THE IDENTIFICATION OF PHENYLTHIOHYDANTOIN AMINO ACIDS BY HIGH PRESSURE LIQUID CHROMATOGRAPHY

by

Carol Badger McCullough

Submitted in Partial Fulfillment of the Requirements
for the Degree of
Master of Science
in the
Chemistry
Program

June 2, 1980
Advisor Date

June 11, 1980
Dean of the Graduate School Date

YOUNGSTOWN STATE UNIVERSITY

June, 1980

YOUNGSTOWN STATE UNIVERSITY

Graduate School

THESIS

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Science

TITLE

"Analysis and Identification of Phenylthiohydantoin Amino Acids by High Pressure Liquid Chromatography"

PRESENTED BY Carol McCullough

ACCEPTED BY THE DEPARTMENT OF Chemistry

Major Professor	June 2, 1980 Date
Marvin Lukin	June 17, 1980 Date
Elmer Foldvary	June 20, 1980 Date
Dean, Graduate School	June 24, 1980 Date

ABSTRACT

THE IDENTIFICATION OF PHENYLTHIOHYDANTOIN AMINO ACIDS BY HIGH PRESSURE LIQUID CHROMATOGRAPHY

Carol Badger McCullough

Master of Science

Youngstown State University, 1980

A method of analysis of phenylthiohydantoin amino acids was developed using high pressure liquid chromatography. This analysis is useful both quantitatively and qualitatively in protein sequencing, as the residues of the protein degradation are converted to the stable phenylthiohydantoin derivative before identification.

When using an automated sequencer to cleave a protein via the Edman degradation, the time of analysis is critical so that the identification is synchronized with the degradation. A gradient elution profile, in which all 20 phenylthiohydantoin amino acids may be identified, has been developed. The analysis requires 25 minutes, which is adequate to allow immediate feedback to the sequencer. Reversed phase chromatography, using a methanol/sodium acetate buffer mixture as the mobile phase was found to be adequate for this purpose, in that it provides a fast, sensitive, and accurate method of identification of phenylthiohydantoins.

ACKNOWLEDGEMENTS

I gratefully acknowledge the contributions, instruction, and encouragement given to me by my major advisor, Dr. J. Michael Kehoe, under whose guidance I was given the opportunity to utilize the research facilities at the Northeastern Ohio Universities College of Medicine. I would also like to thank graduate committee members Dr. Marvin Lukin and Dr. Elmer Foldvary for their support and criticisms of the manuscript.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Definition
ATZ	Anilinothiazolinone
BSA	N,O-bis (trimethylsilyl) acetamide
BuCl	Butyl chloride
CN	Cyanopropylsilane
DMAA	Dimethylallylamine buffer
GLC	Gas liquid chromatography
HCl	Hydrochloric acid
HFBA	Heptafluorobutyric acid
HPLC	High pressure liquid chromatography
I.D.	Internal diameter
N-terminal	Amino terminal end of a protein
PITC	Phenylisothiocyanate
PTC	Phenylthiocarbamyl
PTH	Phenylthiohydantoin
RP-8	Reversed phase column, octyl groups attached to silica core

(See Appendix of 3 letter abbreviations and 1 letter codes of amino acids)

LIST OF SYMBOLS

Symbol	Definition	Units
K'	Capacity Factor	None
R	Resolution	None
t_o	Retention time of an unretained compound	Min.
t_r	Retention time of a given peak	Min.
t_2, t_1	Retention distances of peaks 1,2	cm
w_1,w_2	Peak base widths of peaks 1,2	cm

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Researchers in widely diverse areas of the medical sciences and biochemistry have developed interests in knowing the amino acid sequence of a protein. This knowledge of the primary sequence can greatly enhance the information obtained about the biochemical and structure/function relationships of the protein in question. Sequencing may be accomplished in several ways, including X-ray diffraction or mass spectrometry. However, the most popular method is chemical via the Edman degradation, using the automated protein sequencer. Automated protein sequencing may be separated into three distinct steps. First, the cleavage of the initial amino acid from the intact protein via the Edman process (Edman, 1950), producing the anilinothiazolinone (ATZ) derivative of the N-terminal amino acid. The manual conversion of the unstable ATZ derivative to the phenylthiohydantoin (PTH) derivative is the second step of the sequencing process. Finally, the PTH amino acid is analyzed and quantitated. It is this third and final step that is the primary subject of this study: the development of a sensitive and unambiguous method of identification of PTH amino acids by high pressure liquid chromatography.

In order to develop a useful method, one must remain within the following restraints so that the result is applicable to automated protein sequencing. The time of PTH

identification is of extreme importance; that is, the analysis of the results should keep in step with the time sequence of the automated cleavage so that efficiency is increased and loss of expensive sequencer reagents is minimized. In order to achieve this, a total chromatogram of 30 minutes or less per sample is required.

The cost of analysis is another important consideration. The investigator must compare the costs of several eluting reagents and the differences in the chromatogram patterns produced by these reagents against the desired results. Likewise, columns must be investigated in terms of initial cost, lifetime, and effectiveness of separation. Ideally, the chromatogram that does not compromise sensitivity, efficiency, or time of analysis at the lowest possible cost is chosen.

High pressure liquid chromatography as an analytical tool is relatively new. Giddings was one of the earliest investigators to develop this method of analysis. In the early 1960's, he demonstrated that liquid chromatography could be based on the same theory that was already developed for gas-liquid chromatography (Giddings, 1965). During the last part of that decade, the first high performance liquid chromatographs were described by Kirkland (1969), Huber (1969), and Horvath et. al. (1967).

HPLC as an analytical tool has several advantages. Because its column system operates under high pressures which in turn affect flow, HPLC can offer analysis times as short

as or better than GLC (5-30 minutes). Moreover, some compounds, such as thermally labile compounds or highly polar compounds, that either cannot be analyzed at the high temperatures of GLC or need to be derivatized first, are much more suited to analysis by HPLC. In addition, HPLC can be an extremely sensitive method, depending upon the detector, as well as a reliably reproducible method. Accurate quantitation by external or internal standards is possible. Because of these distinct advantages, HPLC has had a positive impact in many areas, including pharmaceutical, environmental, forensic, and clinical sciences, as well as in research. Therefore, when developing a new method of analysis for PTH amino acids, in which time of analysis, sensitivity, and accuracy are crucial factors, high pressure liquid chromatography appears at this time to be the method of choice.

Historically, however, other methods of analysis of PTHs have been demonstrated. Pehr Edman (1950), used paper chromatography to identify the cleaved phenylthiocarbamyl (PTC) amino acids. Since that time, several additional methods of amino acid identification have been investigated. A classic method for PTH amino acid identification is thin layer chromatography. Summers and coworkers (1973) reported detection of sub-nanomole amounts of PTH amino acids using a fluorescent indicator on polyamide sheets. Kulbe (1974) improved upon this technique using a two dimensional separation on double-faced polyamide sheets, showing the complete identification in about 30 minutes. Thin-layer chromatography,

although relatively sensitive, is not without its disadvantages. Contaminants often contribute to substantial interference and quantitation is not accurate.

It is also possible to hydrolyze thiazolinones back to their parent amino acids before analysis with an amino acid analyzer. Smithies and coworkers (1971) reported two methods of hydrolysis: an acid hydrolysis and an alkaline hydrolysis. Using these methods, all amino acids can be identified and quantitated, with the exception of cysteine, that is not distinguishable from serine without some further manipulation. Although most residues may be identified with a high degree of sensitivity, the time required for hydrolysis is this method's most serious drawback. Most amino acid analyzers require over 1 hour for a complete amino acid profile. In addition, the acid hydrolysis itself can require 20 hours and the alkaline hydrolysis 35 hours. Due to the large amount of time consumed in preparing the residue for identification, it is impossible to keep pace with the automatic sequencer.

A few researchers have also promoted the use of various indirect methods of identification: that is, a portion of the sample protein is extracted at the end of each degradation cycle and analyzed. C.H.W. Hirs et. al. (1960) wrote of a "subtractive" method which involved a hydrolysis and quantitative amino acid analysis at the completion of each cycle of the Edman process. Gray and Hartley (1963) reported the use of the dansyl technique to do N-terminal

studies at the completion of each cycle. Indirect methods of analysis have some serious drawbacks: the amount of sample needed is large, the inconvenience of extracting sample at the end of each cycle makes automation less advantageous, the methods are only practical with small peptides, and the results are not easily quantitated.

It has also been reported that chemical ionization mass spectrometry may be used to effectively identify phenylthiohydantoins (Fales et. al., 1971). It is possible to interface the mass spectrometer to the sequencer, which eliminates the need for conversion (Lovins et. al., 1972). However, although sensitivity is high and ambiguity low in this case, the costs of analysis are prohibitive for most laboratories.

A limited number of researchers also use radioisotopes to aid in phenylthiohydantoin identification. Bridgen (1975) reported high levels of sensitivity using an automated solid phase method featuring [^{35}S] phenylisothiocyanate. The radioactively labeled derivatives were then identified by comparison of autoradiographs of thin layer chromatography plates with non-radioactive internal marker phenylthiohydantoins. The three day incubation period before the autoradiograph can be developed is a serious hindrance when the investigator is interested in immediate feedback to decide whether or not to continue with the sequencing.

When sequencing a protein that is isolated from live, metabolically active tissue, the fresh tissue may be first

incubated with a single tritium labeled amino acid. The protein is then extracted from the tissue and sequenced. A surge in the radioactivity at any single step would indicate that the amino acid at that step was the one with which the protein was labeled. This method was described by Huang et. al (1975) in the sequence analysis of proparathyroid hormone. Although again a sensitive method, the time and tediousness of incubating each sample with potentially all 20 amino acids individually, as well as the costs of analysis, present an unfavorable situation.

Until recently, probably the most widely used method of PTH amino acid identification was gas-liquid chromatography (GLC). Pisano and Bronzert (1969) described a gas chromatographic procedure that quantitates and identifies all amino acids except arginine. A two column system was used, so that nonpolar PTHs are chromatographed with best results under different conditions than the relatively nonvolatile, polar compounds. A third group of amino acids with polar side chains (aspartic, glutamic, and cysteic acids) were further derivatized with the addition of N,O-bis(trimethylsilyl)acetamide (BSA). It has been noted by other authors that trimethylsilylation promotes alteration of retention behavior and better peak shapes (Harman et. al., 1968), and that serine and threonine are too labile to be accurately quantitated under these conditions (Hermodson et. al., 1972). Advances have been made over the original GLC procedures, including improved stationary phases, use of helium as a carrier gas,

and high temperature conditioning (Pisano et. al., 1972). However, quantitation of some derivatives is still not acceptably reproducible and PTH-arginine cannot be identified at all. The need to trimethylsilylate some PTHs involves a second 50 minute chromatograph which increases time and costs of identification.

As early as 1973, reports were published recommending the use of high pressure liquid chromatography for identification and quantitation of PTH amino acids. Zimmerman and Pisano (1973) realized the potential of using HPLC, reporting that it was 10 to 20 times more sensitive than GLC. In this early report, a reverse phase C-18 column was used in a gradient elution scheme involving 0.01 M sodium acetate buffer, pH 3.8, and acetonitrile. As with most pilot projects, resolution was not good. Separation of all phenylthiohydantoins except PTH-His and PTH-Arg was reported two years later, using a Hexane/methanol/propanol gradient system and a 40 minute run (Matthews et. al., 1975). (For 3 letter abbreviations of amino acids, see Appendix A.) In this study, some peaks coeluted, but the authors used the fact that some amino acids (e.g., Thr, Lys) give doublet peaks to resolve the coelution problem. According to a study published the following year, HPLC has several advantages over TLC or GLC for PTH amino acid analysis: (1) better identification is possible due to increased sensitivity, except for Met/Val, (2) sample can be recovered, intact, for further analysis, and (3) it is the most reliable method of identification per dollar spent (Downing and Mann, 1976).

Zimmerman et. al.(1976) reported the use of better column packings, e.g., Zorbax ODS (Dupont) and Permaphase ETH (Dupont), which improved resolution and sensitivity of detection of phenylthiohydantoins over previous reports. Both columns were operated simultaneously and the residues that were not separated by one column were easily resolved with its complimentary column in under 20 minutes. The water soluble residues, PTH-His and PTK-Arg, were analyzed separately.

Since 1976, several reports of a single column method of analysis have been published, using a variety of different solvents, temperatures, and columns. Zimmerman et. al.(1977) recommended a shorter, wider Zorbax ODS column at 62°C, using a gradient elution with 0.01 M sodium acetate buffer, pH 4.5, and acetonitrile. Under these conditions, all 30 amino acids were eluted in less than 20 minutes. However, serine and glutamine coeluted. Using the same column, 0.02 M sodium acetate buffer, pH 5.0, and acetonitrile, it was reported that all amino acids can be eluted and distinguished in 26 minutes (Hunkapillar and Hood, 1978). It was further noted in this report that threonine is better detected at 313 nm as dehydrothreonine, rather than at 254 nm, which demands the use of either a second detector or a variable wavelength detector. Margolies and Brauer (1978) used a Microbondapak C_{18} column (Waters Associates) to separate phenylthiohydantoins. A gradient elution using two solvents: (1) 10% acetonitrile/ 0.01 M sodium acetate, pH 4.0, and (2) 90% acetonitrile/ 0.01 M sodium acetate, pH 4.0, was effective for PTH identification. However, in several cases, including the more hydrophilic

residues, which elute at the beginning of the run at the same elution time, and the pairs Pro/Met and Trp/Phe, which coelute, a second run was needed for identification. In addition, the PTH-amino acids from the aqueous phase required additional time. A reversed phase C_8 column was also reported to be superior, using an isocratic elution of 25% acetonitrile, 75% 0.01 M sodium acetate, pH 4.6 (Abrahamsson et. al., 1978). The 75 minutes that is required for elution under these conditions is much too long to be compatible with the automatic sequencer.

Methanol, rather than acetonitrile, has also been used as the strong solvent in a HPLC gradient elution scheme. According to one report, a C_{18} Bondapak column at room temperature could be used with reasonable separation results in 20 minutes (Brown et. al., 1978). The solvents used were the following: (1) 1 liter 10% methanol in water, 2.5 ml glacial acetic acid, 50 µl acetone, adjusted to pH 4.0 with sodium hydroxide, and (2) 1 liter 90% methanol, 0.25 ml glacial acetic acid. Under these conditions and a 15 minute linear gradient from 5% solvent (2) to 45% solvent (2), most amino acids were distinguishable except Val and Met, which coeluted, and Leu, Phe, and Ile, which were poorly separated. The use of methanol rather than acetonitrile is favorable for several reasons: (1) it is less toxic than acetonitrile, (2) it is 10 times less expensive than acetonitrile, and (3) Trp is completely isolated (Zeeuws and Strosberg, 1978).

The latest report suggests the use of a pH 5.4 sodium

acetate buffer with a 17:3 mixture of methanol and acetonitrile (Johnson et. al., 1979). These authors also suggested the use of cyanopropylsilane (CN) columns, due to greater column lifetime, higher sensitivity, increased column to column reproducibility, and more reliable resolution of Arg and His. A greater than 90% resolution of 19 phenylthiohydantoins in 35 minutes was reported.

In summary, several analytical methods have been employed historically to analyze phenylthiohydantoins. However, for either reasons of cost, time of analysis, or sensitivity, these methods of analysis are not ideal especially to be complementary with the sophistication of the automatic sequencer. High pressure liquid chromatography shows promise of being the current method of choice. However, because it is a very young procedure, an optimal program for the elution of PTH amino acids needs to be developed. The sequencing process can be no better than its detection method; a poor method of analysis may rob the automatic sequencer of its effectiveness in terms of time and sensitivity. Therefore, it is of considerable importance when engaging in automated sequencing to have the best possible means of identification and quantitation. It was the goal of this study to develop such a method, also keeping in mind the restraints of cost and time of analysis, and limitations of the instrumentation available. A number of variables involved in optimizing the separation and identification of PTH derivatives by high pressure liquid chromatography have been examined in the present study.

CHAPTER II

INSTRUMENTATION

High Pressure Liquid Chromatograph

In general, the high pressure liquid chromatograph is of simple design, comprised of an injection system, high pressure pump, column, detector, and recorder. In addition to these basic components, other options such as integrators, in-line filters, flow dampeners, an oven (to control temperature), and solvent reservoirs may be added to increase the efficiency of operation (Pryde and Gilbert, 1979).

