Instrumentation in Amino Acid Sequence Analysis

edited by

R. N. Perham

Department of Biochemistry
and St. John's College, Cambridge

1975

Academic Press

London New York San Francisco

A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS INC. (LONDON) LTD
24/28 Oval Road
London NW1

U.S. Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York
New York 10003

Library of Congress Catalog Card Number 75 34615
ISBN 0 12 551250 3

Printed in Great Britain by
William Clowes & Sons Limited
London, Colchester and Beccles

CONTRIBUTORS

BENSON, J. R. *Durrum Chemical Corporation, 3950 Fabian Way, Palo Alto, California 94303, U.S.A.*

BONNER, A. G. *Department of Chemistry, Boston University, Boston, Massachusetts 02215, U.S.A.*

BRIDGEN, J. *MRC Laboratory of Molecular Biology, Cambridge*

DELL, A. *MRC Laboratory of Molecular Biology and University Chemical Laboratory, Cambridge*

HORN, M. J. *Department of Chemistry, Boston University, Boston, Massachusetts 02215, U.S.A.*

GRAFFEO, A. P. *North-eastern University, Boston, Massachusetts, U.S.A.*

KARGER, B. L. *North-eastern University, Boston, Massachusetts, U.S.A.*

LAURSEN, R. A. *Department of Chemistry, Boston University, Boston, Massachusetts 02215, U.S.A.*

MORRIS, H. R. *MRC Laboratory of Molecular Biology and University Chemical Laboratory, Cambridge*

WATERFIELD, M. D. *Imperial Cancer Research Fund Laboratories, London*

PREFACE

When, some four years ago, I received an invitation to edit a book on automated methods of amino acid sequence analysis, I felt bound to reply that the project was premature. The subject was in a phase of rapid development and critics at that time were far from unanimous in their praise of the emerging technology. But the smoke of battle has now cleared somewhat, leaving—and there is no contradiction here—both sides masters of the field. The automated methods have vindicated their authors and given ample demonstration of their power and growing versatility. At the same time, there has been a wider realization of their high cost, both in capital outlay and in day-to-day running. The proponents of the manual methods can therefore rest assured that for many purposes such methods will be hard to displace.

Thus, the time now seems ripe to try to draw together expert commentary on the various automated methods. Our intention has been to produce a book with its emphasis on the practical aspects, but which is more than a mere laboratory manual. The contributors were commissioned to describe what their machines can do and to point out where the methods stand in need of further improvement. This they have done with admirable clarity and candour and with the implied hope that readers will also be stimulated to try to improve the techniques. The problems cover a wide range of science, from chemistry to engineering, and solutions may well come from those who are not themselves applying automated methods of sequence analysis.

The first method to be automated was in fact amino acid analysis by ion-exchange chromatography, in the now classic work of Moore and Stein. In the opening chapter, Dr. Benson documents the impressive advances in speed and sensitivity that have recently taken place. Dr. Waterfield and Dr. Bridgen then describe the use and potential of the liquid-phase sequencer, based on the pioneering work of Edman in its chemistry and its mechanics. It might be worth recalling here that the peptide degradation with phenylisothiocyanate that bears his name was first published in 1950, which qualifies it for a well-earned silver jubilee. The same degradation is

put to good use in the solid-phase sequencer described in Chapter 3 by Dr. Laursen, Dr. Bonner and Dr. Horn. This chapter provides detail enough for the construction of a home-made sequencer, something that has been successfully undertaken in a number of laboratories. The authors of Chapters 2 and 3 concur in their assessment of the complementarity of the liquid- and solid-phase sequencers, the former being preferred for proteins and peptides in excess of 30 residues and the latter best suited to the shorter peptides.

Both types of sequencer generate phenylthiohydantoins and Chapter 4 is devoted to their identification. Perhaps the most interesting aspect of this work is the promising application of high-pressure-liquid chromatography but it is clear that a battery of techniques, including the use of the amino acid analyser, must be deployed for unequivocal identification. If one is to use sequencers, there is no relief from expense even here.

In Chapter 5, Dr. Morris and Dr. Dell give an account of the theory and practice of mass spectrometry in amino acid sequence analysis. This approach has undergone enormous improvements in the past few years and has now advanced to the point at which it must be regarded as one of the methods of choice, particularly when unusual amino acids are present. Indeed, as improvements continue to be made in mixture analysis and alternative methods of ionization, it is easy to see mass spectrometry coming to occupy a commanding position in peptide sequence work. The prospect of coupling mass spectrometric analysis to peptide separation by gas chromatography is an added attraction that now seems likely to be realized.

