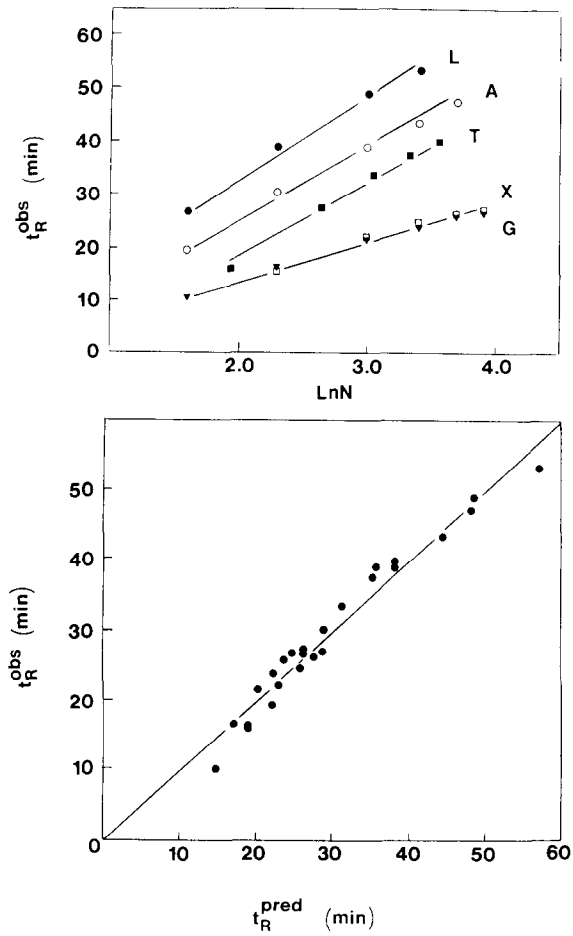

Fig. 3. Top: observed peptide retention time, t_R^{obs}, *versus* ln$N$ (where N is the number of residues in a peptide). Bottom: correlation of predicted and observed peptide retention times in RPC. The predicted retention times (t_R^{pred}) were calculated as described in the text. Results shown are for five series of peptide polymers (X, G, A, L and T series; see Table I for peptide sequences) on a SynChropak RP-P C_{18} column (250 × 4.6 mm I.D.). Conditions are given in Fig. 2.

on a C_{18} column at pH 2.0 (linear AB gradient at 1% B/min and 1 ml/min, where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile). The effect of increasing peptide length on the retention times of these peptides decreased progressively with each addition of a five-residue (G, A and L series), seven-residue (T series) or ten-residue (X series) repeating unit. The exponential nature of the relationship between peptide retention time and peptide chain length of the five series of peptide polymers is illustrated in Fig. 3 (top). Plotting the observed peptide retention time *versus* the logarithm of the number of residues (N) resulted in straight lines with different slopes, depending on the hydrophobicity of a particular peptide polymer series.

The major breakthrough in our ability to predict retention times for peptides with 2–50 residues has resulted from recognizing the intimate relationship between peptide hydrophobicity and chain length and their combined effect on peptide retention behaviour[3,19]. Mant *et al.*[19] demonstrated that plotting the difference between predicted [τ, equivalent to the sum of the retention coefficients of Guo *et al.*[16] (ΣR_c) for the amino acid residues in the peptide, plus the time correction (t_s) for an internal peptide standard] and observed (t_R^{obs}) peptide retention times of series of peptide polymers *versus* the product of peptide hydrophobicity and the logarithm of the number of residues ($\Sigma R_c \ln N$) resulted in a straight-line plot. This relationship held true for peptides of widely differing hydrophobicity and chain length, and was consistent for RPC columns of varying dimensions and hydrophobic functionalities (*n*-alkyl chain length) and ligand density. A similar plot, with a slight modification, was made for the peptide polymers shown in Fig. 2. Instead of $\tau - t_R^{obs}$, $\Sigma R_c - t_R^{obs}$ was plotted against $\Sigma R_c \ln N$. The retention behaviour of the five mixtures of peptide polymers was then predicted using the slope and intercept from this plot. Fig. 3 (bottom) shows the good correlation between the predicted and observed retention times of the polymer series, once peptide chain length has been taken into account. The success of this predictive approach is particularly impressive considering the wide range of peptide chain length and hydrophobicity covered by the peptide polymers and forms the basis for simulation of reversed-phase peptide elution profiles by ProDigest-LC. It should be noted that, at present, ProDigest-LC cannot handle modified amino acids (glycosylated, phosphorylated, etc.). However, the finding of a peptide with a large deviation of predicted from observed retention behaviour can be used as an indication of potential residue modifications.

