PARTIAL PRIMARY STRUCTURE OF GAS VACUOLE MEMBRANES FROM MICROCYSTIS AERUGINOSA

Dissertation for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
PAMELA J. WEATHERS
1974

LIBRARY
Michigan State
University

This is to certify that the

thesis entitled

PARTIAL PRIMARY STRUCTURE OF
GAS VACUOLE MEMBRANES
FROM MICROCYSTIS AERUGINOSA

presented by

Pamela J. Weathers

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Botany

Michael Doot

Major professor

Date 11/15/74

O-7639

BINDING BY
HOAG & SONS'
BOOK BINDERY INC.
LIBRARY BINDERS
SPRINGPORT, MICHIGAN

C 2076

ABSTRACT

PARTIAL PRIMARY STRUCTURE OF GAS VACUOLE MEMBRANES FROM MICROCYSTIS AERUOGINOSA

By

Pamela J. Weathers

The primary structure and hydrophobic characteristics of the gas vacuole membrane were investigated in order to determine if this simple membrane, containing only one protein, typifies integral membrane proteins as defined by Singer and Nicholson (17).

Peptides resulting from trypsin digestion and N-Bromosuccinimide treatment of gas vacuole protein were sequenced using a variety of methods. Sequence analysis from the amino terminus was performed using automated Edman degradation on a Beckman Sequencer Model 890C. Modification of the peptides to prevent their loss from the reaction cup was effected by making the carboxyl terminus more hydrophilic. This was accomplished by using either 4-sulphophenyl isothiocyanate for lysine, or water soluble carbodiimide and naphthalene disulfonic acid for other amino acids. I improved the old carbodiimide procedure by using additional aliquots of carbodiimide, especially for insoluble peptides. Sequence analysis from the carboxyl end of peptides was performed using carboxypeptidase C.

All sequences are shown below and are preceded by their designated names:

T2A3a: Ala-Val-Glu-Lys

T1P1b: Tyr-Ala-Glu-Ala-Val-Gly-Leu-Thr-Glu-Ser-Ala-Ala-Val-Pro-(15 residues)-Arg-Tyr-Ala-Glu-Ala-Val-Gly-Leu-Thr-Glu-(Ser)-Ala-(Pro)-(Val)-Ala-Ala

T1P1a: Ser-Ala-Glu-Ala-Val-Gly-Leu-Thr-Glu-Val-(Ile)-(Ala)-(x number of residues)

T2A2: Gly-Ile-Val-Ile-(Asp)-(Ala)-Ala-Arg

T2A3b: Ile-Leu-Asp-Lys

T2A4: Lys

NIA: Ala-Glu-Ala-Val-Gly-Leu-Thr-Glu-(Ser)-Ala-(Pro)-(Val)-Ala-Ala

NPT: Ala-Val-(Val)-(Val)-Leu-Val-(Val)-Ile-(Ile)-Leu-(Leu)-Ala-(Leu)-(Val)-(Ile)-(x number of residues)

Peptide T2A3a is the amino terminus of the protein; peptide T1P1b, the carboxyl terminus. Peptide NPT and all others reside somewhere between the two terminal peptides.

Two aspects of the sequence analysis are intriguing. Firstly, the sequence of peptide NPT is one of the most hydrophobic sequences of a protein yet described; such a long, aliphatic stretch is rare. Based on the amino acid sequence of a protein, current methods (32) permit determination of the regions of secondary structure (helix, sheet, turn) in a protein with 85% certainty. Application of these methods shows that the amino acid sequence of the NPT peptide would equally favor either helix or sheet formation.

The second important feature of the primary structure of the gas vacuole protein is the presence of the thrice repeating octapeptide:

Ala-Glu-Ala-Val-Gly-Leu-Thr-Glu

A determination of the secondary structure, as above, indicates that this sequence strongly favors helix formation. It is generally inferred (85) that the presence of a repeating sequence in a polypeptide suggests its use as a structural building block.

Based on the presence of peptide NPT, a functional model for the gas vacuole membrane was proposed. Gases pass through the membrane by diffusion. Since many diatomic gases are apolar, passage through an apolar milieu would provide a path of least resistance. Such an apolar milieu would exist in the aliphatic, amino terminal portion of the peptide NPT. The assumption is made that a subunit substructure exists in the membrane. To allow gas to move in and out without a conformational change in the membrane (for which there is no evidence at present), either the gas must pass through the intermolecular space in a protein or through pores. The lining of either of these passageways would be the aliphatic portion of the NPT peptide. Large and/or polar molecules would be restricted from passing through this area based on their size (at least 3 $\mathring{A}$), and charge.