There are two primary varieties of injection available on liquid chromatographs: the syringe injection system and injection valves. The instrument available for this study was a Hewlett Packard High Pressure Liquid Chromatograph, Model 1084B, which is equipped with an automatic sampling system that will accomodate a maximum of sixty 2 ml vials. In addition, the automatic sampler combines with a variable volume syringe injector, adjustable for sampling volumes from 1 µliter to 200 µliters, to provide up to 99 automatic consecutive injections per vial.

Several pumps have also been used in a high pressure chromatographic system a few of them being a direct gas pressure system, a constant pressure pump, a reciprocating pump, or a syringe pump (Pryde and Gilbert, 1979). The 1084 fea-

tures two single piston diaphragm pumps with a common drive. The pumps include a damping system, to prevent pulses of flow, and a feedback flow control. The system operates up to a maximum pressure of 400 bar (5800 lb/in^2) and has a 10 ml per minute maximum flow rate per pump. Dual pumps provide a multilinear, continuously variable gradient system that is time programmable in 0.01 minute increments. Hewlett Packard guarantees a flow regulation stability of $\pm$ 1%, and combinations of flow and gradient programming are possible.

Depending upon the mode of separation and its use, the column varies accordingly with length, width, and column packing used. There are several modes of separation that can be used with HPLC: adsorption chromatography, liquid-liquid partition chromatography, reversed phase chromatography, ion exchange chromatography, ion pair partition chromatography, exclusion chromatography, and affinity chromatography (Pryde and Gilbert, 1979). Reversed phase chromatography, in which the column packing consists of a hydrophobic carbon chain bonded to a silica core was chosen for this study. The mobile phase (e.g., a water/methanol mixture) is passed over this support. Polar molecules that have little affinity for the support will elute more quickly than non-polar molecules. The detailed mechanism of the action of reversed phase chromatography has not yet been satisfactorily explained. In the 1084B the column compartment will accommodate up to twelve 25 cm columns and has an oven which controls temperature 5^{o} above ambient to 100°C.

The detector system of the 1084B is comprised of a fixed wavelength (254 nanometers) ultraviolet absorption detector. The use of a variable wavelength detector or an external detector of any sort (e.g., refractive index, fluorometric, electrochemical, or electron capture detector) is possible with this liquid chromatograph, but was not available for this study.

The two solvent reservoirs may be independently heated to a maximum of 100°C, degassed with a vacuum pump, or purged with an inert gas. The 1084B is a processor based, keyboard controlled liquid chromatograph. Rate of flow, solvent gradient, integrator parameters, chart speed, method of quantitation, and external contact changes can be implemented automatically on a time basis during an analysis. In addition, flow rates, percent of solvent "B", column and pump pressures, oven and solvent temperatures , and detector signals can be plotted on the thermal-writer printer/plotter.

The operator may choose one of four calculation methods for the final report: area percent, normalization, external standard, or internal standard. Several integrate functions, including three methods of baseline resetting, peak skimming, peak slicing, tangent skimming, and area bunching may also be chosen at the discretion of the operator.

A box diagram of the component parts of the flow pattern of the Hewlett-Packard HPLC may be found in Figure 1.

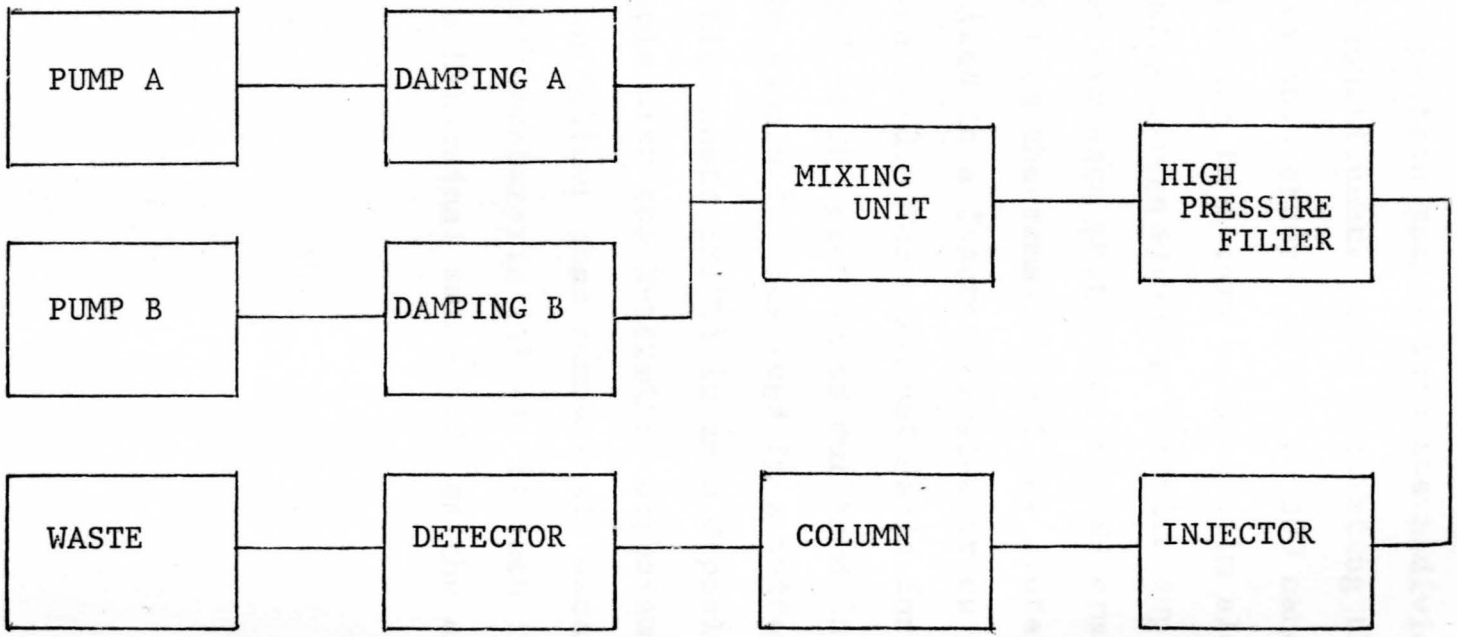


Fig. 1. Flow pattern of Hewlett Packard High Pressure Liquid Chromatograph, Model 1084B

Automated Protein Sequencer

The Beckman 890C Sequencer was used to automatically degrade the protein sample into its individual sequential amino acid constituents, thus generating the samples for analysis for this study. A 200 to 300 nanomole protein sample was dried in a thin uniform film along the inside of a magnetically-driven spinning reaction cup, subjected to a series of reagents that cause the N-terminal amino acid to be cleaved from the remainder of the protein or peptide, and then deposited in a fraction collector as the ATZ derivative of the amino acid. The chemical basis for this "Edman degradation" of the protein is outlined in Figure 2. After the protein sample is dissolved in a buffer and dried, phenylisothiocyanate (PITC) is then deposited in the reaction cup to couple with the N-terminal alpha-amino group. A benzene wash follows that removes all excess PITC. A strong acid, heptafluorobutyric (HFBA), is then introduced which cleaves the N-terminal amino acid as the ATZ derivative.

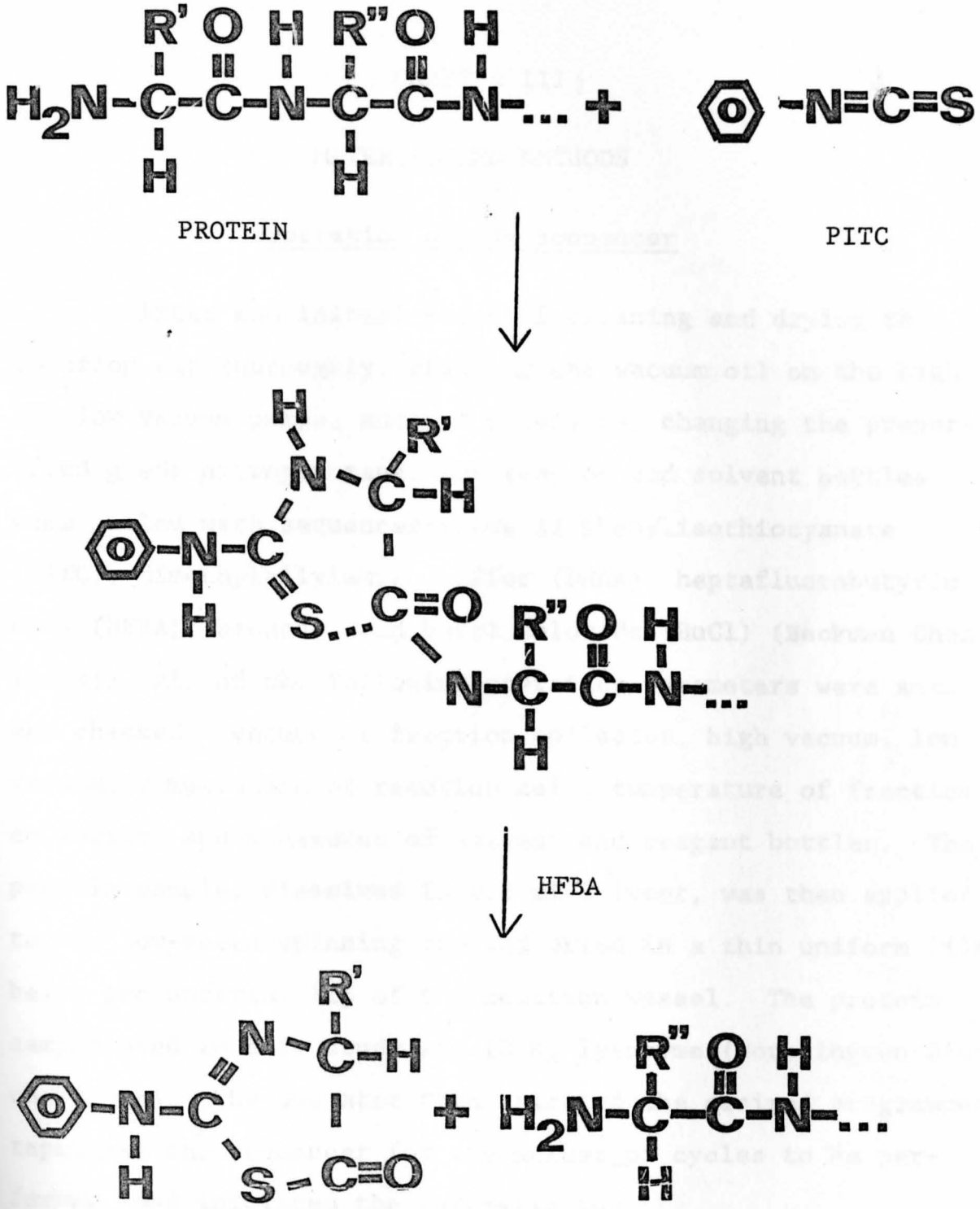


Fig. 2: Edman Degradation

CHAPTER III

MATERIAL AND METHODS

Operation of the Sequencer

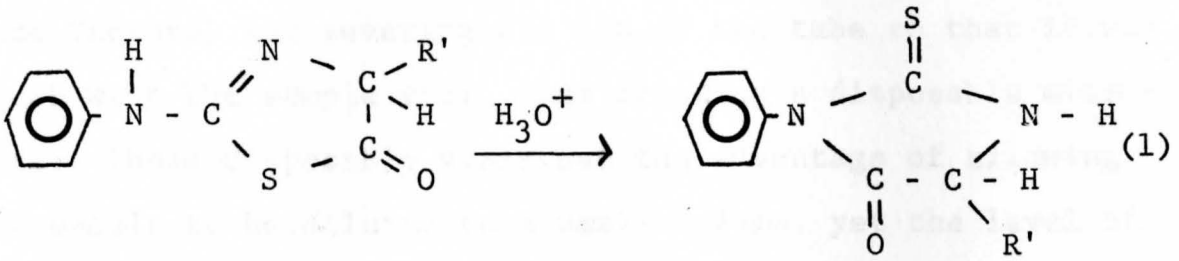
After the initial steps of cleaning and drying the reaction cup thoroughly, changing the vacuum oil on the high and low vacuum pumps, and, if necessary, changing the prepurified grade nitrogen tank, the reagent and solvent bottles were filled with sequencer-grade 5% phenylisothiocyanate (PITC), dimethylallylamine buffer (DMAA), heptafluorobutyric acid (HFBA), benzene, and butyl chloride (BuCl) (Beckman Chemicals). All of the following operating parameters were set and checked: vacuum of fraction collector, high vacuum, low vacuum, temperature of reaction cell, temperature of fraction collector, and pressures of solvent and reagent bottles. The protein sample, dissolved in 0.5 ml solvent, was then applied to the low-speed spinning cup and dried in a thin uniform film below the undercut lip of the reaction vessel. The protein sample used in this study was 10 mg lysozyme (Worthington Biochemicals). The operator then selected the desired programmed tape, set the sequencer for the number of cycles to be performed, and initiated the automatic run.

At the end of the sequence, the vacuum on the fraction collector was broken and fraction tubes sequentially containing the N-terminal ATZ amino acids of the protein sample were re-

moved, converted, and analyzed for identification.

Conversion of Anilinothiazolinone Derivative Amino Acids Into Phenylthiohydantoin Amino Acids

Amino acids were removed from the sequencer in the form of the relatively unstable anilinothiazolinone (ATZ), which must be converted to the phenylthiohydantoin (PTH) derivative before identification and quantitation.



ATZ PTH (1)

The fraction tube was removed from the sequencer and the contents dried under nitrogen before 0.2 ml of 1.0 N hydrochloric acid (HCl) (Fisher Scientific) was added. The tube was then stoppered with an inert silicone stopper, mixed on a vortex mixer, and incubated at 80°C for 10 minutes. The dried residue was then extracted twice with 0.7 ml acetone free ethyl acetate (Beckman Chemicals) and reduced to dryness at 50°C using a prepurified grade of nitrogen. This product was then ready for analysis.

Three amino acids, arginine, histidine, and cysteic acid, remain in the aqueous phase throughout the extraction. In these cases, if no positive identification was made in the organic phase, the aqueous phase was diluted with 1 ml deionized water, lyophilized, dried, dissolved in methanol, and then chromatographed for identification.

Sample Preparation

After cleavage from the intact protein and conversion to the PTH derivative, all dried samples were first diluted in a standard volume (100 µliter) of HPLC grade methanol (Fisher Scientific). Dilution to such a small volume was possible due to the insertion of a size 400 µliter polypropylene centrifuge tube into the standard 2 ml sample vial (Hewlett Packard) and severing the top of the tube so that it was flush with the sample vial, thus creating a disposable microvial. These disposable vials had the advantage of allowing the sample to be diluted to a small volume, yet the level of sample was adequate to allow contact with the variable volume injector. Otherwise, the sample would have necessarily been diluted to at least 0.5 ml before injection. A more concentrated sample is preferred because a smaller injection volume is required, erasing possibilities of overloading the column. Also, the detection system need not be greatly attenuated, eliminating a large amount of background noise and allowing greater flexibility when attempting to detect nanomole quantities. Standard 11 mm teflon lined caps were then crimped onto the disposable vials to prevent sample evaporation.

Standard PTH Amino Acid Preparation

All PTH amino acid standards were obtained from Pierce Chemical Company (Rockford, Illinois), dissolved in HPLC grade methanol (Fisher Scientific), and stored in con-

centrations of 1 nanomole per µliter at sub-zero temperatures until use. Upon thawing a 10 µliter aliquot (10 nanomole) of the standard was injected onto the column at an attenuation of $2^8 \times 10^{-4}$ absorbance units per cm.

Buffer Preparation

The 0.01 M sodium acetate buffer, pH 5.0, that was used for most of this study, was prepared in the following manner: 14.8 ml of a stock 0.2 M acetic acid (11.38 ml glacial acetic acid per liter) (Fisher Scientific) was added to 35.2 ml stock 0.2 M sodium acetate (27.2 g $NaC_2H_3O_2 \cdot 3H_2O$/liter) (Fisher Scientific) and then diluted with deionized water to 1 liter. The pH of the buffer was checked with an Orion Model 701 digital pH meter and adjusted accordingly. All buffers were kept refrigerated until used.

Other sodium acetate buffers of various pHs used in the optimum pH portion of this study were prepared likewise according to Table 1.