Prediction of effect of varying run conditions on peptide elution profiles. A knowledge of the effect of varying run conditions on peptide elution profiles simulated under one particular set of conditions was vital for the development of ProDigest-LC as a flexible and practical aid to the researcher. Manipulation of parameters such as flow-rate and, especially, gradient-rate are common approaches to optimizing peptide separations.

A mixture of four synthetic undecapeptide standards was subjected to linear AB gradient elution (eluent A = 0.1% aqueous TFA and eluent B = 0.1% TFA in acetonitrile, pH 2.0) at gradient-rates of 0.5, 1.0 and 2.0% B/min and flow-rates of 0.5, 1.0 and 2.0 ml/min. The total sample loads ranged from 2 to 40 nmol of each peptide. Absorbance was measured at 210 and 280 nm (to detect tyrosine absorbance in two of the peptides)[3]. Data from these chromatographic experiments were used to derive empirical equations predicting the effects of experimental parameters (sample

size, flow-rate and gradient-rate) on peptide retention times, peak heights, peak widths and resolution. Simulated reversed-phase elution profiles generated by the program are based on the mobile phase (see above) used to derive the equations.

Simulation of reversed-phase peptide elution profiles by ProDigest-LC

Calibration of reversed-phase column. On accessing the Standards data menu (Fig. 1, screen B), the operator can create a Standards file or load/view/edit Standard data files in the program. The purpose of the Standards data file is to provide the researcher with the option of either using the default file in the program or entering data derived from standards chromatographed on his or her columns with his or her instrumentation. For example, after chromatographing a set of standards on a particular column, the amount injected, peak heights, peak widths, retention times and other parameters necessary for the program to adjust the predicted elution profiles to the researcher's particular column can be entered.

The method of correcting for different columns and instrumentation is based on the retention behaviour of a range of peptides of different chain lengths and hydrophobicities. This approach was deemed more accurate than the use of a single peptide standard which was utilized in our previous studies[3,19]. It is impractical to have to run several series of peptide polymers (Figs. 2 and 3) each time a column is calibrated. However, this is unnecessary, as the excellent correlation obtained from such plots for peptides differing widely in chain length and hydrophobicity[19] suggested that only a single series of peptide polymer standards was necessary for the program to generate the equation for different reversed-phase columns. The standards chosen for this calibration role were the five peptides of the X series of peptide polymers (Table I). A set of these standards is supplied by Synthetic Peptides with each ProDigest-LC program.

Fig. 1, screen D, shows the reversed-phase Standards file (CTM1) (middle row of figures) created for part of this study. The parameters shown were based on the observed reversed-phase elution profile of the five peptide polymer standards on a C_{18} column using a linear AB gradient (1% B/min and 1 ml/min), where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile. The figures showing amount of peptide injected (8.00 nmol) and peak height and peak width obtained at this sample load (186 mAU and 0.18 min, respectively) refer to the ten-residue peptide standard (10X) in the polymer series. The observed retention times of 10X, 20X, 30X, 40X and 50X (Table I) were 15.39, 21.16, 23.58, 24.87 and 25.68 min, respectively. The required resolution was set at 0.80; t_0 and t_g refer to the column dead time (time for unretained compounds to be eluted from the column) and the gradient delay time (time for the mobile phase to travel from the proportioning valve to the detector, via pump, injection loop and column), respectively.

The Standards file shown in Fig. 1, screen D, was used for the simulations shown in Fig. 6.

Prediction of peptide retention times at varying gradient-rates and flow-rates. The retention time prediction capabilities of ProDigest-LC were subjected to a stringent test by application to RPC of the model peptide mixture shown in Fig. 4. This sample mixture contained peptides varying significantly in size (11–50 residues), charge (+1 to +8 net charge) and hydrophobicity. In addition, the peptides varied considerably in their degree of secondary structure (α-helix). Fig. 5 illustrates the circular di-