The presence of the peptide NPT internal to the amino and carboxyl termini of the gas vacuole protein, provides the protein

with an amphipathic nature. Based on this and a comparison of the relative polarities of gas vacuole protein to other integral membrane proteins, it is concluded that the gas vacuole protein is a true membrane protein of the integral type. Perhaps, this membrane is even a prototype for more complex membranes.

PARTIAL PRIMARY STRUCTURE OF GAS VACUOLE MEMBRANES FROM *MICROCYSTIS AERUGINOSA*

By

Pamela J. Weathers

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Botany and Plant Pathology

1974

DEDICATION

to my husband, Larry, and

our Oakhill families

ACKNOWLEDGMENTS

The author wishes to express her gratitude to her major professor, Dr. Michael Jost, for his guidance throughout these investigations. Special appreciation is extended to Dr. Derek T.A. Lamport for his extensive advice, use of his laboratory equipment, and encouragement throughout these studies. The constructive criticism and advice of the committee members Dr. Alfred Haug and Dr. Robert Bandurski is also appreciated. The author also wishes to thank Dr. Judith Foster of the Boston University School of Medicine for her helpful discussions and guidance in sequencing procedures. The technical assistance of Nell Brittain, Ray Sculley, Laura Katona, and Mary Shimamoto is gratefully acknowledged.

This work was supported under Contract No. AT(11-1)-1338 with the U.S. Atomic Energy Commission.

TABLE OF CONTENTS

Page

RESULTS

DISCUSSION

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES--continued

LIST OF ABBREVIATIONS

GVP	gas vacuole protein
CD	circular dichroism
IR	infra red analysis
EDC	N-ethyl, N^{1}-(-3-dimethylamino propyl carbodiimide) HCl
ANS	2-amino-1, 5 naphthalene disulfonic acid
SPITC	4-sulphophenyl isothiocyanate
PTH	phenylthiohydantion
TMS	trimethylsilylated derivative
NBS	N-Bromosuccinimide
DNS	dansylated derivatives
DMAA	dimethylallylamine-trifluoroacetic acid in pyridine-water
Tris	tris (hydroxymethyl) aminomethane
SDS	sodium dodecyl sulfate
DAP I	dipeptidyl aminopeptidase I

INTRODUCTION

1.1 Membrane Structure: Fluid Mosaic Model. Biological membranes play a crucial role in many cellular phenomena. Concerning the spatial arrangements of membrane constituents, relatively little is known. Therefore, models have been proposed. One of the most favored is the fluid mosaic model proposed by Lenard and Singer in 1966 (1). This model accepts the basic lipid bilayer as originally proposed by Danielli, Davson and Robertson (2,3) with proteins inserted into the bilayer in a mosaic fashion. Some proteins may penetrate partially, whereas others may span the entire bilayer. In addition, the model suggests that proteins can migrate laterally in the membrane depending on the fluidity of the bilayer.

Evidence for each of these basic tenets of the model comes from various membrane systems. Although there has always been experimental data supporting a lipid bilayer (1, 4, 5, 6), its existence only recently has been demonstrated unequivocally by X-ray diffraction (7) for Acholeplasma membranes.

The concepts of fluidity and lateral mobility of lipids and proteins in the membrane have been discussed recently in a review by Singer (8). It has, for example, been shown for membranes of A. laidlawii that temperatures at or near the phase transition temperature of the membrane lipids can affect the surface pattern of associated membrane proteins. This and other evidence (9, 10, 11) also establishes that membranes contain proteins which are partially or wholly intercalated into the lipid bilayer.

Although a fluid mosaic model is generally accepted in an architectural sense, the nature of the interaction occurring between the lipid and protein moieties of the membrane is still under consideration. However, covalent linkages are generally excluded as a major type of interaction between lipid and protein within the membrane (12). On the contrary, these interactions must be weak and can be mediated by charge-charge interactions, hydrogen bonding, van der Waals interactions, and hydrophobic bonding (1, 12, 13, 14).

1.2 The Nature of Membrane Proteins. Analysis of the secondary structure of the proteins of intact membranes shows that they contain about 40% helix which suggest that the proteins associated with the membrane are largely globular in shape (12, 15, 16, 17). However, since these data were obtained using circular dichroism (CD), optical anomalies arising from light scattering by particulate systems preclude exact calculation of the amount of secondary structure in the membranes (17).

