METHODS IN MOLECULAR BIOLOGY™

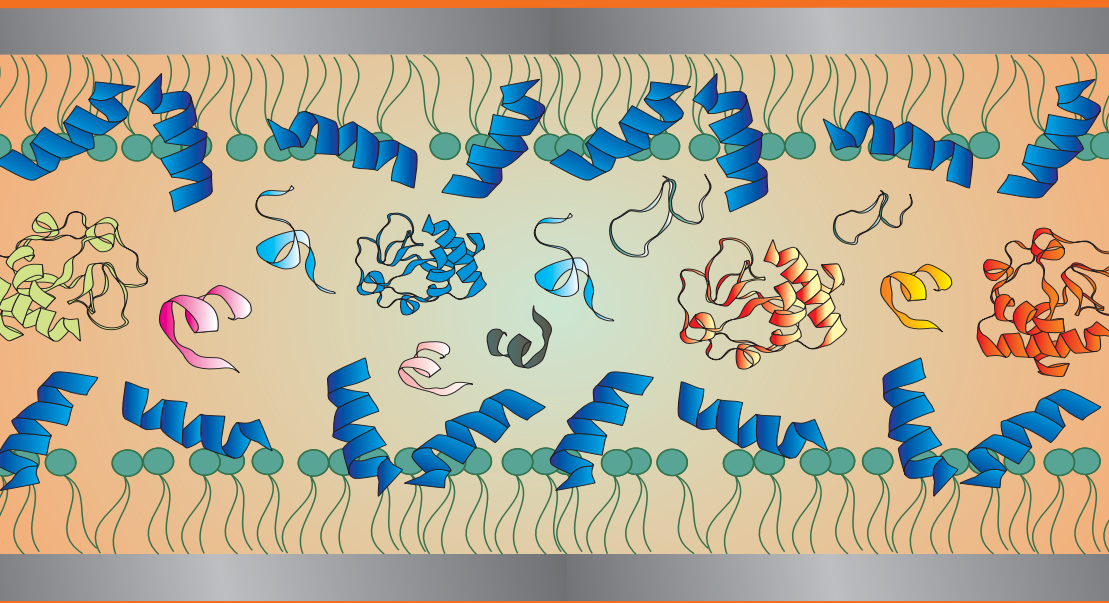
Volume 251

HPLC of Peptides and Proteins

Methods and Protocols

Edited by

Marie-Isabel Aguilar



HUMANA PRESS

1

HPLC of Peptides and Proteins

Basic Theory and Methodology

Marie-Isabel Aguilar

1. Introduction

High-performance liquid chromatography (HPLC) is now firmly established as the premier technique for the analysis and purification of a wide range of molecules. In particular, HPLC in its various modes has become the central technique in the characterization of peptides and proteins and has, therefore, played a critical role in the rapid advances in the biological and biomedical sciences over the last 10 years.

The enormous success of HPLC can be attributed to a number of inherent features associated with reproducibility, ease of selectivity manipulation, and generally high recoveries. The most significant feature is the excellent resolution that can be achieved under a wide range of conditions for very closely related molecules, as well as structurally quite distinct molecules. This arises from the fact that all interactive modes of chromatography are based on recognition forces that can be subtly manipulated through changes in the elution conditions that are specific for the particular mode of chromatography. Peptides and proteins interact with the chromatographic surface in an orientation-specific manner, in which their retention time is determined by the molecular composition of specific contact regions. For larger polypeptides and proteins that adopt a significant degree of secondary and tertiary structure, the chromatographic contact region comprises a small proportion of the total molecular surface. Hence, the unique orientation of a peptide or protein at a particular stationary phase surface forms the basis of the exquisite selectivity that can be achieved with HPLC techniques. All biological processes depend on specific

From: *Methods in Molecular Biology, vol. 251, HPLC of Peptides and Proteins: Methods and Protocols*
Edited by: M.-I. Aguilar © Humana Press Inc., Totowa, NJ

interactions between molecules and affinity chromatography exploits these specific interactions to allow the purification of a biomolecule on the basis of its biological function or individual chemical structure. In contrast reversed phase HPLC, ion-exchange and hydrophobic interaction chromatography separate peptides and proteins on the basis of differences in surface hydrophobicity or surface charge. These techniques therefore allow the separation of complex mixtures whereas affinity chromatography normally results in the purification of one or a small number of closely related components of a mixture.