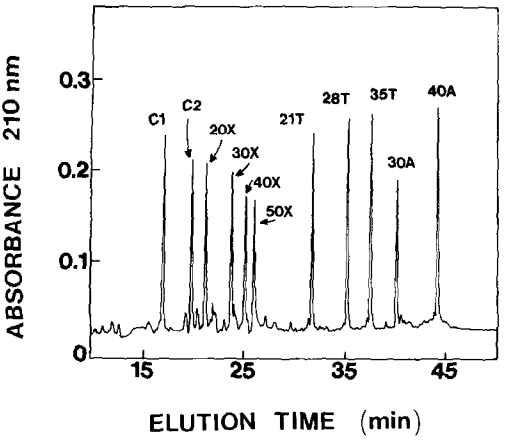

Fig. 4. RPC of a mixture of synthetic peptides. Column: Aquapore RP300 C_8 (220 × 4.6 mm I.D.). Conditions as in Fig. 2. The sequences of the peptides are shown in Table I. This peptide mixture was subjected to variations in both gradient-rate and flow-rate; subsequent observed *versus* predicted (by ProDigest-LC) retention times are given in Table II.

chroism (CD) spectra of several of the peptides in the test mixture. The spectra were measured at pH 2.0 in the presence of trifluoroethanol (TFE) (50% TFE in 0.1% aqueous TFA), a solvent that induces helicity in a single-chain potentially α-helical polypeptide[25]. These conditions were designed to mimic the tendency of the hydrophobic stationary phase in RPC to induce α-helicity in potentially helical molecules[8]. Although only peptides 30L and 40A of the L and A series of peptides, respectively, are shown in Fig 5A, other peptides in these series also showed greater or lesser

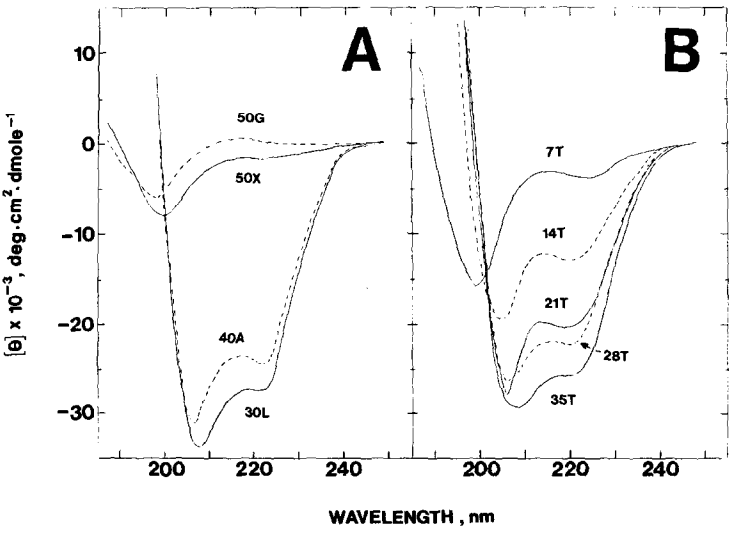

Fig. 5. Circular dichroism spectra of synthetic peptides at pH 2.0 [0.1% aqueous TFA containing 50% (v/v) trifluoroethanol]. The sequences of peptides 50G, 50X, 40A and 30L (A) and 7T–35T (B) are shown in Table I.

α-helical character. Thus, the test mixture shown in Fig. 4 contains peptides that exhibit secondary structure ranging from random coil (*e.g.*, 50X, 50G) to considerable α-helicity (*e.g.*, 30L, 35T), with varying degrees of α-helical content between these extremes.

The peptide mixture was subjected to linear AB gradient elution on a C_8 column at all combinations of gradient-rates of 0.5%, 1.0% and 2.0% B/min and flow-rates of 0.5, 1.0 and 2.0 ml/min (eluent A was 0.1% aqueous TFA and eluent B was 0.1% TFA in acetonitrile). These parameters were chosen as being typical of the optimum range of conditions used for analytical peptide separations in RPC. Prior to the runs, the reversed-phase peptide standards, 10X–50X (Table I), were chromatographed on the C_8 column (1% B/min and 1 ml/min) and the corresponding retention times were subsequently programmed into the Standards data file for RPC (Fig. 1, screen D). The gradient delay time (t_g) for the C_8 column (3.6 min) was also programmed into the Standards file; this value is critical when the effects of gradient-rates other than 1% B/min are being simulated.