There is a large body of evidence which suggests that proteins of biological membranes are asymmetrically distributed across a membrane. No evidence for a corresponding distribution of lipids has yet been found (8). Consequently, only examples of protein asymmetry will be mentioned here. These examples are membranes of Sarcoplasmic reticulum (11), the cytochrome b_5 and NAD-cytochrome b_5 reductase proteins of microsomal membranes (18, 19), and glycophorin of the erythrocyte membrane (20). Singer and Nicolson (17) have categorized membrane proteins as either peripheral or integral.

Peripheral proteins require only mild treatment to dissociate them from membranes; they dissociate free of lipids; and they are relatively soluble in neutral aqueous buffers. Integral proteins, which make up the major portion (more than 70%) of the proteins in most membranes, require very drastic treatments to dissociate them from membranes; once isolated, they may remain associated with lipids; and if free of lipids, they are usually highly insoluble in neutral aqueous buffers. Only the integral membrane proteins are critical to the structural integrity of the membrane (17). In addition, there is no convincing evidence that only one predominant type of membrane structural protein exists (17, 21). Because peripheral membrane proteins are less involved in membrane structure than integral membrane proteins, they will not be discussed further.

1.3 Integral Membrane Proteins. One of the most striking features of all integral membrane proteins studied, so far, is that they are extremely hydrophobic. The majority of integral membrane proteins have a polarity, as determined by their amino acid composition, which is significantly lower (on the average 10-15%) than that of most soluble proteins (22). However, hydrophobicity is not always due to a higher percentage of apolar amino acids. For example, penicillinase, a lipoprotein, is considered an integral membrane protein, yet its hydrophobic properties are due to the lipid moiety rather than an abundance of hydrophobic amino acids (23).

Little is known about the role of hydrophobic bonds (24), interactions involving nonpolar amino acids with nonpolar side chains.

Several theories, though, have been espoused relating hydrophobicity to thermal stability. Scheraga et al. (25) and Bigelow (26) suggested that hydrophobic bonds might play a crucial role in the thermal stability of proteins. However, studies of proteins from thermophilic organisms (27) and calculation of the hydrophobicity index, based on amino acid composition, relative to the thermal stability of a variety of proteins (28, 29) have demonstrated that no apparent relationship exists.

Thus, the extreme hydrophobicity of integral membrane proteins is probably not due to the need for thermal stability of the membrane. Rather, hydrophobicity is probably the mechanism whereby the protein can successfully merge with or span the lipid bilayer (8, 15, 17). To incorporate with an apolar matrix, integral membrane proteins must be amphipathic, i.e., they must contain polar and apolar regions in their primary structure. For example, glycophorin from the erythrocyte membrane has been recently shown to have an amphipathic nature whereby the hydrophobic mid-portion of the protein spans the lipid matrix such that the amino terminal portion of the molecule is exposed to the exterior of the cell, while the carboxyl terminal segment extends towards the interior of the cell (20, 30, 31). The cytochrome b_5 protein has a very hydrophobic peptide, which is firmly bound to the lipid bilayer of the membrane (18). However, no evidence exists to suggest that this protein spans the lipid bilayer. Rather, it apparently is anchored in the membrane by its hydrophobic peptide (19). This type of partial penetration of the hydrophobic portion of a protein is also found for NAD cytochrome b_5 reductase (19). These

observations suggest that hydrophobic bonding must play a crucial role in lipid-protein interactions (8, 12, 13, 14, 15, 17, 20).

1.4 Structural Analysis of Integral Membrane Proteins. Analysis of the primary structure of many integral membrane proteins should determine if these proteins are commonly amphipathic. The approximate secondary structure of the apolar regions can be determined using eith CD or recent predictive methods requiring only an amino acid sequence (32). The latter method, devoid of computer calculations, makes use of empirically derived rules for predicting the initiation and termination of helix and sheet regions in proteins with 85% certainty. Other methods for predicting the occurrence of secondary structure in proteins had been attempted in the past, however, these were primarily concerned with helix determination and their reliability was poor (33, 34, 35, 36, 37, 38). Nevertheless, the ability to predict the secondary structure of proteins will provide a basis for devising molecular models for the hydrophobic interactions between integral membrane proteins and the lipid bilayer in a membrane.