TABLE 1

PREPARATION OF SODIUM ACETATE BUFFERS

Vol. 0.2 M Acetic Acid	Vol. 0.2 M Sodium Acetate	pH
44.0 ml	6.0 ml	3.8
36.8 ml	13.2 ml	4.2
25.5 ml	24.5 ml	4.6
14.8 ml	35.2 ml	5.0
8.8 ml	41.2 ml	5.4
2.3 ml	47.7 ml	5.8

Operation of the High Pressure Liquid Chromatograph

Before use, all buffers were passed through a 0.22 micron filter to prevent blockages in the small-bore stainless steel tubing due to particulate matter. This procedure also helped to decrease possibilities of bacterial growth. Solvent A was a 0.01 M sodium acetate buffer which was degassed for 30 minutes on the instrument at 60°C prior to use. HPLC grade methanol prefiltered by and purchased from Fisher Scientific served as solvent B. This solvent was degassed for 30 minutes at 40°C.

After filling and degassing the solvents, the flow was activated and allowed to equilibrate for 10 to 15 minutes to insure a stable pressure level. All operating parameters, including oven temperature, % solvent B, attenuation, number of injections per vial, number of sample vials, detector signal, length of time programs, acceptable minimum and maximum pressure levels before pump shut-off, and rate of flow, were then selected using the microprocessor-based keyboard. After the initial stabilization of all parameters, the first injection was initiated.

The shut-down procedure for the high pressure liquid chromatograph included a 15 minute flush of the column with 100% methanol, followed by pump and instrument turn-off.

Column

The reversed phase column that was used for this study was a Zorbax C-8 column (Dupont), 4.6 mm. I.D. by 25 cm long. As the name implies, the packing is made up of a chain of 8 carbons bonded to a silica support. A monolayer of C-8 is assured by the use of a monofunctional silane reagent. As indicated by the manufacturer, Zorbax C-8 is useful for separations over a wide range of compounds extending from water soluble to hydrocarbon soluble. Such is the case when working with a variety of phenylthiohydantoins.

Method Development on the HPLC

The ideal elution pattern determination for mixtures of various PTH derivatives involves the study of different elution gradient systems to optimize the separation and differentiation of the various derivatives. When sequencing a protein, there are 20 amino acids that have to be identified and quantitated, in hopefully the least possible time, expense, and ambiguity, and with the greatest possible sensitivity. When trying to separate a mixture of 20 molecules in under 25 minutes, the primary problem is one of resolution. Two or more of the phenylthiohydantoins will coelute until an ideal gradient elution time program is obtained. The sequencer itself, due to the sequencing program, will produce some interfering substances that will show up on the chromatograph. Thus, a second problem: to develop a program not only in

which the PTH's elute independently of each other, but one in which they elute separately from non-PTH organic artifacts generated by the sequencer. A scouting gradient from 10% to 90% B in 20 minutes was run first. It was noted where, within the gradient, the phenylthiohydantoins eluted and the gradient was narrowed accordingly by the operator.

After an acceptable linear gradient system is found in which all PTH amino acids are eluted in a short amount of time, the operator may then change several parameters within the operating limits of the instrument in order to optimize the separation. A complete understanding of the principles of high pressure liquid chromatography plays an important role at this point of the method development, so that each run is not merely the result of trial and error, but rather an educated experiment aimed at a particular result. The operator at this point of the method should be careful to change only one parameter at a time so that a logical order of events may occur. The parameters that may be varied with the 1084B to effect the best baseline separation are % solvent B, flow, oven temperature, and pH of the weak solvent A. Since the determination of the optimum pH of the system is the second major section of this study, it will not be discussed at this time.

The 1084B may be time programmed to form any series or combinations of %B in addition to timed changes in rate of flow that may effectively separate two peaks that would co-elute in an isocratic run. The operator then uses these tools

to achieve the desired separation.

The temperature of the system also affects the retention times of the phenylthiohydantoins. Changing the temperature is less effective in separating peaks that coelute than time programming the flow system because the temperature change effects all of the samples more or less uniformly; that is, in most cases, if the operator raises the temperature of the column, the retention times of all the peaks will be reduced. The temperature of the system must, however, remain constant. For purposes of this study, the temperature of the column was kept at 40°C.

A second approach to the method development was a differential analysis of several elution systems. The effect of the initial %B in a gradient system on the resolution of the polar amino acids was determined by plotting initial %B vs. a capacity factor, sometimes called column capacity ratio, which is a measure of solute retention. This capacity factor is defined by

$$K' = \frac{t_r - t_o}{t_o} \qquad (2)$$

where t_r is the retention time of a given peak and to is the retention time of an unretained compound. Since retention times may vary with small day-to-day differences in flow, the use of K' values, which correct for flow variations due to the incorporation of the retention time of an unretained compound, is preferred. A series of different initial %B values, from 5% to 29%, were examined. Likewise, the effect of the final %B on the separation of the hydrophobic amino acids was

studied, using final %B values of 41% to 69%.

Isocratic systems at different concentrations of strong solvent were then studied, in order to ascertain at what %B level each amino acid is best eluted. Several factors were kept constant during these analyses, including the oven temperature (40°C), time of program (30 minutes), flow rate (2 ml per minute), and pH of the buffer (pH 5.0).

The resolution of the peaks in a chromatogram was then calculated in several cases to determine whether or not a better chromatogram had been developed. The resolution of two peaks is defined as

$$R = 2\left(\frac{t_2 - t_1}{W_1 + W_2}\right) \qquad (3)$$

where t_2 and t_1 are the retention distances and W_1 and W_2 are the peak base widths of the peaks in question (Fig. 3). It is important to note that t_2, t_1, W_1, and W_2 must all be in the same units (e.g., distance). A "R" value of 1.5 constitutes effective baseline separation for two Gaussian peaks.

All of the information obtained from the gradient elution studies above was combined to produce a chromatogram that displayed the optimum separation of 20 phenylthiohydantoin derivatives. After this was achieved, the operator then programmed into the microprocessor the desired integrator parameters and method of quantitation of PTH amino acids.

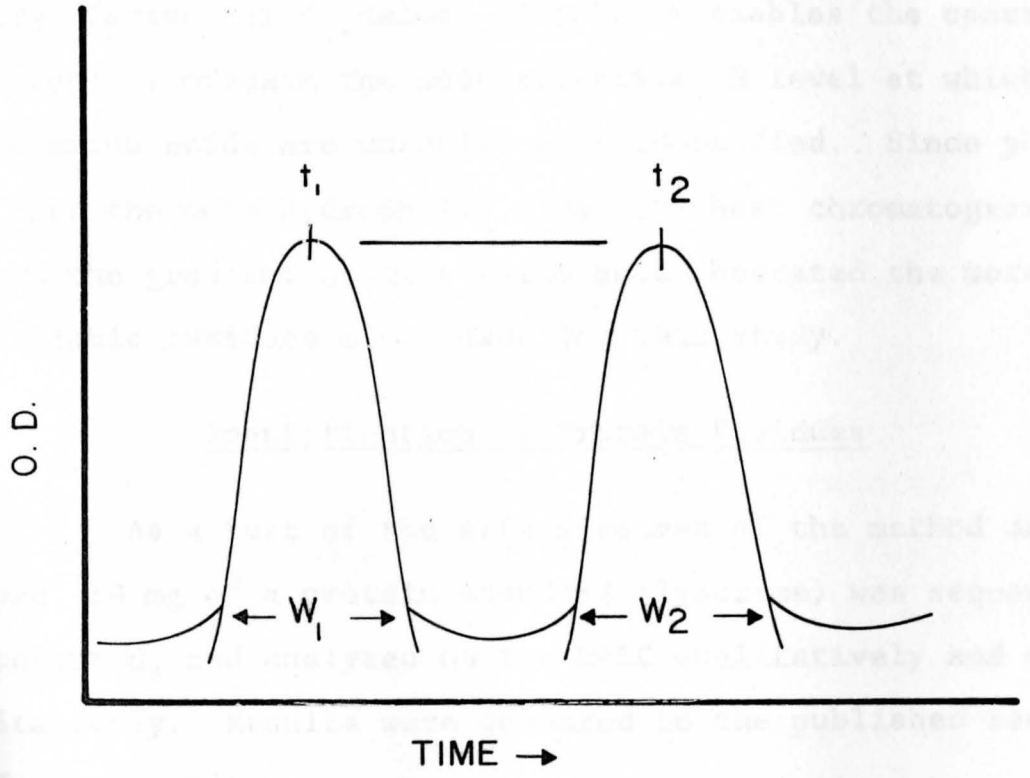


FIG. 3 RETENTION DISTANCES AND PEAK BASE WIDTHS OF TWO GAUSSIAN PEAKS

pH Studies

In order to determine the optimum pH level for maximum separation of 20 PTH amino acids, each amino acid was tested with a variety of sodium acetate buffers, differing only in pH, keeping all other factors constant. The retention times at each pH level were noted and the most desirable pH level was selected using a plot of pH level vs. the capacity factor, or K' value. This plot enables the operator to readily discern the most effective pH level at which all PTH amino acids are unequivocally identified. Since pH will affect the more hydrophilic PTHs, the best chromatogram from the gradient studies which best separated the more hydrophobic residues was chosen for this study.

Identification of Protein Residues

As a test of the effectiveness of the method developed, 10 mg of a protein standard (lysozyme) was sequenced, converted, and analyzed on the HPLC qualitatively and quantitatively. Results were compared to the published sequence of this protein.

CHAPTER IV

RESULTS

The results of this study will be developed according to the following format:

1. Results of a progressive gradient system analysis
2. Results of a differential analysis of several HPLC systems
3. Results of pH studies
4. Testing of the method developed by anaylsis of protein sequence residues

In all cases, the printout of the Hewlett Packard High Pressure Liquid Chromatograph is to be interpreted as explained in Table 2. A few points should be noted. In many printouts, solvent A and solvent B are not heated; they are, rather at room temperature. The explanation is that after the initial "degassing" procedure, in which a vacuum is pulled on the heated solvents, the heater is in most cases turned off. This has no effect on any results because even heated solvents cool in the stainless steel tubing before they reach the column. Therefore, only the oven temperature surrounding the column is critical. Another consideration is the time programming. At the end of every timed program, there is a 2 to 3 minute isocratic period in which the mobile phase composition is returned to its original state. This is to insure flow and column pressure stability and equilibration

TABLE 2

PRINTOUT OF HEWLETT PACKARD HIGH PRESSURE LIQUID CHROMATOGRAPH

PRINTOUT			INTERPRETATION
FLOW	2.00	1.96	Flow set at 2 ml/min; Reads at 1.96 ml/min.
%B	18.0	17.7	%B set at 18%B; Reads at 17.7%B
COLUMN P		243	Column pressure reads 243 bar
MAX P	400		Maximum pressure allowable before pump shut-off
MIN P	0		Minimum pressure allowable before pump shut-off
S-TEMP A	60	23	Temp. Solvent A set at 60°C; reads 23°C
S-TEMP B	40	21	Temp. Solvent B set at 40°C; reads 21°C
OVEN TEMP	40	40	Oven temp. set at 40°C; reads 40°C
UV SGNL			Detector is a UV Spectrophotometer, 254 nm
CHT SPD	1.00		Chart paper moves at 1 cm per minute
ZERO	10.0		Baseline located at 10% of full scale
ATTN 2↑	8		Attenuation at 2^8=256 x 10^{-4}absorbance units per cm
AREA REJ	5000		Peak areas less than 5000 not reported
SLP SENS	1.00		Slope sensitivity = 1
8.00 %B		18.0	Mobile phase composition = 18%B at 8 minutes
13.00 %B		40.0	Mobile phase composition = 40%B at 13 minutes
15.00 %B		50.0	Mobile phase composition = 50%B at 15 minutes
21.00 %B		50.0	Mobile phase composition = 50%B at 21 minutes
23.00 %B		18.0	Mobile phase composition = 18%B at 23 minutes
25.00 %B		18.0	Mobile phase composition = 18%B at 25 minutes
25.00 STOP			Stop program at 25 minutes

of these parameters is necessary for good reproducibility. Figure 4 illustrates the mobile phase composition as programmed in the example in Table 2.

Figure 5 illustrates a sample chromatogram in which all of the PTH-amino acid peaks are labelled appropriately. It should be noted that the numbers above the peaks are retention times in minutes, and that glitches appearing at the baseline of a peak indicate where integration of the area under the peak begins and ends. Individual PTH amino acid standards were run to confirm the naming of each peak in the mixture.

Before beginning the method development, several conditions were set as the basis of the study. A reversed phase column, C-3, was used throughout, due to its versatility and applicability to this system. A methanol/sodium acetate buffer mixture was chosen to be the mobile phase: methanol, because of its relatively low cost and toxicity, and a buffer because of its ability to maintain pH. It was also assumed at the onset of this study that a gradient elution would be necessary due to the varying polarities of phenylthiohydantions. Finally, the temperature of the column was maintained at 40°C at all times.

Progressive Gradient System Analysis

The first approach taken in the method development of the analysis of PTHs was the progressive application of several gradient systems in a logical manner to the high