A comparison of observed peptide retention times and retention times predicted by ProDigest-LC (Table II) indicates excellent predictive accuracy by the program. The average errors at each combination of gradient-rate and flow-rate were small and, indeed, the error range for individual peptides essentially represented the maximum errors likely to be experienced by the researcher for most applications, considering the unusual nature of this test peptide mixture. The largest individual errors (up to 5.7 min for 40A) occurred during RPC under conditions where all eleven peptides in the test mixture were eluted from the column only after a long run time, *e.g.*, 0.5% B/min and 0.5 ml/min (88.0 min run time), 0.5% B/min and 1.0 ml/min (84.3 min run time) and 0.5% B/min and 2.0 ml/min (83.8 min run time). Hence even these errors were small in comparison with the total run time of the peptides. The program was never intended to predict peptide retention times down to an accuracy of 0.1 min, an average error of *ca.* 2 min appearing more realistic. The fact that this was achieved with such a difficult peptide mixture was extremely gratifying, particularly when one considers what is required from the program, *i.e.*, to predict peptide retention times over a range of gradient-rates and flow-rates with no prior information about the peptides except their amino acid composition. It is considerably more straightforward to predict the effect of variations in gradient-rate and flow-rate from observed peptide retention times[7]. However, we felt that it was important to test what the limitations of the ProDigest-LC program are at present. In addition, many, and perhaps most, average chemical or proteolytic protein digests would be likely to produce a mixture of peptides exhibiting similar extremes of size, hydrophobicity and α-helicity to that shown in Fig. 4. We are confident that we can refine the program further for even greater predictive accuracy.

The results shown in Table II (and in Fig. 3, bottom panel) also illustrate the important point that peptide conformation *per se* does not necessarily preclude the accurate prediction of peptide retention behaviour in RPC. Indeed, the predicted retention times of peptides exhibiting considerable α-helicity (*e.g.*, 30L, 40A and 21T–35T in Fig. 4) showed very satisfactory accuracy. These results suggest that, if the conformation of a peptide does not present a preferred binding site, such as can be found in the unusual situation of an amphipathic helix, where hydrophobic and non-hydrophobic residues are clustered on opposite sides of an α-helix, its reversed-

TABLE II

PREDICTED *VERSUS* OBSERVED PEPTIDE RETENTION TIMES WHEN VARYING THE FLOW-RATE AND GRADIENT-RATE

Peptide[a]	*Flow-rate 0.5 ml/min*[b]									*Flow-rate 1.0 ml/min*									*Flow-rate 2.0 ml/min*								
	0.5% B/min[b]			*1.0% B/min*			*2.0% B/min*			*0.5% B/min*			*1.0% B/min*			*2.0% B/min*			*0.5% B/min*			*1.0% B/min*			*2.0% B/min*		
	τ[c]	t_R^{obs}[d]	Δt[e]	τ	t_R^{obs}	Δt	τ	t_R^{obs}	Δt	τ	t_R^{obs}	Δt	τ	t_R^{obs}	Δt	τ	t_R^{obs}	Δt	τ	t_R^{obs}	Δt	τ	t_R^{obs}	Δt	τ	t_R^{obs}	Δt
C1	32.2	33.8	1.6	19.7	22.6	2.9	13.4	16.0	2.6	28.6	26.2	2.4	16.1	17.0	0.9	9.8	11.4	1.6	26.8	22.9	3.9	14.3	13.1	1.2	8.0	8.5	0.5
C2	35.2	39.4	4.2	21.2	25.1	3.9	14.2	17.1	2.9	31.6	32.2	0.6	17.6	19.8	2.1	10.6	12.6	2.0	29.8	27.1	2.7	15.8	16.1	0.3	8.8	9.9	1.1
20X	40.4	42.1	1.7	23.8	26.2	2.4	15.5	17.5	2.0	36.8	35.6	1.2	20.2	21.1	0.9	11.9	13.2	1.3	35.0	31.1	3.9	18.4	17.9	0.5	10.1	10.6	0.5
30X	46.2	47.3	1.1	26.7	28.5	1.8	17.0	18.5	1.5	42.6	41.2	1.4	23.1	23.7	0.6	13.4	14.3	0.9	40.8	37.3	3.5	21.3	20.7	0.6	11.6	11.9	0.3
40X	50.2	50.1	0.1	28.7	29.7	1.0	18.0	19.0	1.0	46.6	44.4	2.2	25.1	25.1	0	14.4	14.9	0.5	44.8	40.8	4.0	23.3	22.3	1.0	12.6	12.6	0
50X	52.9	51.9	1.0	30.0	30.5	0.5	18.6	19.3	0.7	49.3	46.4	2.9	26.4	26.0	0.4	15.0	15.3	0.3	47.5	43.0	4.5	24.6	23.3	1.3	13.2	13.0	0.2
21T	57.7	62.0	4.3	32.4	36.6	4.2	19.8	22.5	2.7	54.1	57.0	2.9	28.8	31.7	2.9	16.2	18.3	2.1	52.3	52.8	0.5	27.0	28.5	1.5	14.4	15.9	1.5
28T	65.8	70.3	4.5	36.5	39.9	3.4	21.9	24.0	2.1	62.2	64.3	2.1	32.9	35.2	2.3	18.3	20.0	1.7	60.4	60.5	0.1	31.1	32.2	1.1	16.5	17.7	1.2
35T	72.2	74.9	2.7	39.7	42.1	2.4	23.4	25.1	1.7	68.6	69.2	0.6	36.1	37.5	1.4	19.8	21.1	1.3	66.8	65.6	1.2	34.3	34.7	0.4	18.0	18.8	0.8
30A	84.6	80.2	4.2	45.9	44.9	1.0	26.5	26.5	0	81.0	76.3	4.7	42.3	40.1	2.2	22.9	22.4	0.5	79.2	74.5	4.7	40.5	37.2	3.3	21.1	20.2	0.9
40A	93.6	88.0	5.6	50.4	48.6	1.8	28.8	28.3	0.5	90.0	84.3	5.7	46.8	44.0	2.8	25.2	24.3	0.9	88.2	83.8	4.4	45.0	41.3	3.7	23.4	22.2	1.2
Average error			2.8			2.2			1.6			2.4			1.5			1.2			3.0			1.3			0.7