Reversed-phase chromatography (RPC) is arguably the most commonly used mode of separation for peptides, although ion-exchange (IEC) and size exclusion (SEC) chromatography also find application. The three-dimensional structure of proteins can be sensitive to the often harsh conditions employed in RPC, and as a consequence, RPC is employed less for the isolation of proteins where it is important to recover the protein in a biologically active form. IEC, SEC, and affinity chromatography are therefore the most commonly used modes for proteins, but RPC and hydrophobic interaction (HIC) chromatography are also employed.

HPLC is extremely versatile for the isolation of peptides and proteins from a wide variety of synthetic or biological sources. The number of applications of HPLC in peptide and protein purification continue to expand at an extremely rapid rate. Solid-phase peptide synthesis and recombinant DNA techniques have allowed the production of large quantities of peptides and proteins which need to be highly purified. The design of multidimensional purification schemes to achieve high levels of product purity further highlight the power of HPLC techniques in the analysis and isolation of peptide and proteins samples. The complexity of the mixture to be chromatographed depends on the nature of the source and the degree of preliminary clean-up that can be performed. In the case of synthetic peptides, RPC is generally employed both for the initial analysis and the final large scale purification. The isolation of proteins from a biological cocktail however, often requires a combination of techniques to produce a homogenous sample. HPLC techniques are then introduced at the later stages following initial precipitation, clarification and preliminary separations using soft gel. Purification protocols therefore need to be tailored to the specific target molecule. The key factor that underpins the development of a successful separation protocol is the ability to manipulate the retention of the target molecule so that it can be resolved from other contaminating components. This chapter thus provides an outline of the general theory of chromatography and the factors that control both the retention time and peakwidth of solutes undergoing separation in terms of the parameters that control resolution. This information can then be used to understand the approaches used to perform

separations with specific modes of chromatography as outlined in the remaining chapters in this book.

2. The Molecular Basis of Separation

The separation of a mixture of peptides and proteins in interactive modes of chromatography arises from the differential adsorption of each solute according to their respective affinity for the immobilized stationary phase. Thus, when a particular molecule has a very high affinity for a specific stationary phase, i.e., when the equilibrium distribution coefficient K is high, then that solute is retained to a greater extent than another molecule with a lower affinity for the stationary phase. The degree and nature of the binding affinity is clearly dependent on the structure of the solute and the immobilized ligands. For example, in the case of RPC and HIC, binding is mediated predominantly through hydrophobic interactions between the solute and the immobilized *n*-alkyl ligands. In IEC, the binding is through electrostatic interactions, whereas in different modes of affinity chromatography, binding involves a mixture of hydrophobic, electrostatic, and polar forces. In the case of size exclusion chromatography, the differential movement along the column is a result of the extent to which each solute can permeate the porous structure of the stationary phase.

An additional factor that influences the appearance and relative separation of a peak is the degree of bandbroadening of the solute band during migration through the column. Thus, as it moves down the column, the solute band broadens as a consequence of a number of factors including longitudinal diffusion, brownian motion, eddy diffusion, and mobile phase and stagnant phase mass transfer. These effects result in bandbroadening that generally increases with increasing residence time in the column. The resulting degree of separation or selectivity between constituent solutes in a mixture is thus a subtle interplay between the relative affinity of the molecules for the stationary phase and the degree of diffusive processes that occur during separation.

3. Retention and Bandwidth Relationships

The time taken for a solute to pass though a chromatographic column is referred to as the retention time t_r. This retention time is measured as the time taken by the solute, following injection, to emerge from the column and to be detected as illustrated in **Fig. 1**. In order to allow retention times to be compared to different columns or under different conditions, the retention time of a solute is normally compared with the retention time of a molecule which is not retained on the specific column of interest. This allows the unitless capacity factor k' of a solute to be expressed in terms of the retention time t_r, through the relationship