1.5 Description of the Membrane System Used in This Study: The Gas Vacuole Membrane. Since the determination of secondary structure in proteins relies on the primary structure of the molecule, sequence analysis of integral membrane proteins is necessary. In this thesis, attention has been focused on one particular membrane system, the gas vacuole membrane from Microcystis aeruginosa Kuetz. emend Elenkin.

This membrane was chosen because it is simple. It contains only one protein species, no carbohydrate, and no lipid (39). In addition, the protein contains a high proportion of hydrophobic, but no sulfur-containing amino acids (39). Yet, the gas vacuole membrane functions as a semi-permeable barrier whereby only gases are permitted to diffuse through the membrane (40). Two functions have been attributed to the membranes: providing bouyancy to the cells (40,41), and light shielding of the photosynthetic lamellae (42).

The molecular weight is reported as 14,000 (39), however, there remains some confusion about this value. Preliminary X-ray data have suggested that a repeating "unit cell" (molecular weight 7,800) might exist (41). On the other hand, polyacrylamide gel electrophoresis at pH 2 and 4 indicated one 14,300 molecular weight species (39), whereas SDS disc gel electrophoresis at pH 8.5 indicated a 21,500 molecular weight species (43). Together, these results could indicate that a 7,000 molecular weight species exists which aggregates during polyacrylamide gel electrophoresis.

Unpublished X-ray data have indicated an asymmetric electron density profile (41). This was interpreted to mean that hydrophobic amino acids, which are less electron-dense than the hydrophilic amino acids, predominate on the inner (gas-facing) surface, whereas the hydrophilic amino acids predominate on the outer (cytoplasm-facing) surface (41). In addition, several prominent wide-angle reflections were observed which characterize an extensive β·structure: a feature supported by CD analysis (44). Based on this, it was proposed that two layers of cross β-structure along each rib might account for the

rigidity of the membrane. Blaurock and Stockenius (41) have also suggested that this cross β-conformation would be such that along any rib, molecules would be joined by hydrogen bonding between the backbones of the chains. This last conjecture on the presence of intrarib hydrogen bonding is independently supported by others (43, 45). No X-ray data are presented suggesting either the presence or absence of helix in the membranes. However, IR (46) and CD (44) analyses have indicated that some helix is present.

Analyses of the GVP from other species have been reported (47, 48). Chemical information beyond a molecular weight, and amino acid composition remains incomplete. For a more thorough discussion on the structure and function of gas vacuoles, see Walsby's review (41).

1.6 Implications of Sequence Data. Sequence analysis, though often difficult, can provide not only the primary but also secondary structure of a protein. In addition, accumulation of the primary structures of many related proteins e.g. the cytochromes c, provides information for devising phylogenetic relationships among different organisms (49). The genetic and evolutionary significance of changes which occur in the primary and secondary structures of various proteins is discussed in several reports and will not be considered here (49, 50, 51, 52).

1.7 Techniques Available for Sequence Analysis. Hydrophobic proteins are especially difficult to sequence due to their insolubility; up to now, only one hydrophobic protein, glycophorin, has been sequenced (30). If peptides can be generated from a protein for sequence

analysis, many methods are available to assist in the determination of the actual structure. These methods can be enzymatic, chemical or physico-chemical.

Enzymatic methods are those which use either carboxpeptidases A, B, or C (of which type C is the best (53)), or aminopeptidases. These enzymes sequentially degrade a peptide from either the carboxyl or amino terminal end. Providing the sequence contains no repeating residues and the peptide is very pure, these methods work. Another enzymatic approach to sequence analysis makes use of dipeptidyl aminopeptidase I (DAPI) which essentially cleaves peptides from their amino terminus into dipeptides (54). Removal of the amino terminus of the peptide and subsequent retreatment with DAP I produces an overlapping set of dipeptides. The dipeptides are then separated and sequenced using chemical methods or mass spectroscopy (55). Major disadvantages of the method are the difficulty of separation of dipeptides, and limitation on the size of peptide to be sequenced (maximally 20-30 residues) (54).