[a] Peptide sequences are shown in Table I.

[b] The peptide mixture was chromatographed on an Aquapore RP300 C_8 column (220 × 4.6 mm I.D.) under linear AB gradient elution conditions, where eluent A was 0.1% aqueous TFA and eluent B was 0.1% TFA in acetonitrile, pH 2.0.

[c] τ denotes peptide retention times (min) predicted by ProDigest-LC.

[d] t_R^{obs} denotes observed peptide retention times (min).

[e] Δt denotes error (min) between predicted and observed peptide retention times.

phase retention behaviour should be predictable. Work is currently in progress to allow even for the situations where preferred binding sites are present in a polypeptide (*e.g.*, amphipathic α-helix) and the early results appear promising.

Simulation of peptide elution profiles at varying gradient-rates and flow-rates. ProDigest-LC was applied to predict the effect of varying gradient-rate and flow-rate on the elution profile of a mixture of five decapeptide analogues closely related in hydrophobicity: I1, I2, S2, S3 and S4 (Table I). The hydrophobicity of the peptides increases only slightly between S2 and S4; between S2 and S3 there is a change from an α-H to a β-CH_3 group and between S3 and S4 there is a change from a β-CH_3 group to two methyl groups attached to the β-CH group. The hydrophobicity variations between I1, I2 and S2 are even more subtle. There is a change of only an isoleucine to a leucine residue between I1 and I2 and between I2 and S2. Guo *et al.*[16] demonstrated that leucine is slightly more hydrophobic than isoleucine, although