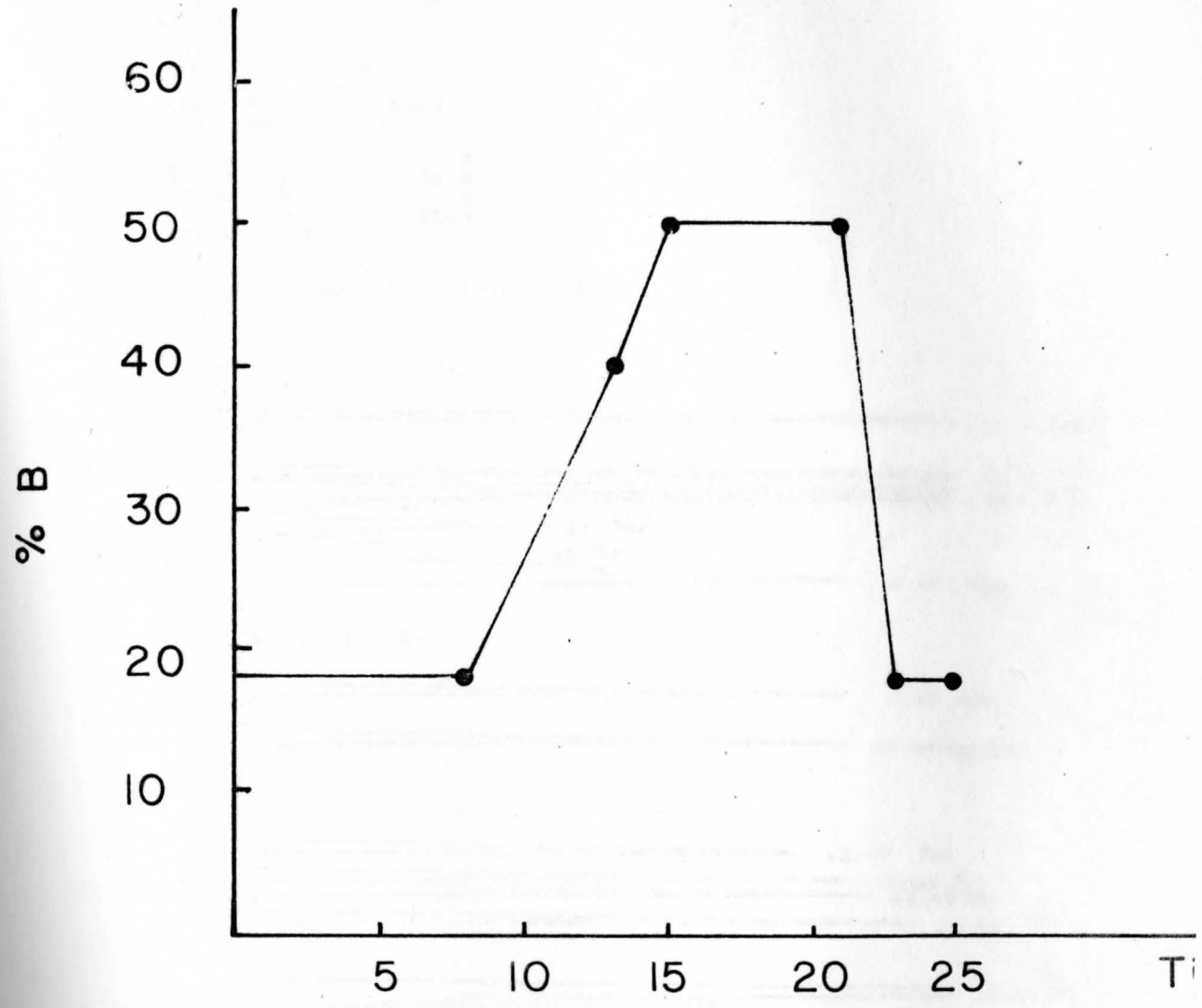


FIG. 4 MOBILE PHASE COMPOSITION
TYPICAL TIME PROGRAM

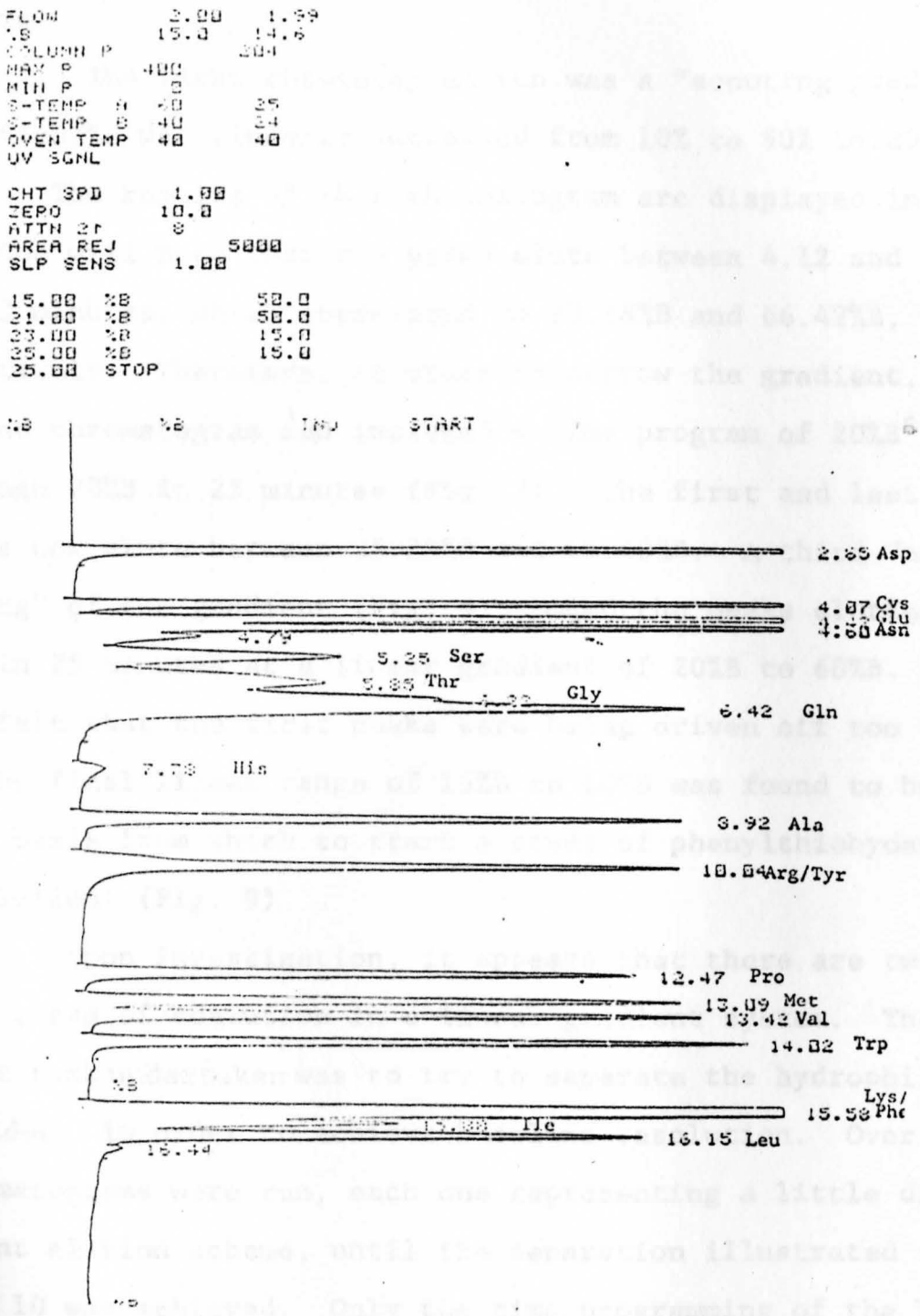


Fig. 5. Sample HPLC Chromatogram of Phenylthiohydantoins

pressure liquid chromatograph in order to separate the 20 phenylthiohydantoins according to the best possible elution scheme.

The first chromatogram run was a "scouting gradient," in which %B was linearly increased from 10% to 90% in 25 minutes. The results of this chromatogram are displayed in Fig. 6. One will note that the peaks elute between 4.12 and 17.63 minutes, which correspond to 23.18%B and 66.42%B, respectively. Therefore, in order to narrow the gradient, the second chromatogram run included a time program of 20%B through 70%B in 25 minutes (Fig. 7). The first and last peaks now elute between 25.28%B and 59.40%B. A third "narrowing" of the gradient (Fig. 8) showed the peaks eluting within 25 minutes at a linear gradient of 20%B to 60%B. It was felt that the first peaks were being driven off too fast, so the final linear range of 15%B to 60%B was found to be a good basis from which to start a study of phenylthiohydantoin resolution. (Fig. 9)

Upon investigation, it appears that there are two main areas of coelution in a linear gradient system. The first task undertaken was to try to separate the hydrophilic residues in order to achieve baseline resolution. Over 100 chromatograms were run, each one representing a little different elution scheme, until the separation illustrated in Fig. 10 was achieved. Only the time programming of the %B concentration was altered during this portion of the method development. It was found that an isocratic mobile phase composition of 18%B for 8 minutes allowed Thr/Gly/Gln to be

```
FLOW          2.00    1.98
%B           10.0    10.0
COLUMN P            199
MAX P      400
MIN P        0
S-TEMP  A   60       24
S-TEMP  B   40       22
OVEN TEMP   40       40
UV SGNL

CHT SPD       1.00
ZERO         10.0
ATTN 2↑       8
AREA REJ           5000
SLP SENS      1.00

25.00  %B            90.0
27.00  %B            10.0
30.00  %B            10.0
30.00  STOP
```

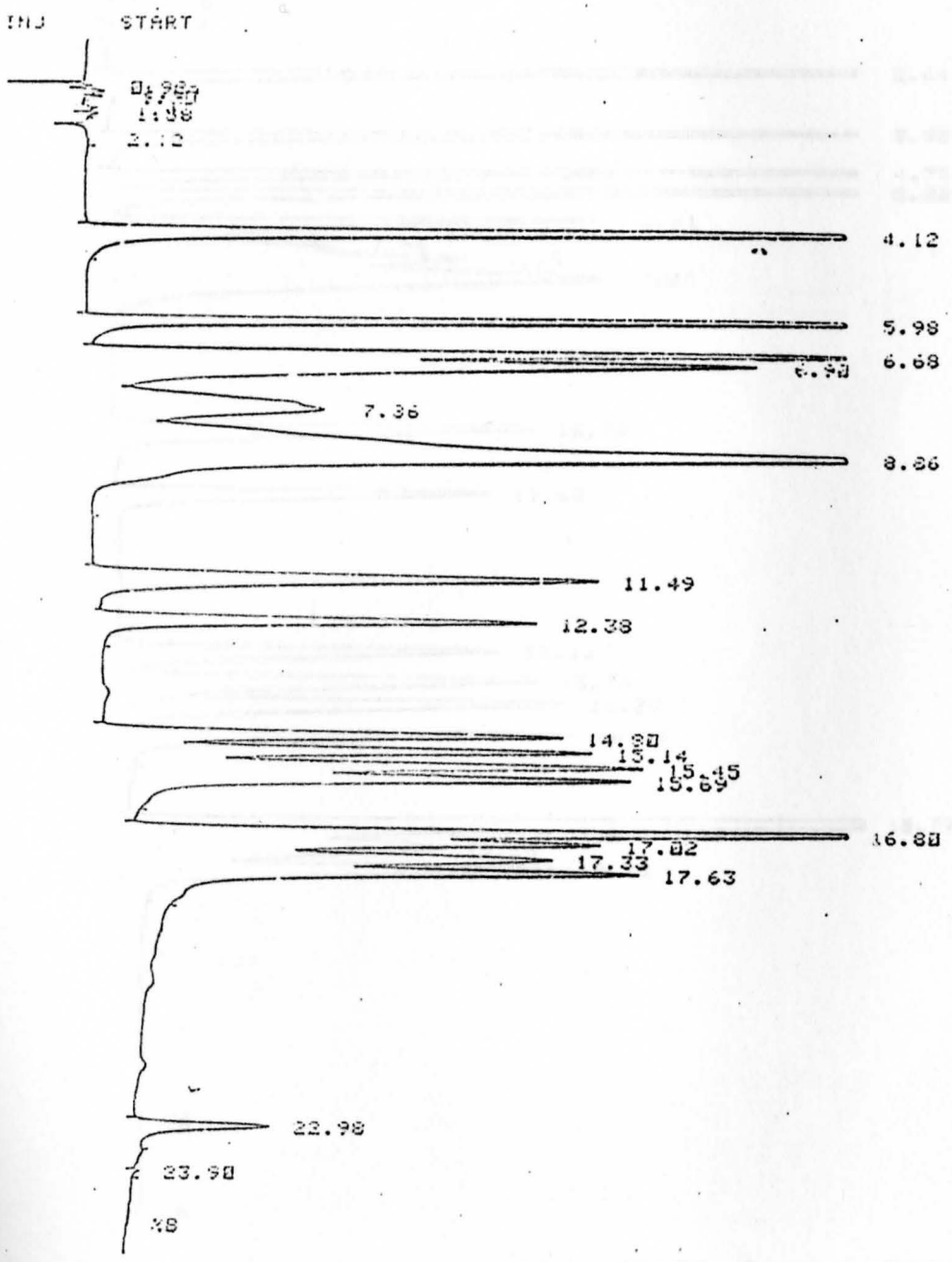


Fig. 6. Gradient Elution Profile (Scouting Gradient)

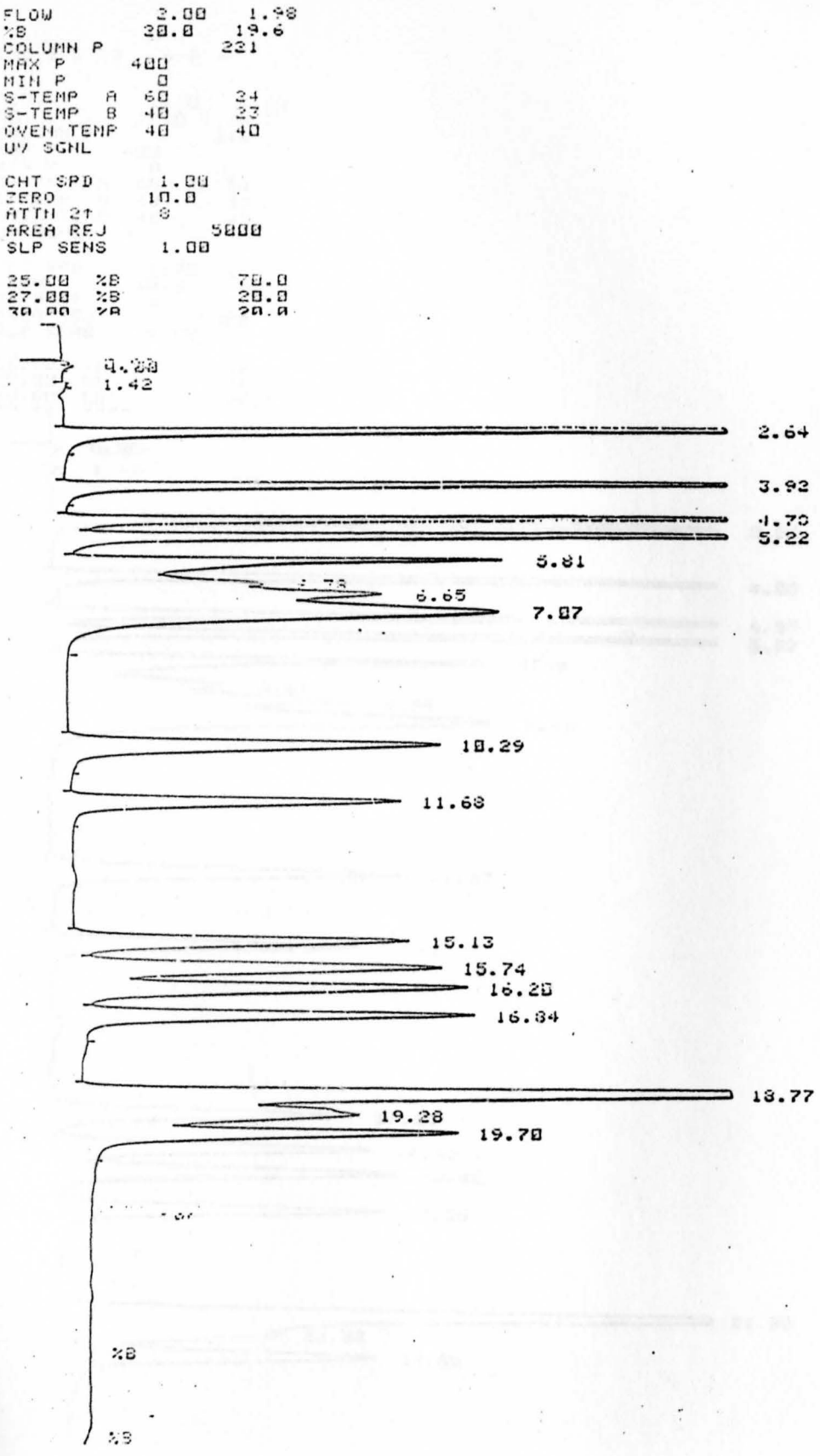


Fig. 7. Gradient Elution Profile (20%B through 70%B)

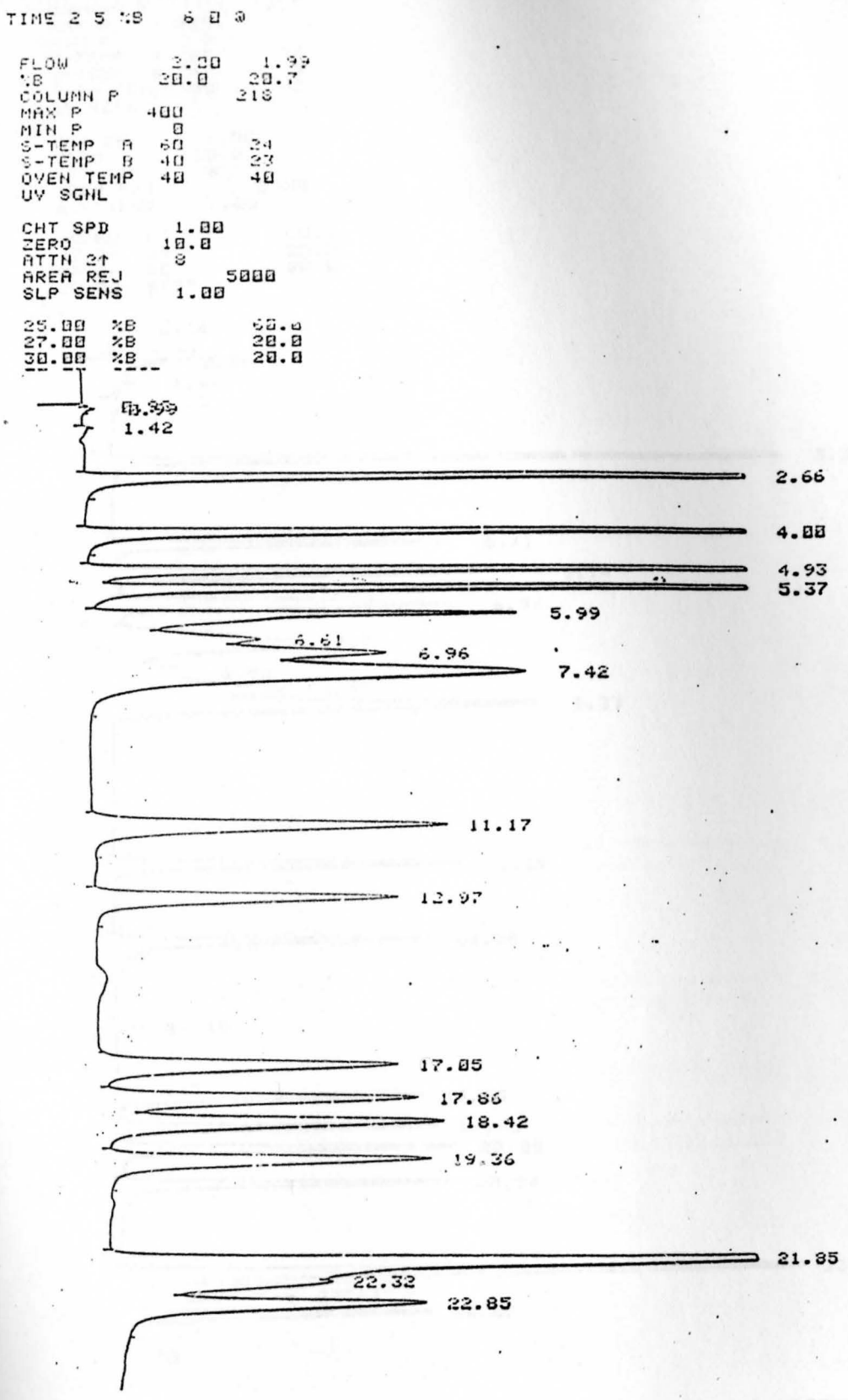


Fig. 8. Gradient Elution Profile (20%B through 60%B)

```
FLOW          2.00    1.99
%B           15.0    15.0
COLUMN P           215
MAX P       400
MIN P         0
S-TEMP   A   60       24
S-TEMP   B   40       23
OVEN TEMP    40       40
UV SGNL

CHT SPD       1.00
ZERO         10.0
ATTN 2↑       8
AREA REJ           5000
SLP SENS      1.00

25.00  %B            60.0
27.00  %B            20.0
30.00  %B            20.0
30.00  STOP
```