There can be no doubt, then, that automated methods of sequence analysis have demonstrated their utility. The only question that hangs over them is one of cost. Mass spectrometers, which can be found already in most chemical laboratories, may be at a small advantage here but the sums involved in buying a machine to dedicate to sequence work are daunting, to say the least. For those laboratories not specializing in sequence analysis, therefore, manual methods are likely to continue to provide the most effective and economic way of obtaining the occasional and limited information they may wish to acquire. But, if "Life is one long process of getting tired" (Samuel Butler, *Note Books. Life,* vii), the daily round of the sequence laboratory has been appreciably quickened by the introduction of the instrumentation described in this book. We can be certain that the story will not end here.

Richard Perham
Cambridge, August 1975

ABBREVIATIONS

BSA	*N,O*-bis(trimethylsilyl)acetamide
DCE	dichloroethane
DMAA	*N,N*-dimethylallylamine
DMBA	*N,N*-dimethylbenzylamine
HFBA	heptafluorobutyric acid
MTH	methylthiohydantoin
NCS	*N*-chlorosuccinimide
OPA	*o*-phthalaldehyde
PCA	pyrrolidone carboxylic acid
PITC	phenylisothiocyanate
PTH	phenylthiohydantoin
Quadrol	*N,N,N′,N′*-tetrakis (2-hydroxypropyl)ethylenediamine (reg. trade mark Wyandotte Chemical Corp^{n}.)
TETA	triethylenetetramine
TFA	trifluoroacetic acid
TMS	trimethylsilyl

CONTENTS

1. Some Recent Advances in Amino Acid Analysis

JAMES ROBERT BENSON

Durrum Chemical Corporation, 3950 Fabian Way,
Palo Alto, California 94303, U.S.A.

I. Introduction

The development of the first automatic amino acid analyser by Spackman, Stein and Moore (1) has played a major part in the successful sequence analysis of over three hundred proteins. Scores of protein sequence laboratories have appeared around the world, and thousands of instruments capable of automatically analysing amino acids have been sold. As protein purification techniques have been refined and improved, biochemists have isolated increasing numbers of proteins for which more complete structural characterization is desired. Development of automatic sequencers has provided a means of greatly accelerating primary structure determinations and such sequencing devices are now challenging the speed of conventional amino acid analysers. All this has led to increased demands for shorter analysis times and more complete automation of instrumentation.

Concomitant with interest in faster analyses of amino acids is

interest in higher sensitivity. Most proteins available in gram quantities have already been sequenced but many proteins of biological interest are available only in milligram or microgram amounts. Analyses of amino acids at nanomolar levels (10^{-9}) are now common, but need for picomolar (10^{-12}) and sub-picomolar analyses is becoming apparent. This chapter deals with recent advances in amino acid analysis that pertain to improved sensitivity and increased speed. It is not intended as a review of the entire field, but is a discussion of some of the newer techniques which will most significantly affect amino acid analysis during the next few years.

Ion-exchange chromatography has been the most widely used means of separating amino acids, although gas-liquid chromatography has frequently been adopted. The method originally described by Spackman, Stein and Moore (1) employed dual ion-exchange columns. A long column (50 to 100 cm) was used to separate acidic and neutral amino acids, after which the analysis was terminated, leaving basic amino acids on the resin. An identical sample was applied to a short column (5 to 10 cm) that separated the basic amino acids. Before the next analysis could be accomplished, the long column had to be regenerated to remove the basic amino acids still bound to the resin. Baseline instability and artifacts precluded analysis of all amino acids with only one column.

Objections to the dual column method are several. When high sensitivity is demanded, requirements for two separate samples for a single analysis can be limiting. Analytical precision is compromised because introduction of a second sample increases possibilities of error. Finally, automation of instrumentation is complicated by requirements for additional columns, valves, and programming. A single-column method of analysing amino acids is preferable, provided baseline irregularities can be controlled. The recent development of an effective ammonia trap, improved buffer solution formulations and improved photometers makes successful single-column analyses possible.