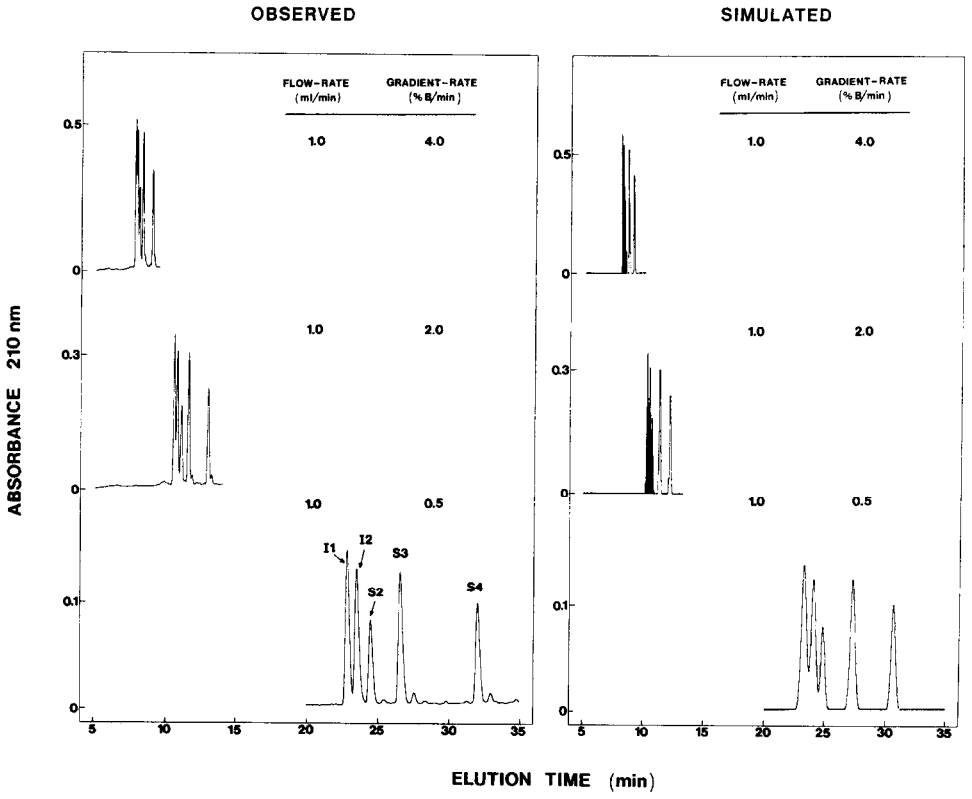

Fig. 6. Computer simulation of the effect of varying gradient-rate on the reversed-phase elution profile of a mixture of synthetic peptide standards. Column: SynChropak RP-P C_{18} (250 × 4.6 mm I.D.). Left: observed RPC elution profiles, obtained with a linear AB gradient (4.0, 2.0 or 0.5% B/min), where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0); flow-rate, 1 ml/min; temperature, 26°C. Right: simulated peptide elution profiles. Standard parameters used to obtain these simulations are shown in Fig. 1, section D; the specified peptide resolution is 0.8. The simulated profile for 0.5% B/min (bottom right) is the same as that shown in Fig. 1, section N. Shaded peak areas in the simulated profiles at 2.0 and 4.0% B/min denote unresolved peptides. The sequences of the five peptides are shown in Table I.

these residues contain the same number of carbon atoms. As isoleucine is β-branched, the β-carbon is close to the peptide backbone and not as available to interact with the hydrophobic stationary phase compared with the conformation of the leucine side-chain. Therefore, in addition to being a good simulation test for ProDigest-LC, the peptide mixture permits a very precise determination of the resolving power of a reversed-phase column.

Fig. 6 shows the observed (left) and simulated (right) elution profiles of the peptide mixture on a C_{18} column at a fixed flow-rate of 1 ml/min and gradient-rates of 4.0, 2.0 and 0.5% B/min (top, middle and bottom elution profiles, respectively). The similarity of the observed and simulated profiles is immediately apparent, with ProDigest-LC successfully simulating the major effects of varying gradient-rate on peptide elution profiles, *i.e.*, an increase in peptide retention times and peak widths, a decrease in peak height and improved peptide resolution with decreasing gradient-rate.

The required minimum peptide resolution when varying the gradient-rate was set at 0.8 (Fig. 1, screen D). This means that adjacent peaks that are not separated to at least this degree of resolution will be shaded. Thus, the simulated peptide elution profile at 0.5% B/min (Fig. 6, bottom right) suggested that all five peptides would be separated to at least a resolution of 0.8 as all peptide peaks were unshaded (the program manipulations required to produce this simulated profile are shown in Fig. 1). In contrast, with an increase in gradient-rate to 2.0% (middle right) and 4.0% B/min (top right), I1, I2 and S2 were shaded, indicating that a minimum resolution of 0.8 between these peptides had not been achieved.

Fig. 7 shows the observed (left) and simulated (right) elution profiles of the five-peptide mixture on the C_{18} column at a fixed gradient-rate of 1.0% B/min and flow-rates of 1.0, 0.5 and 0.2 ml/min (top, middle and bottom elution profiles, respectively). Similar observed and simulated profiles are again demonstrated, with the program accurately simulating the major effects of varying flow-rate on peptide elution profiles, *i.e.*, an increase in peptide retention times and peak heights and a decrease in peptide resolution with decreasing flow-rate. The effect of varying flow-rate on peak widths, unlike the gradient-rate (Fig. 6), is small. The minimum peptide resolution required was set at 0, so that all three simulated elution profiles contained unshaded peptide peaks.

The observed peptide elution profiles in Figs. 6 and 7 exhibited a slightly better resolution than the simulated profiles tended to suggest. However, as stated previously, this is a minor point considering that the program has to simulate the effects of wide variations in flow-rate and gradient-rate during the RPC of a mixture of closely related peptide analogues, with the only information about these peptides being their amino acid composition. The major effects on peptide retention times, resolution, peak heights and peak widths were certainly well predicted. It should also be noted that the researcher may simulate the effects of parameters such as gradient-rate, flow-rate and sample load on the RPC of a particular peptide mixture without ever actually running the sample on a column, simply by using the Standards default file already in the program. This then becomes a major advantage to the researcher.