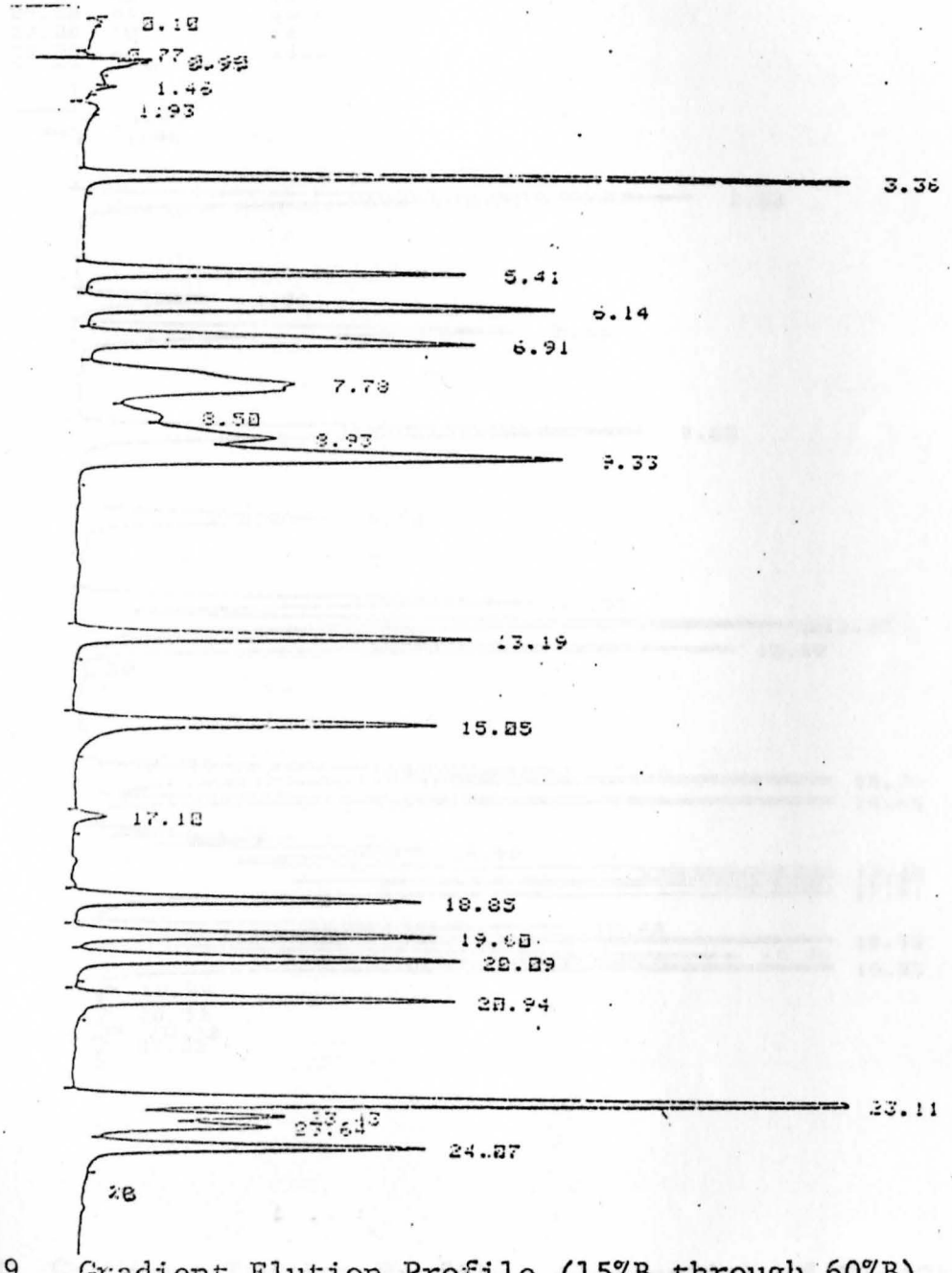


Fig. 9. Gradient Elution Profile (15%B through 60%B)

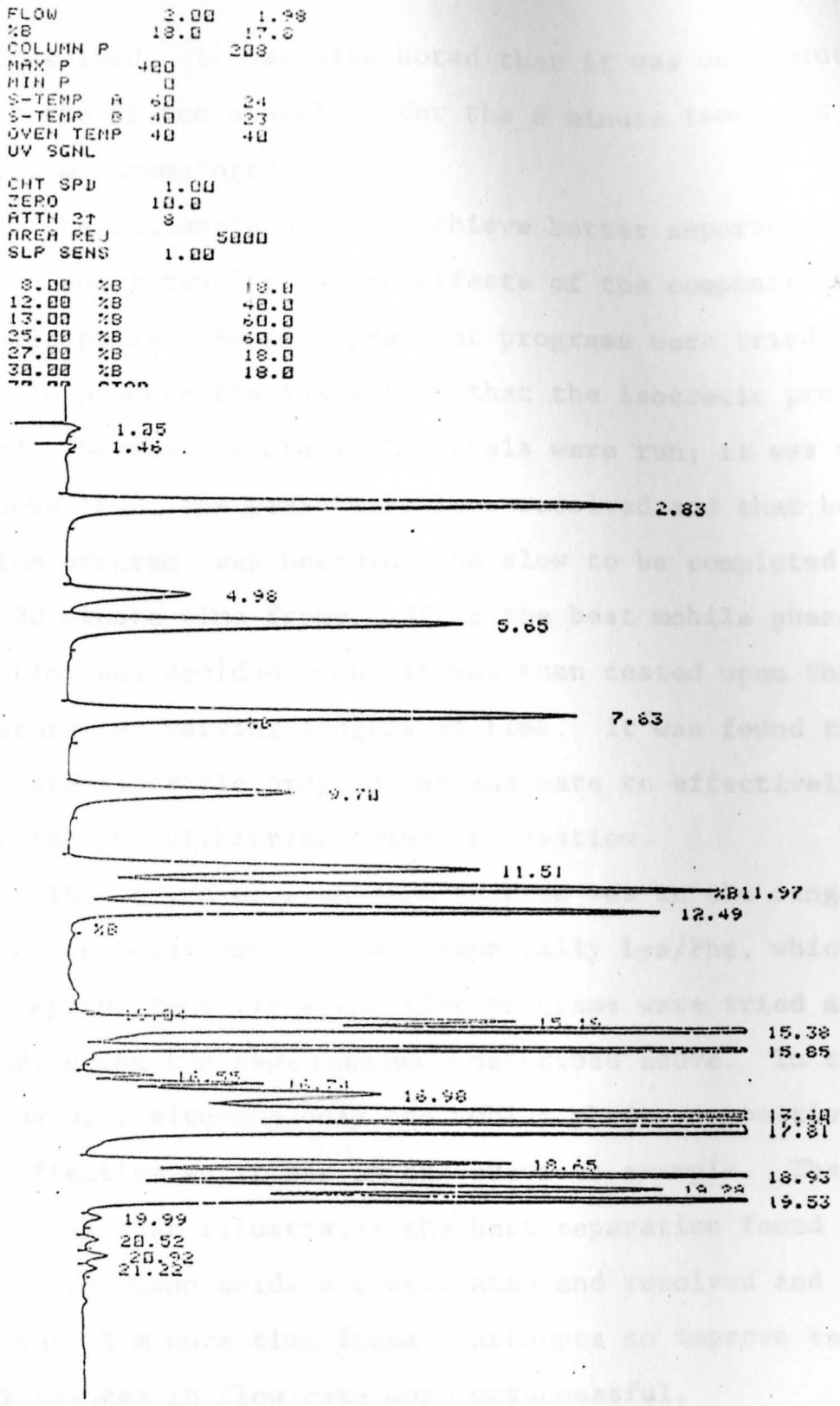


Fig. 10 Gradient Elution Profile. (Separation of Thr/Gly/Gln)

nicely resolved. It was also noted that it was detrimental to raise the %B too sharply after the 8 minute isocratic portion of the chromatogram.

The rationale used to achieve better separation involved an understanding of the effects of the composition of the mobile phase. Several gradient programs were tried, but none could produce the resolution that the isocratic program produced. Several isocratic %B levels were run; it was noted that above 18%B, the peaks were less resolved and that below 18%B, the program was becoming too slow to be completed within the 30 minute time frame. After the best mobile phase composition was decided upon, it was then tested upon the 20 PTH mixture for varying lengths of time. It was found that an 8 minute isocratic program was adequate to effectively resolve the phenylthiohydantoins in question.

The second problem area tackled was in the range of the more non-polar amino acids, expecially Lys/Phe, which co-elute. Again, many different time programs were tried and examined, using the same rationale described above. In this case, however, altering only the mobile phase composition was not as effective as it was in the previous example. The chromatogram that illustrates the best separation found is Fig. 11. All PTH amino acids are separated and resolved and elute within the 25 minute time frame. Attempts to improve resolution by changes in flow rate were unsuccessful.

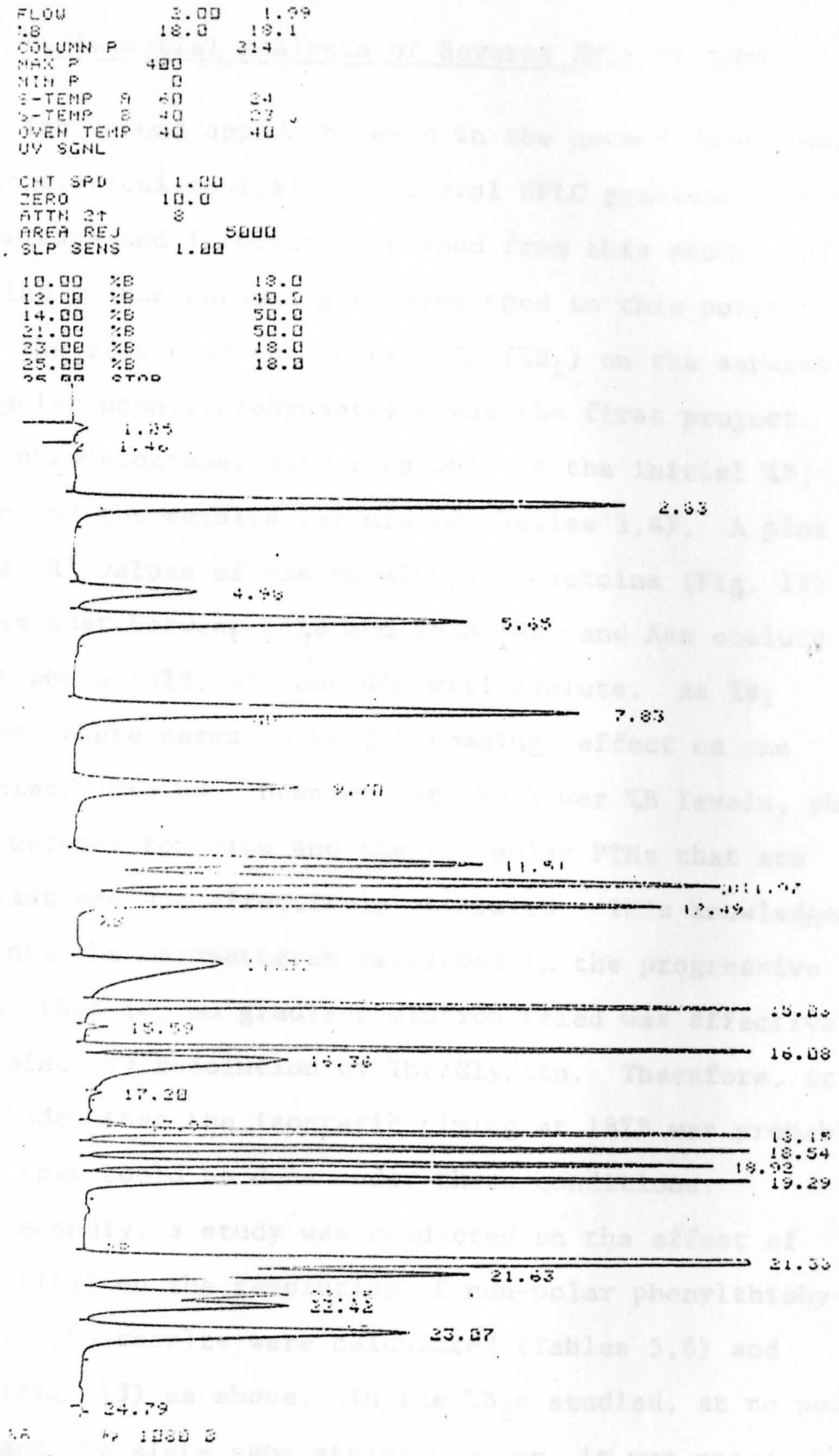


Fig. 11. Gradient Elution Profile (Separation of 20 PTHs)

Differential Analysis of Several HPLC Systems

The second approach taken in the method development was a statistical analysis of several HPLC gradient systems. The knowledge and information gained from this study could be applied to the chromatogram developed to this point.

The effect of the initial %B ($\%B_i$) on the separation of the polar phenylthiohydantoins was the first project. Several chromatograms, differing only in the initial %B, were run and the results calculated (Tables 3,4). A plot of $\%B_i$ vs. K' values of the phenlythiohydantoins (Fig. 12) indicates that between 11%B and 15%B, Glu and Asn coelute, and that above 15%B, Gly and Gln will coelute. As $\%B_i$ increases, there seems to be a "crowding" effect on the phenylthiohydantoins. However, at the lower %B levels, the program becomes too slow and the non-polar PTHs that are eluted last are not effectively separated. This knowledge complements the chromatogram developed in the progressive analysis; that is, no gradient elution tried was effective in the satisfactory resolution of Thr/Gly/Gln. Therefore, it was concluded that the isocratic timing at 18%B was probably the best that could be done under these conditions.

Secondly, a study was conducted on the effect of final %B ($\%B_f$) on the resolution of non-polar phenylthiohydantoins. The results were calculated (Tables 5,6) and plotted (Fig. 13) as above. In the $\%B_f$s studied, at no point did Lys and Phe elute separately; however, it was noted that there was a "tendency" to separate with increasing $\%B_f$.