Piez and Morris (2) developed a single-column method that employed elution with a continuous gradient buffer solution. Hamilton (3) devised a protocol utilizing discrete buffer solutions to elute all amino acids from a single column. A system of discrete buffers is preferred because it is more flexible, more easily defined, and because baseline irregularities are more easily controlled. In addition, gradient methods usually require longer analysis times. For these reasons, only single-column systems employing discrete buffer solutions are discussed in this chapter.

II. Separation Methods

A. ION-EXCHANGE CHROMATOGRAPHY

(1) Principles

Ion-exchange resins for separating amino acids are normally made of polystyrene that has been crosslinked with divinylbenzene. This copolymer is then sulphonated to yield the structure illustrated in Fig. 1. The sulphonic acid groups are strongly anionic and form the sites upon which ion-exchange occurs.

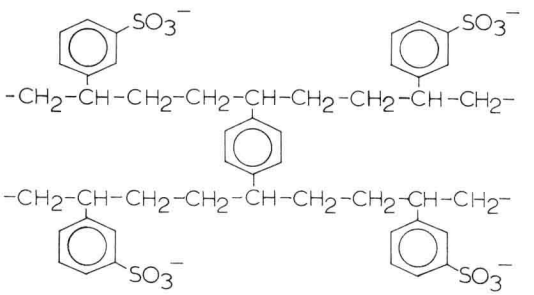

FIG. 1. The structure of sulphonated, cross-linked polystyrene resins. Divinylbenzene is a cross-linking agent that develops a three-dimensional matrix with polystyrene. Treatment of this copolymer with sulphuric acid results in attachment of sulphonic acid functional groups. These form the sites for ion-exchange.

Although amino acids can exist as either anions or cations, it is convenient to place them in cationic form:

$$H_2N{-}\underset{\displaystyle R}{\underset{|}{C}}HCOOH + H^+ \rightleftarrows {}^+H_3N{-}\underset{\displaystyle R}{\underset{|}{C}}HCOOH \qquad \text{(Eqn. 1)}$$

As such, they are attracted to the negatively-charged sulphonic acid groups of the resin. Those amino acids that interact weakly with the resin are poorly retained and elute early from columns; those that interact strongly are more strongly retained and elute later.

The dissociation depicted in Eqn. 1 can be controlled by the pH of the system. At low pH values, the increased concentration of hydrogen ions shifts equilibrium to the right, thereby increasing the concentration of ions that can interact with sulphonic acid groups. Thus, at low pH values, amino acids are more strongly retained by the resin. As the pH is gradually raised, equilibrium shifts to the left,

eventually resulting in amino acids losing their positive charge. Under such conditions, amino acids are not retained by the resin and are eluted from columns.

Although pH is the most important factor in determining relative elution positions, the rate at which amino acids elute from columns is also strongly influenced by the cation concentration in eluting buffer solutions. If sodium citrate buffer solutions are employed, sodium ions compete with cationic amino acids for the ion-exchange sites present in the resin. This is expressed as follows:

$$Na^+ + PSO_3\bar{A} \rightleftharpoons PSO_3Na + \bar{A}^+ \qquad \text{(Eqn. 2)}$$

where $\bar{A}$ represents an amino acid and PSO_3 represents the sulphonic acid group attached to the resin copolymer. Although amino acids have a greater affinity for the resin than sodium ions, the sodium ions are present in much higher concentrations and are continually supplied to the column. As a result, the equilibrium in Eqn. 2 is shifted to the right and the amino acid is displaced ("exchanged"). If buffer solutions containing higher concentrations of sodium ions are utilized, the rate of ion-exchange is increased and amino acids are eluted more rapidly from the resin.

The nature of amino acid side chains greatly influences elution position. Ionizable groups on side chains affect binding, either retarding or accelerating migration of residues within the resin bed. For example, aspartic acid is eluted early because its acidic side chain tends to neutralize the positive charge of the α-amino group, thereby reducing the affinity for the negatively charged resin. However, other types of interaction are also important. For example, a hydrophobic side chain can affect elution behaviour by interacting with non-polar, hydrophobic regions on the resin matrix. This type of interaction is shown by tyrosine (pK_a = 9.1), which is eluted after leucine (pK_a = 9.7). Without the aromatic side chain, elution of tyrosine would occur much earlier. Thus, although resins contain polar, ionizable species, the copolymer matrix remains strongly hydrophobic and can itself interact with amino acids independently of ion-exchange effects. These interactions are not necessarily adverse. Two molecules with identical net charge at a given pH value will not co-elute if their interactions with the resin matrix differ; a separation which would not be possible otherwise is thereby achieved.