Further optimization of reversed-phase peptide elution profiles

If further optimization is required following simulation of peptide elution pro-

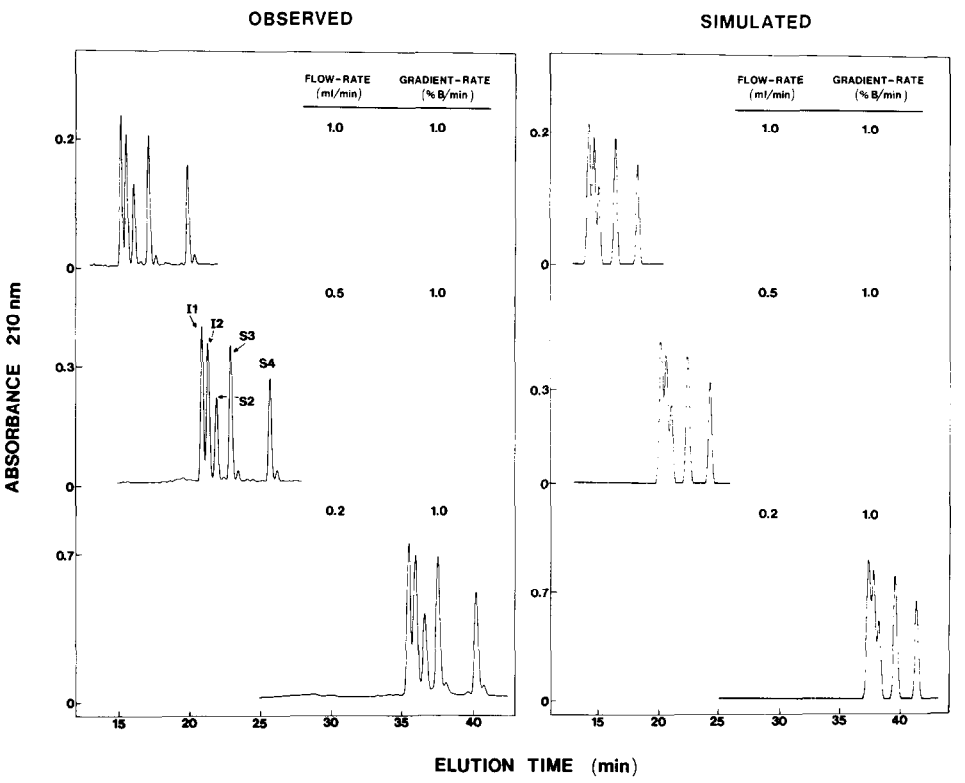

Fig. 7. Computer simulation of the effect of varying flow-rate on the reversed-phase elution profile of a mixture of synthetic peptide standards. Column: SynChropak RP-P C_{18} (250 × 4.6 mm I.D.). Left: observed RPC elution profiles, obtained with a linear AB gradient (1% B/min), where eluent A is 0.1% aqueous TFA and eluent B is 0.1% TFA in acetonitrile (pH 2.0); flow-rate, 1.0, 0.5 or 0.2 ml/min; temperature, 26°C. Right: simulated peptide elution profiles. Standard parameters used to obtain these simulations are the same as those used in Fig. 6, except that the specified peptide resolution was set at 0.0. The sequences of the five peptides are shown in Table I.

files by ProDigest-LC, a complementary optimization program such as DryLab G[26–29] may be employed. ProDigest-LC predicts peptide retention times, at varying gradient-rates and flow-rates, with no prior information about the peptides except their amino acid composition. Thus, ProDigest-LC simulates peptide elution profiles without ever having to carry out an actual chromatographic run. Having manipulated the program until the desired separation has been simulated, the researcher may then carry out the run. From this observed peptide elution profile, the researcher may then decide that further optimization of the chromatographic conditions may be required for the desired separation to be achieved. It is always more accurate to simulate the effect of changes in run parameters on observed elution profiles than to base these simulations on previously predicted profiles, and DryLab G has been designed specifically for this purpose. This program requires a minimum of two experimental runs, following which the effects on the peptide elution profile of manipulating parameters such as gradient time, gradient shape (multi-segmented gradients) and flow-rate may be simulated until the desired resolution is achieved.