TABLE 3
RETENTION TIMES OF HYDROPHILIC PHENYLTHIOHYDANTOINS WITH VARIATIONS IN INITIAL MOBILE PHASE COMPOSITION

	Retention Times (Minutes)								
$\%B_i$*	Asp	Cys	Glu	Asn	Ser	Thr	Gly	Gln	T_o**
5%	6.03	8.81	9.67	10.46	11.10	12.41	12.91	13.31	0.94
7%	5.26	7.96	8.87	9.27	10.26	11.60	12.20	12.50	1.03
9%	4.95	7.69	8.51	8.86	9.84	11.13	11.79	12.27	1.01
11%	4.44	7.04	7.96	7.96	9.16	10.41	11.16	11.55	1.00
13%	3.98	6.44	7.44	7.44	8.49	9.67	10.48	10.81	1.00
15%	3.60	5.76	6.82	6.92	7.79	8.92	9.75	9.75	0.99
17%	3.17	5.08	6.18	6.39	7.08	8.07	9.00	9.00	0.96
19%	2.95	4.59	5.63	5.90	6.65	7.48	8.40	8.40	0.97
21%	2.69	4.07	5.03	5.38	6.11	6.71	7.46	7.46	0.93
23%	2.31	3.40	4.30	4.75	5.51	5.91	6.78	6.78	0.93
25%	2.14	2.99	3.81	4.30	5.02	4.48	6.16	6.16	0.93
27%	2.01	2.65	3.38	3.90	4.56	5.03	5.56	5.56	0.91
29%	1.91	2.39	3.02	3.56	4.15	4.60	4.99	4.99	0.90

* Initial Mobile Phase Composition
** Retention time of an unretained compound

TABLE 4

CAPACITY FACTORS OF HYDROPHILIC PHENYLTHIOHYDANTOINS WITH VARIATIONS IN INITIAL MOBILE PHASE COMPOSITION

	Capacity Factors (K')							
$\%B_i$*	Asp	Cys	Glu	Asn	Ser	Thr	Gly	Gln
5%	5.41	8.38	9.29	10.13	10.81	12.20	12.73	13.16
7%	4.11	6.73	7.61	8.00	8.06	10.26	10.85	11.14
9%	3.90	6.61	7.43	7.77	8.74	10.02	10.67	11.15
11%	3.44	6.04	6.96	6.96	8.16	9.41	10.16	10.55
13%	2.98	5.44	6.44	6.44	7.49	8.67	9.48	9.81
15%	2.64	4.82	5.89	5.99	6.87	8.01	8.85	8.85
17%	2.30	4.29	5.44	5.66	6.38	7.41	8.38	8.38
19%	2.04	3.73	4.80	5.08	5.86	6.71	7.66	7.66
21%	1.89	3.20	4.41	4.78	5.57	6.22	7.02	7.02
23%	1.48	2.66	3.62	4.11	4.92	5.35	6.29	6.29
25%	1.30	2.22	3.10	3.62	4.40	4.89	5.62	5.62
27%	1.21	1.91	2.71	3.29	4.01	4.53	5.11	5.11
29%	1.12	1.66	2.36	2.96	3.61	4.11	4.54	4.51

*Initial Mobile Phase Composition

FIG. 12

EFFECT OF % Bi ON SEPARATION OF HYDROPHILIC PTH AAs

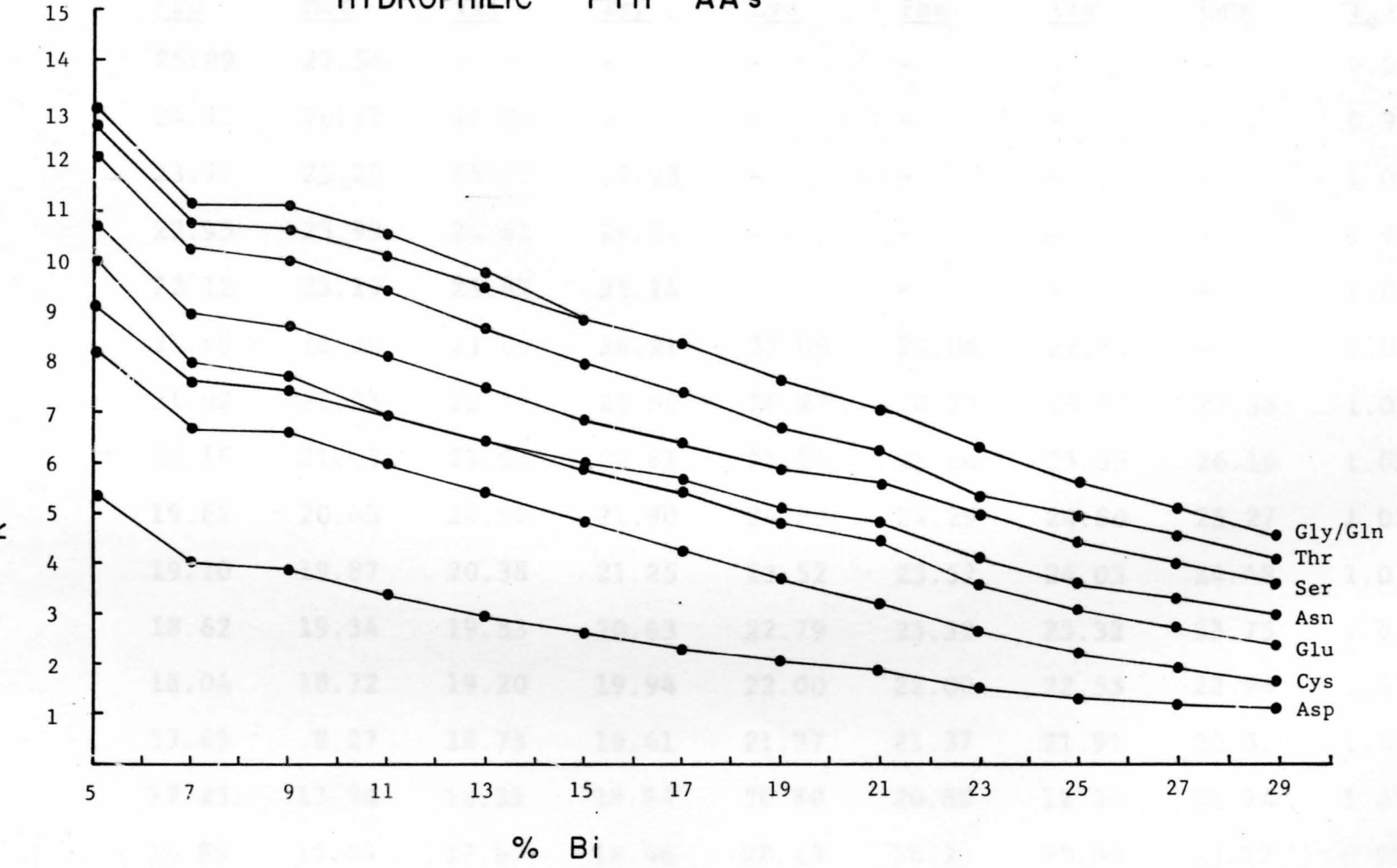


TABLE 5

RETENTION TIMES OF HYDROPHOBIC PHENYLTHIOHYDANTOINS WITH VARIATIONS IN FINAL MOBILE PHASE COMPOSITION

	Retention Times (Minutes)								
$\%B_f$*	Pro	Met	Val	Trp	Lys	Phe	Ile	Leu	T_o**
41%	25.99	27.54	-	-	-	-	-	-	0.96
43%	24.93	26.27	27.04	-	-	-	-	-	0.99
45%	23.97	25.20	25.92	27.23	-	-	-	-	1.00
47%	22.95	23.99	24.61	26.31	-	-	-	-	1.01
49%	22.12	23.19	23.82	25.14	-	-	-	-	1.01
51%	21.40	22.39	23.00	24.21	27.08	27.08	27.68	-	1.02
53%	21.02	21.93	22.52	23.59	26.27	26.27	26.79	27.38	1.03
55%	20.16	21.03	21.58	22.61	25.14	25.14	25.65	26.15	1.03
57%	19.62	20.43	20.96	21.90	24.29	24.29	24.80	25.27	1.02
59%	19.10	19.87	20.38	21.25	23.52	23.52	24.03	24.48	1.01
61%	18.62	19.34	19.83	20.63	22.79	23.32	23.32	23.75	1.01
63%	18.04	18.72	19.20	19.94	22.00	22.00	22.53	22.94	1.01
65%	17.63	.8.27	18.73	19.41	21.37	21.37	21.91	22.31	1.01
67%	17.25	17.86	18.31	18.94	20.80	20.80	21.36	21.74	1.01
69%	16.85	17.44	17.87	18.46	20.23	20.23	20.80	21.17	0.99

*Final Mobile Phase Composition

**Retention time of an unretained compound

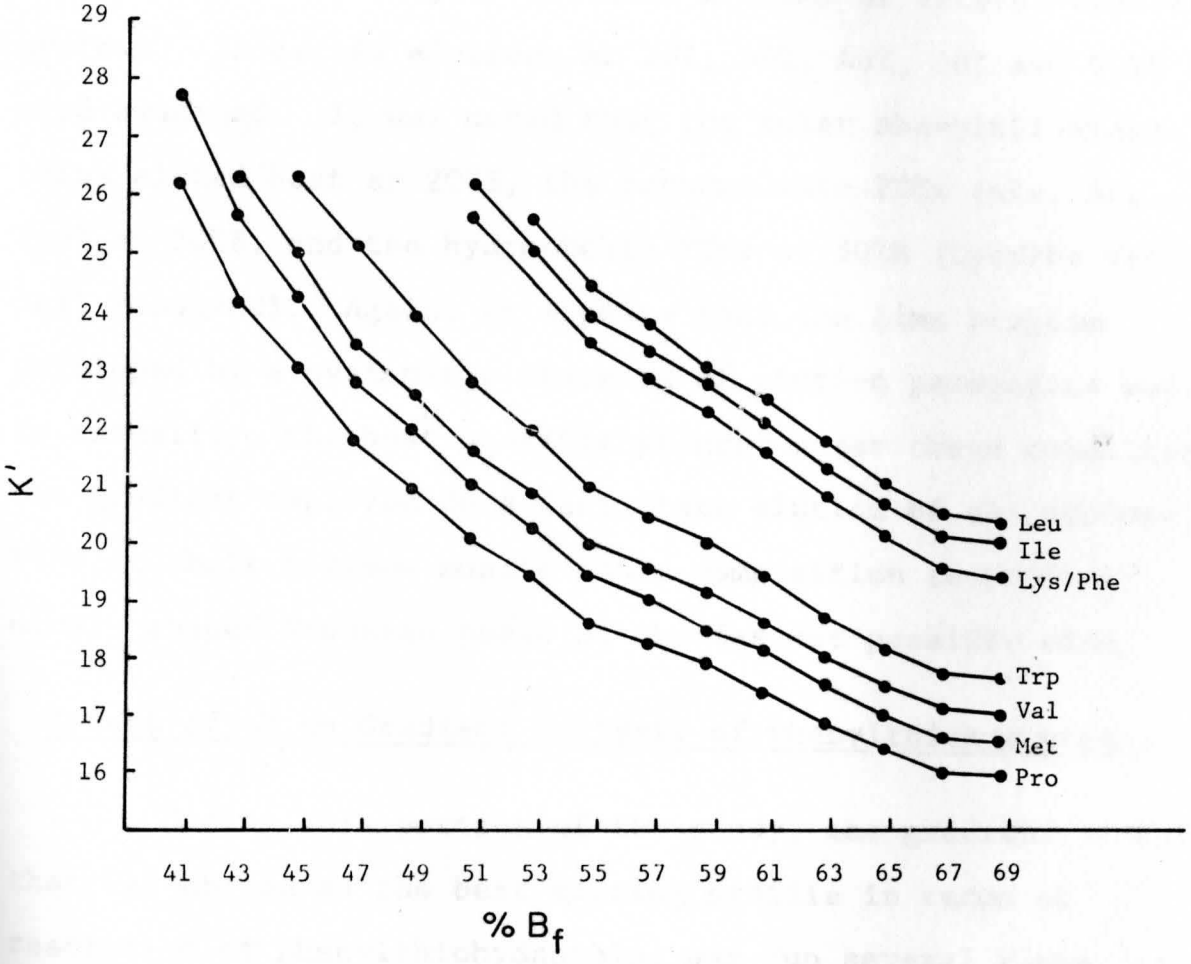
TABLE 6

CAPACITY FACTORS OF HYDROPHOBIC PHENYLTHIOHYDANTOINS WITH VARIATIONS IN FINAL MOBILE PHASE COMPOSITION

	Capacity Factors (K')							
$\%B_f$ *	Pro	Met	Val	Trp	Lys	Phe	Ile	Leu
41%	26.07	27.69	-	-	-	-	-	-
43%	24.18	25.54	26.31	-	-	-	-	-
45%	22.97	24.20	24.92	26.23	-	-	-	-
47%	21.71	22.75	23.37	25.05	-	-	-	-
49%	20.90	21.96	22.58	23.89	-	-	-	-
51%	19.98	20.95	21.55	22.74	25.55	25.55	26.14	-
53%	19.41	20.29	20.86	21.90	24.50	24.50	25.01	25.58
55%	18.57	19.42	19.95	20.95	23.41	23.41	23.90	24.39
57%	18.24	19.03	19.55	20.47	22.81	22.81	23.31	23.77
59%	17.91	18.48	19.18	20.04	22.29	22.29	22.79	23.00
61%	17.44	18.15	18.63	19.43	21.56	21.56	22.09	22.51
63%	16.86	17.53	18.01	18.74	20.78	20.78	21.31	21.74
65%	16.46	17.09	17.54	18.22	20.16	20.16	20.69	21.09
67%	16.08	16.68	17.13	17.75	19.59	19.59	20.15	20.52
69%	16.02	16.62	17.05	17.65	19.43	19.43	20.01	20.38

* Final Mobile Phase Composition

FIG. 13. EFFECT OF FINAL % B ON SEPARATION OF HYDROPHOBIC PTH AAs



Below 50%B, the elution is too slow to allow these PTHs to pass through the column within the 30 minute time period desired. This knowledge, also, supports the method developed to this point; it was felt that no gradient elution at the end of the chromatogram would resolve Lys/Phe as well as the isocratic system developed.

The third project involved a study of several isocratic systems. Isocratic elutions at 20%, 30%, 40%, 50% and 60%B were examined. It was noted that the polar phenylthiohydantoins eluted best at 20%B, the intermediate PTHs (Ala, Arg, Tyr) at 30%B, and the hydrophobic PTHs at 50%B (Lys/Phe very well resolved). Again, it appears that the time program developed by a systematic changing of elution parameters was, in actuality, the best possible program under these conditions. The gradient employed does facilitate elution of the various PTHs at their optimum mobile phase composition to produce nicely shaped Gaussian peaks at the fastest possible time.

Effect of pH on Gradient Analysis of Phenylthiohydantoins

During this portion of the study, the gradient profile that was chosen as the best eluting profile in terms of resolution of phenylthiohydantoins was run several times, each time altering only the pH of the sodium acetate buffer. Table 7 lists the retention times and capacity factors (K') of the PTH residues under these variable conditions. It should be noted that the reversed phase column can be destroyed by pH extremes; therefore, this study was conducted

TABLE 7

RETENTION TIMES AND CAPACITY FACTORS OF PHENYLTHIOHYDANTOINS WITH VARIATIONS IN pH

	Retention Times (minutes)						Capacity Factors (K')					
	pH3.8	pH4.2	pH4.6	pH5.0	pH5.4	pH5.8	pH3.8	pH4.2	pH4.6	pH5.0	pH5.4	pH5.8
Asp	6.85	4.97	3.58	3.14	2.98	2.77	5.65	3.97	2.58	2.05	1.98	1.69
Cys	13.61	7.39	5.94	5.32	5.06	4.87	12.21	6.39	4.94	4.17	4.06	3.73
Glu	14.25	13.51	9.94	6.82	5.44	4.50	12.83	12.51	8.94	5.62	4.44	3.37
Asn	7.59	7.44	7.44	7.48	7.46	7.42	6.37	6.44	6.44	6.26	6.46	6.20
Ser	7.58	7.53	7.63	7.61	7.65	7.74	6.36	6.53	6.63	6.39	6.65	6.51
Thr	9.39	8.91	9.39	9.32	8.97	8.50	8.12	7.91	8.39	8.05	7.97	7.25
Gly	11.12	11.11	11.42	11.54	11.24	11.28	9.80	10.11	10.42	10.20	10.24	9.95
Gln	11.96	12.30	12.38	12.70	12.41	12.47	10.61	11.30	11.38	11.33	11.41	11.11
His	15.73	16.15	15.79	15.00	15.24	14.50	14.27	15.15	14.79	13.55	14.24	13.08
Ala	15.01	15.03	15.11	15.18	15.09	15.06	13.57	14.03	14.11	13.74	14.09	13.62
Tyr	15.68	15.69	15.75	15.73	15.71	15.72	14.22	14.69	14.75	14.27	14.71	14.26
Arg	16.46	16.46	16.55	16.96	16.89	16.87	14.98	15.46	15.55	15.47	15.79	15.38
Pro	17.60	17.61	17.70	17.68	17.66	17.66	16.09	16.61	16.70	16.17	16.66	16.15
Met	17.96	17.96	18.05	18.02	18.00	18.01	16.44	16.96	17.05	16.50	17.00	16.49
Val	18.33	18.33	18.42	18.39	18.34	18.38	16.80	17.33	17.43	16.85	17.34	16.48
Trp	18.63	18.63	18.72	18.68	18.68	18.67	17.09	17.63	17.72	17.14	17.68	17.13
Lys	20.54	20.54	20.66	20.62	20.65	20.61	18.94	19.54	19.66	19.02	19.65	19.01
Phe	20.92	20.93	21.05	21.01	21.00	20.99	19.31	19.93	20.05	19.40	20.00	19.38
Ile	21.45	21.46	21.59	21.56	21.55	21.53	19.83	20.46	20.59	19.93	20.55	19.90
Leu	22.31	22.33	22.46	22.42	22.37	22.39	20.66	21.33	21.46	20.77	21.37	20.74
T_o*	1.03	1.00	1.00	1.03	1.00	1.03	-	-	-	-	-	-

* Retention time of an unretained compound

over the moderate pH range of pH 3.8 to pH 5.8.