In developing a protocol for the quantitative resolution of amino acids, emphasis is generally given to pH and molarity (cation concentration) of buffer solutions. All amino acids found in hydrolysed proteins can be eluted from a single ion-exchange column by

utilizing a buffer solution sequence of (a) increasing pH value and constant molarity, (b) constant pH and increasing molarity, or (c) some combination of the two. Variables such as column temperature, eluent flow velocity and choice of anion play a secondary role in defining a protocol for such a separation.

(2) pH Gradients

Most methods for separating amino acids on a single ion-exchange column rely on sequential application of a series of buffer solutions of increasing pH and molarity. The method of Dus *et al.* (4) utilizes four sodium citrate buffer solutions: the first two solutions each contain 0.20 *M* Na^+; the third contains 1.4 *M* Na^+ (pH 6.4); the fourth contains 0.7 *M* Na^+ (pH 10.9). The high pH of the fourth buffer solution is necessary to ensure elution of the basic amino acids within a reasonable time. Citrate buffer solutions are particularly suited for this application because citric acid has p*K* values of 3.14,

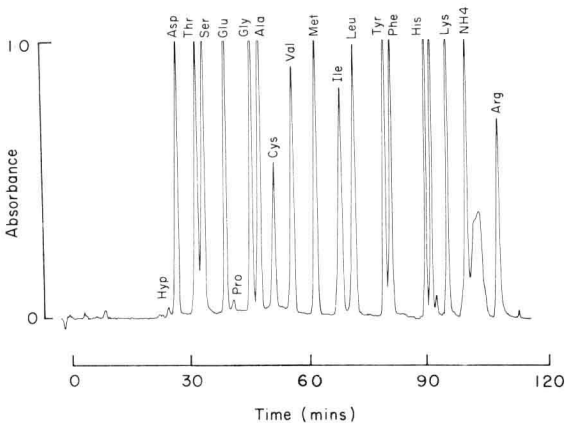

FIG. 2. Amino acid analysis using constant molarity buffer solutions. A calibration standard containing one nanomole of each amino acid was applied to a 0.20 x 25 cm bed of Durrum DC-4A resin. Detection was accomplished with ninhydrin using a 60 mm optical path length flowcell at 570 nm. Four buffer solutions were sequentially applied to the column; pH values ranged from 3.25 to 10.1. Sodium ion concentration was constant at 0.20 *M* for all solutions, which were pumped at constant pressure (see text). The peak appearing between histidine and lysine is hydroxylysine. Artifacts appearing in the region of basic amino acids are caused by buffer solution contaminants. Hydroxyproline is resolved by lowering the pH of the initial buffer solution to about 2.9; then following with the usual program beginning with pH 3.25 buffer. (Chromatogram courtesy of P. E. Hare.)

5.95, and 6.39, thereby providing optimum buffering capacity over this pH range. However, above pH values of 6.5, other anions must be used.

Hare (5) described a constant molarity buffer system that utilizes borate as the anion in the final buffer solution. Sodium ion concentration is kept constant at 0.20 *M* in the solutions used and a stepwise pH gradient ranging from pH 3.25 to pH 10.1 is employed to elute amino acids from a single ion-exchange column. A typical analysis is shown in Fig. 2. With this method, resin beds do not shrink or swell during the analysis and column pressures remain low. Buffer solutions containing high salt concentrations cause resins to contract. Although re-expansion occurs during equilibration with the solutions of low ionic strength, the re-equilibration time can be lengthy with some resin formulations. Also, the continued shrinkage

TABLE I
Constant molarity buffer system (after Hare)

Solution	pH	Sodium Citrate (dihydrate)	Boric Acid	NaOH	NaCl	Volume
A	3.25	19.6 g	—	—	—	1000 ml
B	4.15	19.6 g	—	—	—	1000 ml
C	5.25	19.6 g	—	—	—	1000 ml
D	10.1	—	3.0 g	2.0 g	8.76 g	1000 ml

and expansion of resin beds exposed to alternating low and high molarity buffer solutions causes higher resin density within columns. As a result, pressures are usually higher than in those systems utilizing constant molarity buffer solutions. Another advantage of Hare's system is that regeneration of the resin bed with sodium hydroxide solution is not required because the pH of the borate buffer (pH = 10.1) is sufficient to regenerate the bed. The formulation for Hare's buffer system is shown in Table I.