CONCLUSIONS

The ProDigest-LC computer software program assists researchers in devising methodologies for the analytical separation and purification of biologically active peptides and peptide fragments from enzymatic and chemical digests of proteins by SEC, CEC and RPC. This study has clearly demonstrated the accuracy of the program in predicting peptide retention times in RPC and simulating the effects of variations in gradient-rate and flow-rate on reversed-phase peptide elution profiles. The flexibility and ease of use of ProDigest-LC should ensure its value both as a teaching aid and as an analytical tool for workers involved in peptide and protein research.

ACKNOWLEDGEMENTS

This work was supported by the Medical Research Council of Canada, S.P.I. Synthetic Peptides and the Alberta Heritage Foundation for Medical Research. We thank Vicki Luxton for typing the manuscript.

REFERENCES

1 J. M. R. Parker and R. S. Hodges, *J. Protein Chem.*, 3 (1985) 465.
2 R. S. Hodges, R. J. Heaton, J. M. R. Parker, L. Molday and R. S. Molday, *J. Biol. Chem.*, 263 (1988) 11768.
3 R. S. Hodges, J. M. R. Parker, C. T. Mant and R. R. Sharma, *J. Chromatogr.*, 458 (1988) 147.
4 F. Regnier, *Methods Enzymol.*, 91 (1983) 137.
5 C. T. Mant and R. S. Hodges, *LC, Liq. Chromatogr. HPLC Mag.*, 4 (1986) 250.
6 C. T. Mant and R. S. Hodges, *Chromatographia*, 24 (1987) 805.
7 C. T. Mant and R. S. Hodges, *J. Liq. Chromatogr.*, 12 (1989) 139.
8 C. T. Mant and R. S. Hodges, in K. Gooding and F. Regnier (Editors), *High-Performance Liquid Chromatography of Biological Macromulecules: Methods and Applications*, Marcel Dekker, New York, 1989.
9 J. L. Meek, *Proc. Natl. Acad. Sci. U.S.A.*, 77 (1980) 1632.
10 J. L. Meek and Z. L. Rossetti, *J. Chromatogr.*, 211 (1981) 15.
11 K. J. Wilson, A. Honegger, R. P. Stötzel and G. J. Hughes, *Biochem. J.*, 199 (1981) 31.
12 S.-J. Su, G. Grego, B. Niven and M. T. W. Hearn, *J. Liq. Chromatogr.*, 4 (1981) 1745.
13 T. Sasagawa, T. Okuyama and D. C. Teller, *J. Chromatogr.*, 240 (1982) 329.
14 C. A. Browne, H. P. J. Bennett and S. Solomon, *Anal. Biochem.*, 124 (1982) 201.
15 T. Sasagawa, L. E. Ericsson, D. C. Teller, K. Titani and K. A. Walsh, *J. Chromatogr.*, 307 (1984) 29.
16 D. Guo, C. T. Mant, A. K. Taneja, J. M. R. Parker and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 499.
17 D. Guo, C. T. Mant, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 519.
18 C. T. Mant and R. S. Hodges, in M. T. W. Hearn (Editor), *HPLC of Proteins, Peptides and Polynucleotides*, VCH, Weinheim, in press.
19 C. T. Mant, T. W. L. Burke, J. A. Black and R. S. Hodges, *J. Chromatogr.*, 458 (1988) 193.
20 M. J. O'Hare and E. C. Nice, *J. Chromatogr.*, 171 (1979) 209.
21 E. C. Nice, M. W. Capp, N. Cooke and M. J. O'Hare, *J. Chromatogr.*, 218 (1981) 569.
22 K. J. Wilson, A. Honegger and G. J. Hughes, *Biochem. J.*, 199 (1981) 43.
23 C. T. Wehr, L. Correia and S. R. Abbott, *J. Chromatogr. Sci.*, 20 (1982) 114.
24 S. Y. M. Lau, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 317 (1984) 129.
25 S. Y. M. Lau, A. K. Taneja and R. S. Hodges, *J. Biol. Chem.*, 259 (1984) 13253.
26 J. W. Dolan and L. R. Snyder, *LC · GC, Mag. Liq. Gas Chromatogr.*, 5 (1987) 971.
27 B. F. D. Ghrist, B. S. Cooperman and L. R. Snyder, *J. Chromatogr.*, 459 (1988) 1.
28 B. F. D. Ghrist and L. R. Snyder, *J. Chromatogr.*, 459 (1988) 25.
29 B. F. D. Ghrist and L. R. Snyder, *J. Chromatogr.*, 459 (1988) 43.