Chemical methods of sequence analysis are extensive and only the most recent techniques will be mentioned here (56). All of the most effective methods use the Edman degradation (57). Problems arise, however, in losses of the peptides being sequenced to the organic wash phase, termed washout. Techniques have been devised to overcome this problem by making the carboxyl terminus of the peptide more hydrophilic. If the peptide in question is a tryptic peptide, those terminating in lysine can be modified by using 4-sulfophenylisothiocyanate which is specific for primary amino groups (58,59). For peptides which terminate with other amino acids, modification is

effected with 2-amino- 1, 5-naphthalene disulfonic acid (ANS) using a water soluble carbodiimide (EDC) as the coupling agent (60). The primary amino groups are kept protonated so the ANS, being essentially uncharged, serves as the nucleophile. Polymerization of the peptides is, thus, avoided while amide linkages are formed at the free carboxyl groups. Utilization of these methods, which I have modified, reduces losses of peptides to washout.

Another method which is relatively new is solid phase sequencing (61, 62). This procedure involves attaching the carboxyl terminus of the peptide to an insoluble matrix which then permits sequence analysis without the problem of washout. However, only recently has this technique been developed for practical use.

1.8 Objectives of this Study. The basic objectives which I wished to accomplish in this study were threefold. First, I would like to establish whether or not the gas vacuole protein is, indeed, a true membrane protein. A second objective, which might be expected to follow from amino acid sequence data, is the development of a molecular model for the diffusion of gases into the membrane.

My last objective concerns the hydrophobicity of the membrane. An analysis of the primary structure of glycophorin (30) has provided valuable information on hydrophobic proteins while generating the potential for studying hydrophobic interactions between integral membrane proteins and lipids. It is the purpose of this investigation to increase our knowledge of the hydrophobic properties of the GVP so that it might be used by other workers in the field as a model system in which hydrophobic interactions can be studied.

MATERIALS AND METHODS

2.1 Culture, Harvest, and Lysis. The strain NRC-1 of the Cyanophyte, Microcystis aeruginosa Kuetz. emend. Elenkin, obtained from P. R. Gorham,was grown in ASM-1 medium (63) as described by Jones and Jost (39) to a final density of 10^7 cells/ml. After incubation with benzylpenicillin, the cells were collected with an Amicon CH3 ultra-filtration concentrator fitted with an HICP10 fiber filter. The concentrate was then osmotically lysed (39).

2.2 Purification of Gas Vacuole Membranes. The procedure used is summarized in Figure 1. Immediately after lysis, the lysate, fraction A, was centrifuged in a Sorvall Type SS-34 angle head rotor at 270 x g overnight; for the last 15 minutes of the run the rotor was accelerated to 3,000 x g. The pale blue-green layer, fraction B, at the meniscus of the centrifuge tubes was drawn off and subsequently filtered through a series of Millipore filters of decreasing pore size: 5 μm, 1.2 μm, 0.8 μm, 0.65 μm, 0.45 μm. The filtrate, fraction C, was then processed on a 5 x 50 cm column of Sepharose 4B in 0.01 M Tris-HCl buffer, pH 7.5 containing 0.01 M NaN_3. The milky-white fraction D from the column was layered in 2.5 ml portions onto 9.5 ml of 0.4 M NaCl in centrifuge tubes and spun for 30 minutes at 200,000 x g in a Spinco SW-41 rotor. The narrow white band at the meniscus of the tubes was removed and stored at $4^{\circ}C$ in the presence of 0.01 M NaN_3. The concentration of this purified gas vacuole membrane protein, fraction E, was determined by the absorbancy at 400 nm (45).