The buffer solution A elutes amino acids through cystine; the solution B through leucine; the solution C through tyrosine and phenylalanine and the buffer solution D elutes the remaining amino acids and regenerates the resin bed (see Fig. 2).

(3) Concentration Gradients

Another approach to elution of amino acids from a single ion-exchange bed involves sequential application of a series of buffer solutions in

which each solution contains a higher concentration of cations than the preceding one, but in which pH values remain nearly constant. This system, developed by Benson (6), has the advantage of providing more stable baselines than systems employing pH gradients. A chromatogram illustrating the results of an amino acid analysis performed with this buffer system is shown in Fig. 3.

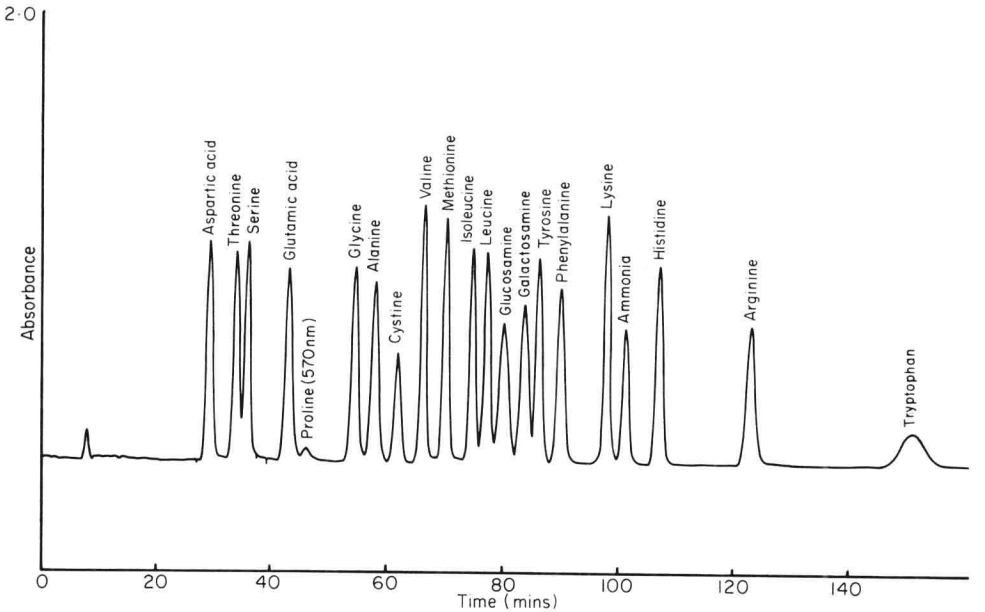

FIG. 3. Analysis of amino acids using iso-pH buffer system. A calibration standard containing 25 nmol of each amino acid was applied to a 0.90 x 30 cm bed of Durrum DC-6A resin. Detection was accomplished with ninhydrin using a 12 mm optical path length flowcell at 570 nm. Three buffer solutions of increasing sodium ion concentration at nearly constant pH values were used as eluents (Pico-Buffer System II, Durrum Chemical Corp., Palo Alto, CA). Column temperature was 45°C, changed to 65°C at 25 min. Eluent flowrates of 70 ml h^{-1} required eluent pressures of 27 to 30 atm. Note that tryptophan elutes after arginine and that the baseline in the region of the basic amino acids is flat.

With this "iso-pH" buffer system, pH values of all buffer solutions are low, and amino acids are eluted by the increasing concentration of sodium ions. As discussed earlier, this has the effect of shifting the equilibrium of Eqn. 2 to the right, thereby releasing the amino acids bound to the resin. The rate of this release, and to some extent the speed of the analysis, can be controlled by sodium ion molarity. A buffer formulation based on this principle is given in Table II.

TABLE II
Iso-pH buffer system (after Benson (6))

Solution	pH	$Na^+(M)$	Sodium Citrate (dihydrate)	NaCl	Volume
A	3.25	0.20	19.6 g	—	1000 ml
B	3.50	0.70	19.6 g	29.2 g	1000 ml
C	3.65	1.60	19.6 g	81.8 g	1000 ml

The buffer solution A elutes amino acids through alanine; the B solution elutes from cystine to leucine and the C solution elutes tyrosine through arginine (see Fig. 3).