Alterations in the pH had little or no effect on the elution times of the more hydrophobic amino acids that elute late in the elution profile. This observation would be expected due to the fact that these non-polar amino acids are driven off the column primarily by methanol, the strong solvent, and are therefore not affected by the sodium acetate buffer.

The retention times of the more polar amino acids, however, are in some cases very drastically changed with variations in pH. Glutamic acid eluted almost 10 minutes later when the pH is decreased from 5.8 to 3.8. Cysteic acid and aspartic acid also demonstrated an increase in retention time with decrease in pH. These increases in retention times at the more acidic pH levels cause a "crowding" effect of the elution pattern of the phenylthiohydantoin chromatogram, due to the fact that the first amino acids elute later and the latter amino acids are not affected.

Analysis of Protein Sequence Residues by HPLC

Lysozyme, an enzymatic protein comprised of 129 amino acids, was placed on the Beckman 890C for sequencing. A 10 mg sample of the protein was used, and the instrument was programmed to cleave the first 10 amino acids. After conversion, these residues were analyzed by HPLC using the method developed in this study. Chromatograms of representative residues are shown on Figs. 14, 15, 16. The amino acid

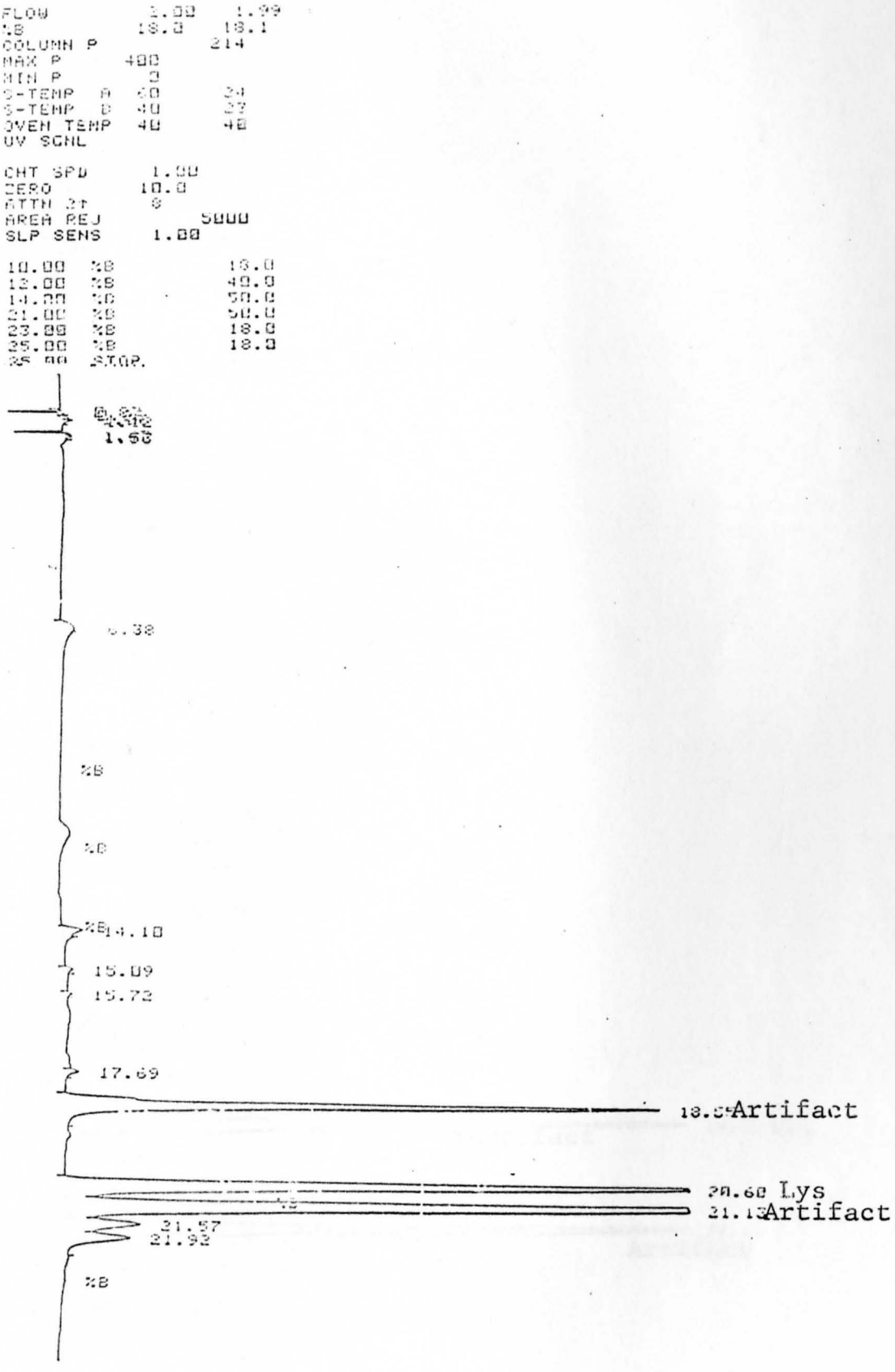


Fig. 14. Gradient Elution of Lysozyme Residue #1.

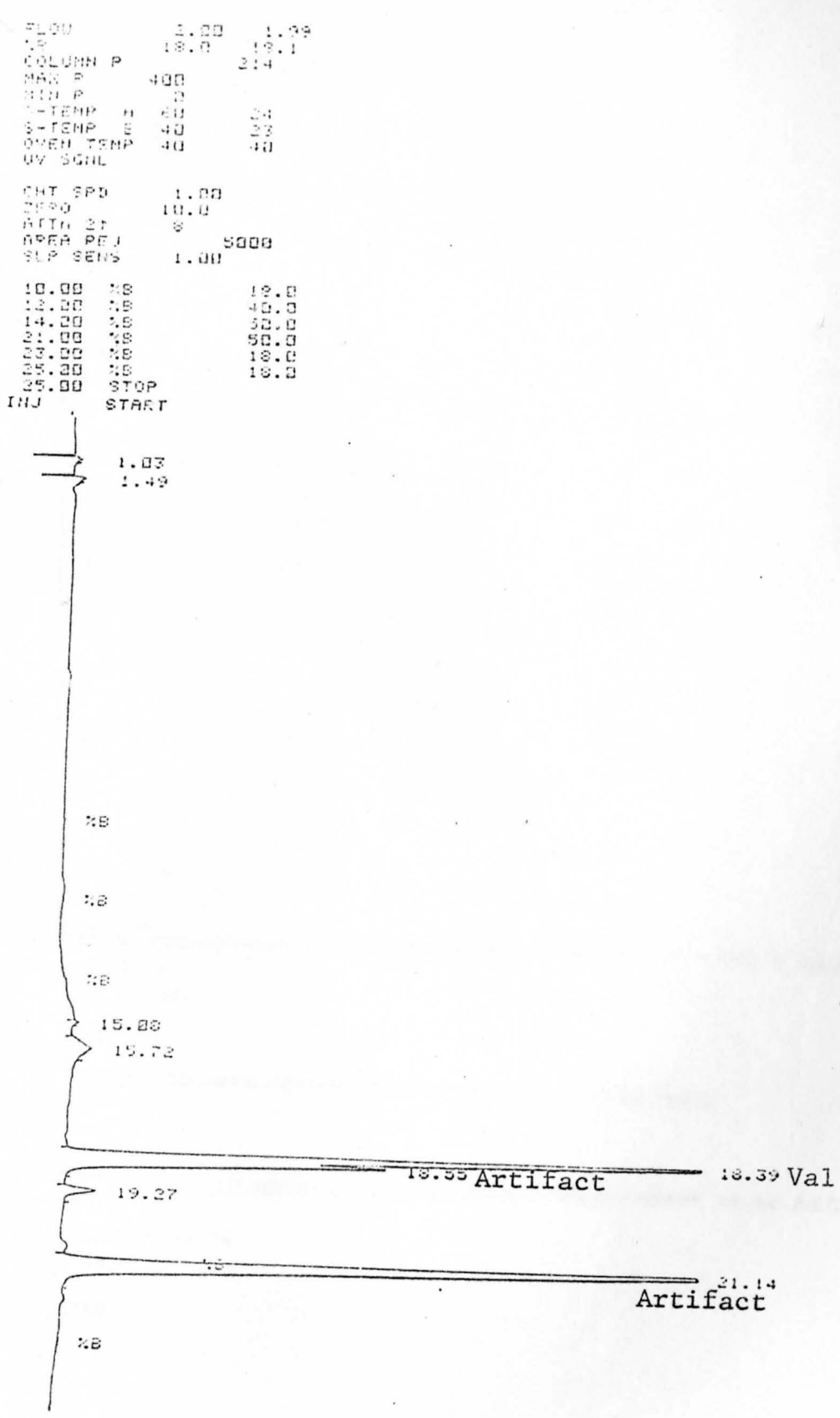


Fig. 15. Gradient Elution of Lysozyme Residue #2.

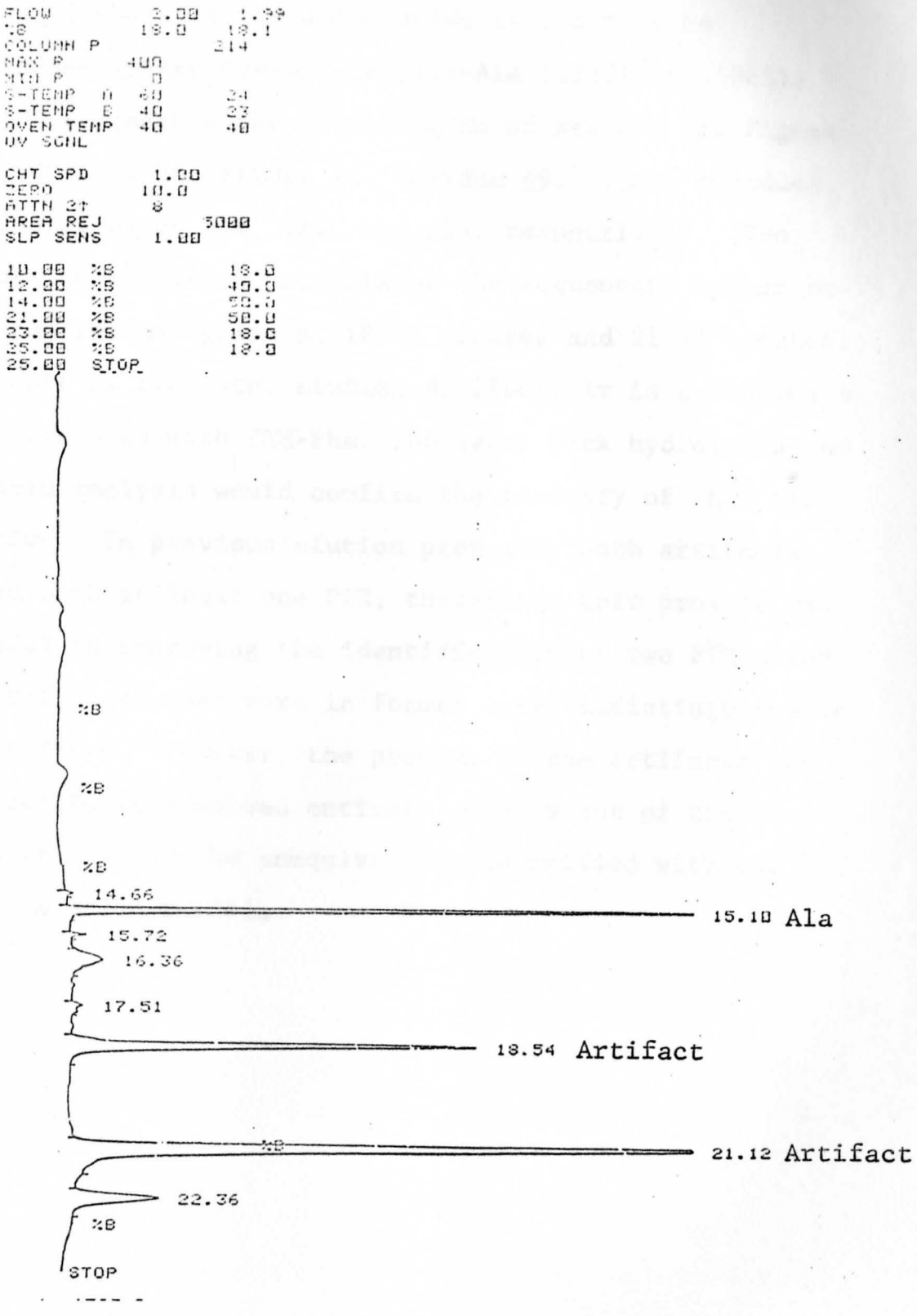


Fig. 16. Gradient Elution of Lysozyme Residue #9.

sequence of the first 10 amino acids is known to be Lys-Val-Phe-Gly-Arg-Cys-Glu-Leu-Ala-Ala (Canfield, 1963). Figure 14 represents the chromatogram of Residue #1, Figure 15, Residue #2, and Figure 16, Residue #9. These residues were identified as Lys, Val, and Ala, respectively. Two artifacts, degradation products of the sequencer, appear on all of the chromatograms at 18.55 minutes and 21.13 minutes. In the case of the later eluting artifact, it is unfortunate that it coelutes with PTH-Phe. However, back hydrolysis and amino acid analysis would confirm the identity of this PTH derivative. In previous elution programs, both artifacts coeluted with at least one PTH; therefore, this project was successful in improving the identification of two PTH amino acids (Val, Lys) that were in former cases indistinguishable from artifacts. However, the problem of the artifacts was not successfully resolved entirely, and 19 out of the 20 amino acids can now be unequivocably identified with one gradient elution profile.

CHAPTER V

DISCUSSION OF RESULTS

This study involved several aspects of the analytical technique of high pressure liquid chromatography. The identification and quantitation of phenylthiohydantoins, the products of protein sequencing, were accomplished. During this study, a variety of HPLC parameters were investigated to determine their effect on the resolution of PTH derivatives of amino acids.

It was found that the mobile phase composition of the system was the single most important parameter, without altering solvents, temperature, or column type. This fact was demonstrated throughout the method development. A change in the %B was in most cases the most effective modification to affect peak resolution (e.g., Lys/Phe, Thr/Gly/Gln).

Changes in flow rate also affect the retention times of solute peaks. However, this parameter was not found to be useful in the resolution of phenylthiohydantoins. For this reason, a constant flow rate of 2 ml per minute was used throughout.

The effect of the pH of the weak solvent (i.e., sodium acetate buffer) on the elution pattern was a major portion of this investigation. This parameter, although having drastic effects on the retention times of some PTHs (e.g.,

Glu, Asp, Cys), was not found to be effective in this study. At the lower pH levels, the crowding of some peaks due to longer retention times caused a loss of resolution. (At pH3.8, Asn and Ser coelute. At pH4.2, Asn and Cys coelute). At the higher pH levels, the Glu and Cys peaks elute very close to each other. The knowledge of the possible detrimental effects of pH extremes to the column, together with the observations noted above, led to the conclusions that a sodium acetate buffer at pH 5.0 was probably the most beneficial in the identification of phenylthiohydantoins and that changes in pH did not improve the resolution of PTH chromatographic peaks.