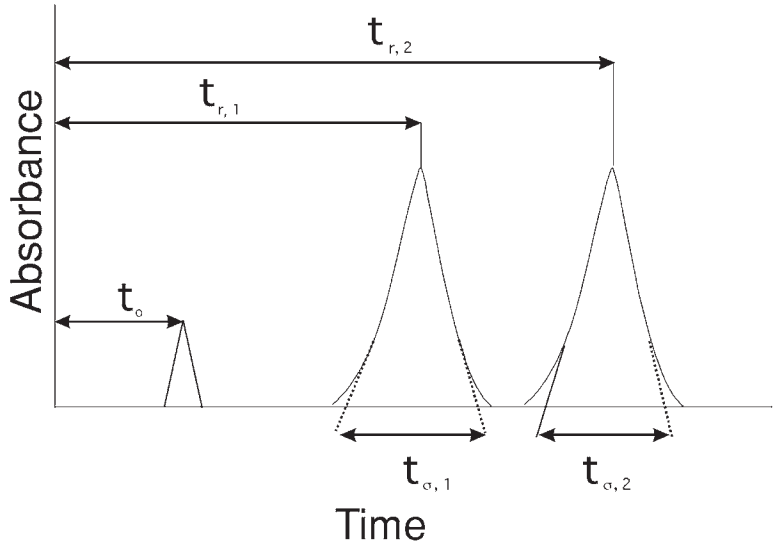


Fig. 1. Diagram of the retention parameters that describe a chromatographic separation. The retention time of a nonretained solute is denoted by t_0, while the retention times of two retained solutes, 1 and 2, are given by $t_{r,1}$ and $t_{r,2}$. The corresponding peak-widths for solutes 1 and 2 are denoted σ,1 and σ,2, and together with the retention times are they used to detrmine the resolution of the separation according to **Eq. 9.**

$$k' = (t_r - t_o) / t_o \tag{1}$$

where t_o is the retention time of a nonretained solute. The capacity factor k' can also be defined as the ratio n_s/n_m where n_s and n_m are the number of moles of solute in the stationary phase and mobile phase respectively as follows:

$$k' = n_s / n_m \tag{2}$$

or alternatively as

$$k' = [X]_s \, V_s / [X]_m \, V_m \tag{3}$$

where $[X]_s$ and $[X]_m$ refer to the concentrations of the solute in the stationary and mobile phases, respectively, and V_s and V_m are the corresponding volumes of the stationary and mobile phases. Since the ratio $[X]_s / [X]_m$ is the equilibrium distribution coefficient K and the ratio V_s / V_m defines the phase ratio Φ of the chromatographic system, the capacity factor can also be expressed as follows:

$$k' = \Phi [X]_s / [X]_m \tag{4}$$

or

$$k' = \Phi K \tag{5}$$

Equation 5 thus formerly describes the direct thermodynamic relationship between the retention of a peptide or protein and its affinity for the stationary phase material.

The practical significance of k' can be related to the selectivity parameter α, defined as the ratio of the capacity factors of two adjacent peaks as follows:

$$\alpha = k'_i / k'_j \tag{6}$$

which allows the definition of a chromatographic elution window in which retention times can be manipulated to maximise the separation of components within a mixture. Clearly, the aim is to obtain as high a value of α as possible, which reflects a high degree of separation between two peaks. The second factor involved in defining the quality of a separation is the peak width σ_t. The degree of peak broadening is directly related to the efficiency of the column and can be expressed in terms of the number of theoretical plates N as follows:

$$N = (t_r)^2 / \sigma_r^2. \tag{7}$$

N can also be expressed in terms of the reduced plate height equivalent h, the column length L, and the particle diameter of the stationary phase material d_p, as

$$N = hL / d_p. \tag{8}$$

The resolution R_s between two components of a mixture, therefore, depends on both selectivity and bandwidth according to

$$R_s = 1 / 4 \sqrt{N} (\alpha - 1)[1 / (1 + k')]. \tag{9}$$

This equation describes the relationship between the quality of a separation and the relative retention, selectivity, and the bandwidth. It also provides the formal basis upon which resolution can be manipulated to achieve a particular level of separation. Thus, when faced with an unsatisfactory separation, the aim is to improve resolution by one of three possible strategies. The first is to increase α as previously and the second, but related, approach is to vary k' within a defined range normally $1 < k' < 10$ through variation in the experimental elution conditions such as solvent strength, separation time, or nature of the immobilized ligand. Third, one can increase N, for example, by using very small particles in microbore or narrow bore columns.

An appreciation of the factors that control the resolution of peptides and proteins in interactive modes of chromatography can assist in the development and manipulation of separation protocols to obtain the desired separation. The optimization of high-resolution separations of peptides and proteins involves the separation of sample components through manipulation of both retention times and solute peak shape. For example, inspection of the schematic separation shown in **Fig. 1** demonstrates baseline separation between the two components

which corresponds to a high value of both selectivity α, and resolution. A scenario can be envisaged where it may be desirable to decrease the retention times of the solutes to allow more rapid analysis times. However, resolution may be sacrificed and the final separation conditions are often likely to be a tradeoff between rate of analysis and quality of separation.

An enormous range of different separation techniques are available for peptide and protein analysis. The challenge facing the scientist who wishes to analyze and/or purify their peptide or protein sample is the selection of the initial separation conditions and subsequent optimisation of the appropriate experimental parameters. The following chapters thus provide a practical guide to performing peptide and protein analyses under a range of different separation modes. In addition, the reader is guided through the experimental options available to achieve a high-resolution separation of a peptide or protein mixture, an exercise which is underpinned by the theoretical relationships provided in this chapter.