(4) Resolution and Speed of Analyses

Once a protocol for separating amino acids has been developed it is possible to adjust certain parameters in order to reduce analysis time without losing resolution. Hamilton *et al.* (7) first formulated a theoretical basis for understanding how chromatographic conditions affect resolution. They demonstrated that the resolution between two amino acids, R_{ab}, is given by

$$R_{ab} \propto \frac{1}{d_p}\left(\frac{Z}{2U_0}\right)^{1/2} \qquad \text{(Eqn. 3)}$$

where d_p is the mean resin particle diameter in centimetres, Z is the bed length in centimetres, and U_0 is the linear eluent flow velocity in cm min^{-1}. U_0 equals the flowrate in ml min^{-1} divided by the cross-sectional area of the bed in cm^2.

Empirically, reduction in analysis time can be accomplished by increasing buffer solution flow-rate or by decreasing resin bed length; however, Eqn. 3 indicates that these changes result in loss of resolution. Consequently, shorter analysis times without concomitant loss in resolution can only be accomplished by using resins with smaller mean particle diameters. Benson (8) showed that Hamilton's expression (Eqn. 3) should be modified to include ϕ, the "packing density" of columns, defined as m/V_c, where m is the mass of resin in grams in the column and V_c is the volume of the empty column in cm^3. He derived

$$t_2 = t_1 \frac{Z_2 \phi_2 U_{01}}{Z_1 \phi_1 U_{02}} \qquad \text{(Eqn. 4)}$$

which defines the relationship between the initial analysis time, t_1, and the revised analysis time, t_2. To prevent loss of resolution as a result of modifying any of the above parameters, the mean particle diameter of the resin must be reduced. The required diameter d_{p2}, is given by

$$d_{p2}^2 = \frac{Z_2 U_{01} \phi_2}{Z_1 U_{02} \phi_1} d_{p1}^2 \qquad \text{(Eqn. 5)}$$

where d_{p1} is the mean particle diameter used in the initial analysis. It is important to recognize that the analysis time for any separation method can be reduced, but only if the resin particles are made smaller. Equation 4 allows the analysis time to be made as short as desired and Eqn. 5 then defines the resin particle diameter required for that new shorter analysis time.

An application of these computations is shown in Fig. 4 in which an amino acid analysis was performed on a single ion-exchange bed in about 40 min with resin having a mean particle diameter of 7 μm. The resin bed length was 30 cm and the linear flow velocity of eluent was in excess of 5 cm min^{-1}.

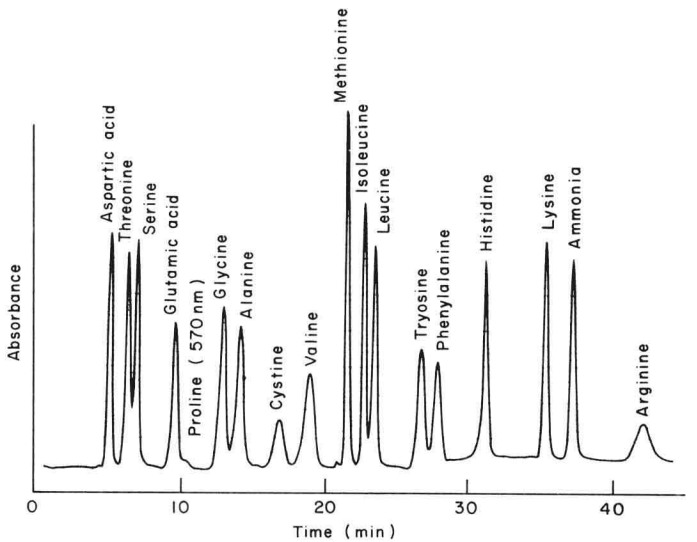

FIG. 4. Amino acid analysis using high-pressure syringe pumps. This analysis was performed with a Durrum Model D-500 Amino Acid Analyzer (Durrum Instrument Corp., Palo Alto, California) using constant displacement syringe pumps. A calibration standard containing 10 nmol of each amino acid was applied to a 0.18 x 30 cm bed of Durrum DC-4A-7 (mean particle diameter equals 7 μm). To achieve this rapid analysis eluent pressures of *ca.* 185 atm were required.