It should be emphasized that the results of the method development using a logical sequential progression of differing the mobile phase composition corresponded quite well with the differential analysis of the gradient systems. In the studies of the effect of initial %B and final %B of an elution scheme, it was concluded that at no particular mobile phase composition could all PTH amino acids be separated. In order to resolve Thr/Gly/Gln, or Lys/Phe, a period of isocratic elution was required. The chromatographic scheme that was developed did, in fact, contain 2 periods in which the %B remained at the same level. Likewise, the mobile phase composition at each point of a PTH peak corresponded well with the isocratic elution studies to produce well-defined, nicely shaped, peaks.

In order to statistically prove that the program developed was superior to previous gradient elution schemes, a comparison of the resolutions (R) of all 20 PTHs of the first "scouting" chromatogram (Fig. 6), an intermediate chromatogram of a linear gradient scheme (Fig. 9), and the final chromatogram (Fig. 11) was conducted. Retention times were collected and peak base widths were measured to facilitate the calculation of resolution (Tables 8,9). Upon inspection, the graph of these resolution results (Fig. 17) indicates that the final chromatograph was, in fact, much better at resolving phenylthiohydantoins. An "R" value of 1.5 constitutes effective baseline separation of 2 peaks. It is important to note that values above 1.5 are of no consequence; that is, baseline separation is the best resolution that one can achieve. "R" values above 1.5 only indicate that the peaks are very widely spread, and, if carried to an extreme, may actually indicate a negative aspect to an effective chromatogram in which time of analysis is important. In the Fig. 6 profile, there are several R values that fall short of 1.5. There are also 4 sets of peaks that coelute; that is, there is virtually no separation (R=0). The Fig. 9 profile improves somewhat over Fig. 6, (e.g., Thr/Gly); however, the Fig. 11 profile is by far the best. No peaks coelute, and only 4 sets of peaks have "R" values less than 1.5; Thr/Gly, Gly/Gln, Tyr/Arg, and Lys/Phe. However, although baseline separation was not achieved in all cases, adequate resolution for unequivocal

TABLE 8

RETENTION TIMES AND PEAK BASE WIDTHS OF PHENYLTHIOHYDANTOIN DERIVATIVES IN 3 REPRESENTATIVE CHROMATOGRAMS

Amino Acid	Retention Time (cm)*			Peak Base Width (cm)		
	Fig.6	Fig.9	Fig.11	Fig.6	Fig.9	Fig.11
Asp	4.12	3.36	2.83	0.25	0.2	0.5
Cys	5.98	5.41	4.98	0.25	0.2	0.5
Glu	6.68	6.14	5.65	0.2	0.3	0.45
Asn	6.9	6.91	7.83	0.5	0.3	0.5
Ser	7.86	7.78	9.70	0.7	0.8	0.45
Thr	8.86	8.50	11.51	0.9	0.5	0.4
		8.93**			0.7	
Gly	8.86	9.33	11.97	0.9	0.55	0.5
Gln	8.86	9.33	12.49	0.9	0.55	0.4
His	11.49	13.19	14.05	0.4	0.3	0.6
Ala	11.49	13.19	15.06	0.4	0.3	0.2
Tyr	12.38	15.05	16.08	0.3	0.25	0.2
Arg	12.38	15.05	16.36	0.3	0.25	0.4
Pro	14.80	18.85	18.15	0.35	0.25	0.2
Met	15.14	19.60	18.54	0.3	0.25	0.2
Val	15.45	20.09	18.92	0.35	0.25	0.2
Trp	15.69	20.94	19.29	0.35	0.25	0.3
Lys	16.80	23.11	21.33	0.3	0.3	0.4
Phe	17.02	23.11	21.63	0.4	0.3	0.4
Ile	17.33	23.43	22.15	0.5	0.35	0.3
		23.64	22.42		0.4	0.3
Leu	17.63	24.07	23.07	0.4	0.3	0.45

*Chart paper moves at 1 cm per min; therefore, retention values may be in time (min.) or in distance (cm).

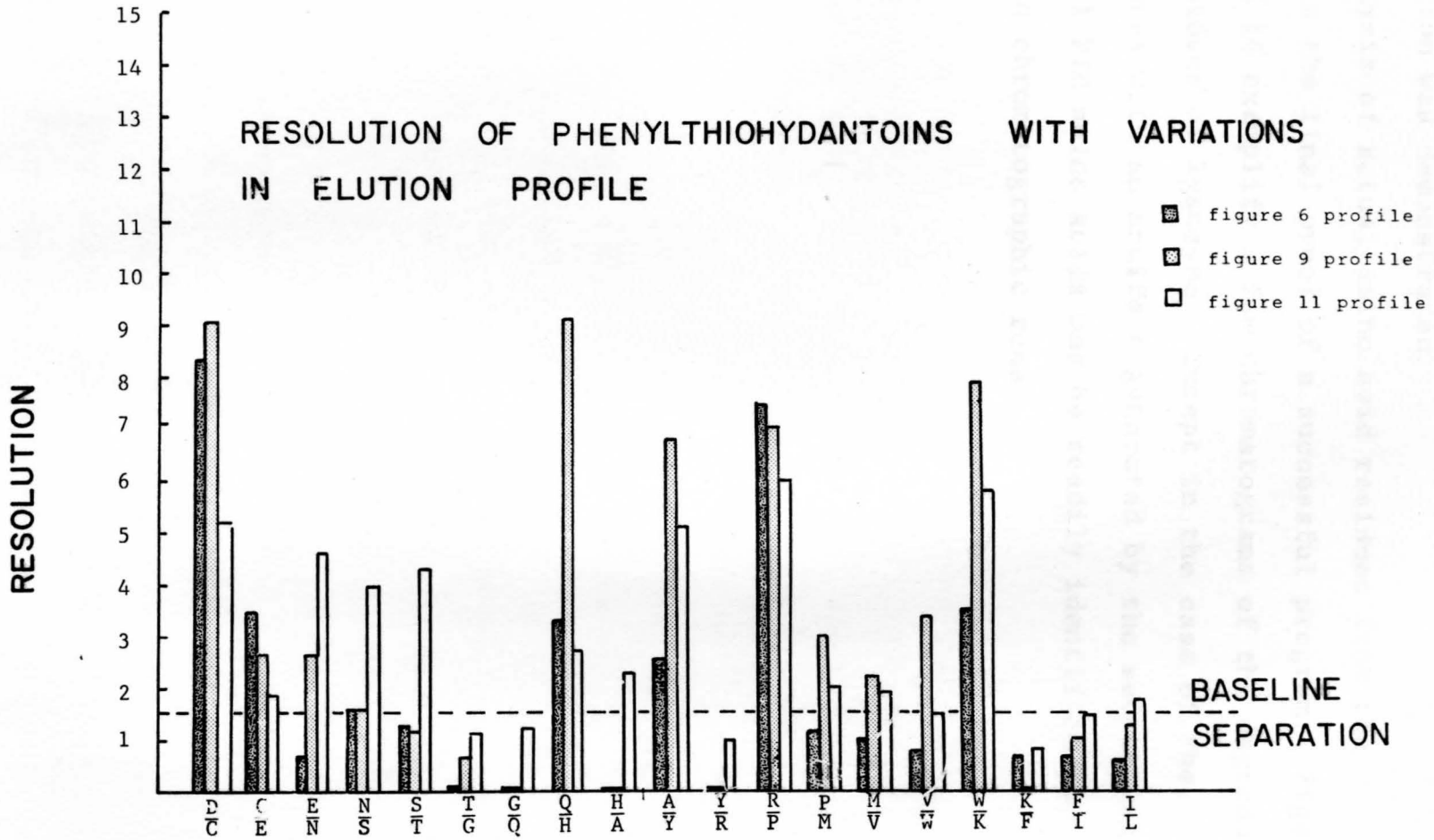
** Elutes as a doublet peak with two retention times. Resolution based on peak of doublet that is closest to second peak of resolution.

TABLE 9

RESOLUTIONS OF PHENYLTHIOHYDANTOIN DERIVATIVES IN 3 REPRESENTATIVE CHROMATOGRAMS

Amino Acids	Resolutions (R)		
	Fig.6	Fig.9	Fig. 11
Asp/Cys	8.27	9.11	5.06
Cys/Glu	3.50	2.65	1.91
Glu/Asn	0.63	2.57	5.49
Asn/Ser	1.60	1.58	3.94
Ser/Thr	1.25	1.11	4.26
Thr/Gly	0	0.64	1.02
Gly/Gln	0	0	1.16
Gln/His	3.29	9.08	2.60
His/Ala	0	0	2.25
Ala/Tyr	2.54	6.76	5.10
Tyr/Arg	0	0	0.93
Arg/Pro	7.45	7.00	5.97
Pro/Met	1.05	3.00	1.95
Met/Val	0.95	1.96	1.90
Val/Trp	0.74	3.40	1.48
Trp/Lys	3.42	7.89	5.83
Lys/Phe	0.63	0	0.75
Phe/Ile	0.69	0.98	1.49
Ile/Leu	0.67	1.23	1.73

FIG. 17


RESOLUTION OF PHENYLTHIOHYDANTOINS WITH VARIATIONS IN ELUTION PROFILE
figure 6 profile
figure 9 profile
figure 11 profile
RESOLUTION
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
BASELINE SEPARATION
D/C
C/E
E/N
N/S
S/T
T/G
G/Q
Q/H
H/A
A/Y
Y/R
R/P
P/M
M/V
V/W
W/K
K/F
F/I
I/L

identification was demonstrated.

Analysis of actual amino acid residues from the sequencer is the final proof of a successful program. Figs. 14, 15, and 16 exemplify a few chromatograms of the degraded protein residues of lysozyme. Except in the case of Phe, which coelutes with an artifact generated by the sequencing process, all PTH amino acids may be readily identified in these single chromatographic runs.

CHAPTER VI

SUMMARY

The identification and quantitation of phenylthiohydantoins is a very important part of the amino acid sequencing process; a method that is relatively fast (under 30 minutes), convenient, accurate, and sensitive is ideal. High pressure liquid chromatography is the method that seems to lend itself best to these restrictions. HPLC is a very versatile tool that was used in this study to analyze PTHs. HPLC is as efficient and sensitive as GLC, the previous method of choice, if not more so. A combination isocratic/gradient elution scheme to effect alterations in the mobile phase composition was developed that was quite successful in the resolution of the 20 phenylthiohydantoin peaks.

It should be noted that during the process of the study, several columns of the same type were used, due to the 3 month average lifetime of a reversed phase column. It was noticed that column-to-column variations do exist, which at times made a homogeneous development quite difficult. However, although minor variations do occur, it was found that the columns did behave quite similarly, so that once major trends were established, they could be applied to similar reversed phase C-8 columns used. For instance, it would be expected that an isocratic period at the beginning and

again at the end of a chromatogram of any C-8 column would, indeed, be necessary for effective separation of PTHs, although there may be small differences in optimum times or mobile phase composition.

Mobile phase composition was found to be the major parameter in this study. Changes in flow and pH, while effective in changing retention times, were not as effective in increasing the resolution of phenylthiohydantoins. These parameters might well be more useful, however, in some alternate system. It is felt that any further development work on this particular system would probably not be profitable in view of the amount of time spent versus the results obtained. A good method of identification and quantitation has been developed, taking into account the restrictions of the instruments available, as well as the restrictions of the analysis, such as solvent (methanol), time (30 minutes), and cost. In order to develop a better analysis method in the future, a variable wavelength detector might be considered. This addition would certainly improve the detection of threonine, which gives a stronger signal at 313 nm as dihydrothreonine.

In summary, a 25 minute elution system using high pressure liquid chromatigraphy has been developed in which all 20 amino acids are separated in one run with adequate resolution to allow unequivocal identification. Methanol, which is cheaper and less toxic than acetonitrile, was selected as the strong solvent. The 25 minute program is

quick enough to keep in step with the automatic sequencer so that immediate feedback is received. The only drawback to this method is that Phe coelutes with an artifact from the sequencer. This problem artifact is one of two that are results of the program used in the sequencer, and would be very difficult to eradicate, except by altering the sequencer program. An increase in the area of this peak would indicate that a Phe was present and could be confirmed by back hydrolysis to the parent amino acid. Therefore, this problem is minimized with a backup amino acid analyzer, and the method developed and described here is considered to be successful, since it provides a sensitive and accurate analytical system useful in the rapid sequence analysis of proteins and peptides.

APPENDIX

LIST OF ABBREVIATIONS OF AMINO ACIDS

Amino Acid	3 Letter Abbreviation	1 Letter Code
Aspartic Acid	Asp	D
Cysteic Acid	Cys	C
Glutamic Acid	Glu	E
Asparagine	Asn	N
Serine	Ser	S
Threonine	Thr	T
Glycine	Gly	G
Glutamine	Gln	Q
Histidine	His	H
Alanine	Ala	A
Tyrosine	Tyr	Y
Arginine	Arg	R
Proline	Pro	P
Methionine	Met	M
Valine	Val	V
Tryptophan	Trp	W
Lysine	Lys	K
Phenylalanine	Phe	F
Isoleucine	Ile	I
Leucine	Leu	L

BIBLIOGRAPHY

Abrahamsson, Monika, Groningsson, Kerstin, and Castensson, Staffan, Journal of Chromatography, 154:313-317 (1978).

Bhown, Ajit S., et.al., Journal of Chromatography, 248:532-535 (1978).

Bridgen, John, FEBS Letters, Vol. 50, No. 2, 159-162 (1975).

Canfield, R.E.. J. Biol. Chem., 238: 2698-2707 (1963).

Downing, Michael R., and Mann, Kenneth G., Analytical Biochemistry, 74: 298-319 (1976).

Edman, Pehr, Acta. Chemica Scandinavica, 4: 283-293 (1950).

Fales, Henry M., et. al., Anal. Biochem., 43: 288-299 (1971).

Giddings, J.C., Dynamics of Chromatography. Part I. Principles and Theory. New York: Marcel Dekker, 1965.

Gray, W.R., and Hartley, B.S., Biochem. J., 89: 379-380 (1963).

Harman, R.E., Patterson, J.L., and Vandenhurval, W.J.A., Analytical Biochemistry, 25: 452-458 (1968).

Hermodson, Mark A., et. al., Biochemistry, Vol. 11, No. 24, 4493-4502 (1972).

Hirs, CHW, Moore, Stanford, and Stein, William H., The Journal of Biological Chemistry, Vol. 235, No. 3, 633-647 (1960).

Horvath, C.G., Preiss, B.A., and Lipsky, S.R., Anal. Chem., 39:1422 (1967).

Huang, Wei-Yong, et. al., Archives of Biochemistry and Biophysics, 166: 67-71 (1975).

Huber, J.F.K., J. Chromatogr. Sci., 7:85 (1969).

Hunkapillar, Michael W., and Hood, L.E., Biochemistry, Vol. 17, No. 11, 2124-2133 (1978).

Johnson, N.D., Hunkapiller, M.W., and Hood, L.E., Analytical Biochemistry, 100: 335-338 (1979).

Kirkland, J.J., J. Chromatogr. Sci., 7:7 (1969).

Kulbe, Klaus D., Analytical Biochemistry, 59: 564-673 (1974).

Lovins, Robert E., et. al, Anal. Biochem., 47: 537-549 (1972).

Margolies, Michael, and Bauer, Andrew, Journal of Chromatography, 148: 429-439 (1978).

Matthews, Eric W., Byfield, Peter G.H., and MacIntyre, Iain, Journal of Chromatography, 110: 369-373 (1973).

Pisano, John J., and Bronzert, Thomas J., The Journal of Biological Chemistry, Vol. 244, No. 20, 5597-5607 (1969).

Pryde, A., and Gilbert, M.T., Applications of High Performance Liquid Chromatography, John Wiley: New York (1979).

Smithies, O., et. al., Biochemistry, Vol. 10, No. 26, 4912-1921 (1971).

Summers, Murray R., Smythers, Gary W., and Oroszlan, Stephen, Analytical Biochemistry, 53: 624-628 (1973).

Zeeuws, R., and Strosberg, A.D., FEBS Letters, Vol. 85, No. 1, 68-72 (1978).

Zimmerman, Carl L., Appella, Ettore, and Pisano, John J., Analytical Biochemistry, 75: 75-85 (1976).

Zimmerman, Carl L., Appella, Ettore, and Pisano, John J., Analytical Biochemistry, 77: 569-573 (1977).