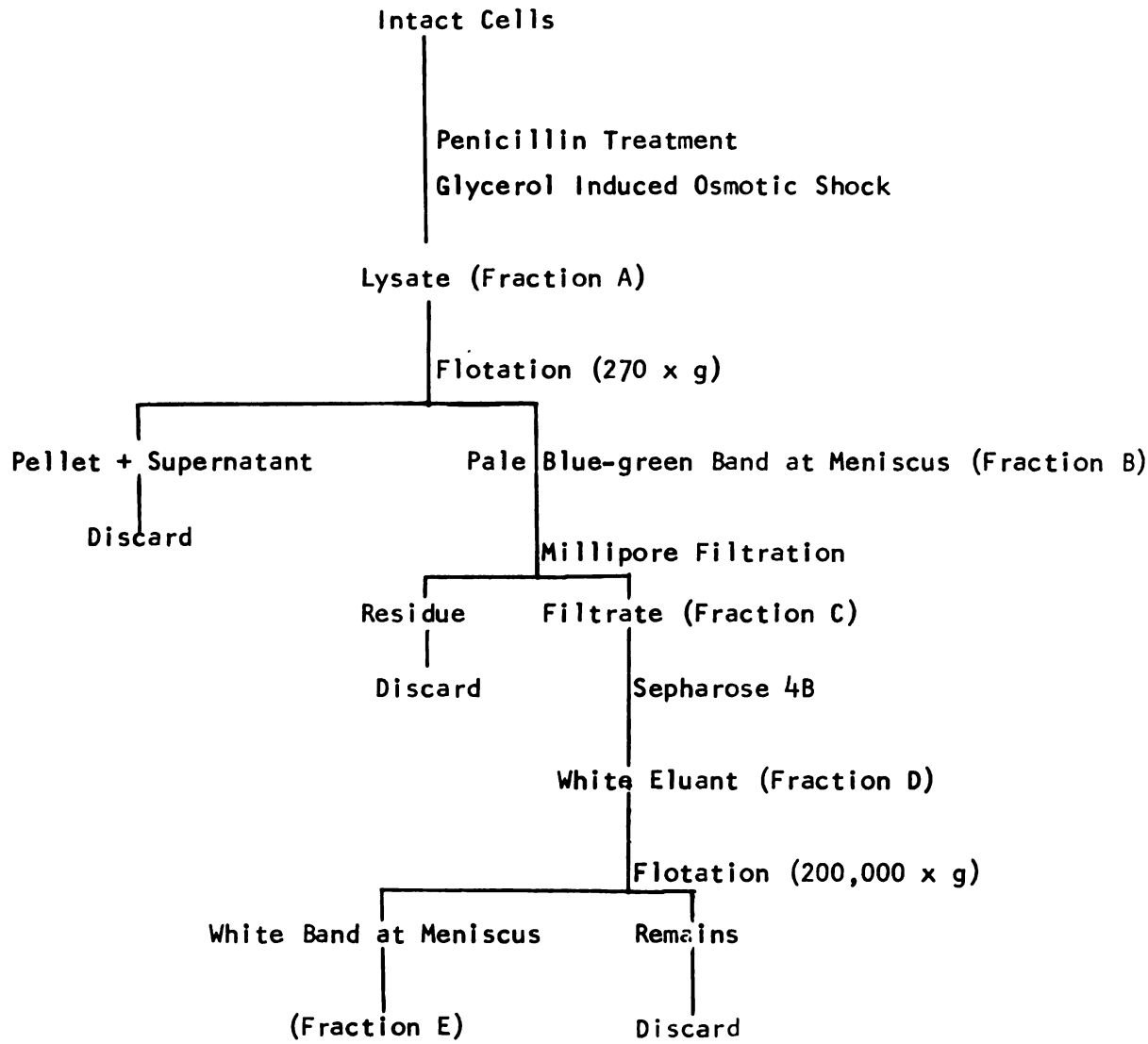


Figure 1. Fractionation Scheme of the Cell Free Preparation of Gas Vacuole Membranes.

2.3 Estimate of the Purity of Fraction E. Cells from a late log phase culture were innoculated into 500 ml of ASM-1 medium to a final density of 6 x 10^6 cells/ml. The $MgSO_4$ in the medium was replaced by $MgCl_2$. To this was added 0.04 mM $Na_2{}^{35}SO_4$ (specific activity 27.8 mC/m mole) and 0.16 m M $Na_2{}^{35}SO_4$. Cells were then grown, harvested, and purified as described previously. At each step of the purification procedure 0.5 ml aliquots were assayed for protein by the method of Lowry et al. (64), and for radioactivity. The scintillation fluid used consisted of 907 ml dioxane, 100 g naphthalene, and 5 g PPO.

2.4 Preparation and Separation of Tryptic Peptides. The procedures used are outlined in Figure 2. Dialysis tubing was prewashed with 88% formic acid. One volume of fraction E containing 100 mg of gas vacuole protein (GVP) was dialyzed against two volumes of 88% formic acid at room temperature. This acid denatured protein was then dialyzed exhaustively against deionized distilled water until the washing fluid reached pH 3. The acid precipitated protein in 0.002 M $CaCl_2$ at pH 7.8 was then trypsinized for 48 hours at 25° C with two aliquots each of 4 mg of trypsin (specific activity 195 U/mg, Worthington). The first aliquot of trypsin was added at the beginning, the second, 18 hours after the start of the digestion. The pH of the reaction mixture was maintained with 0.1 N NaOH delivered by a pH stat or manually. At the end of trypsinization, the volume of the incubation mixture was reduced to dryness by flash evaporation at 40° C. An aliquot of the digestion mixture was dansylated (65),

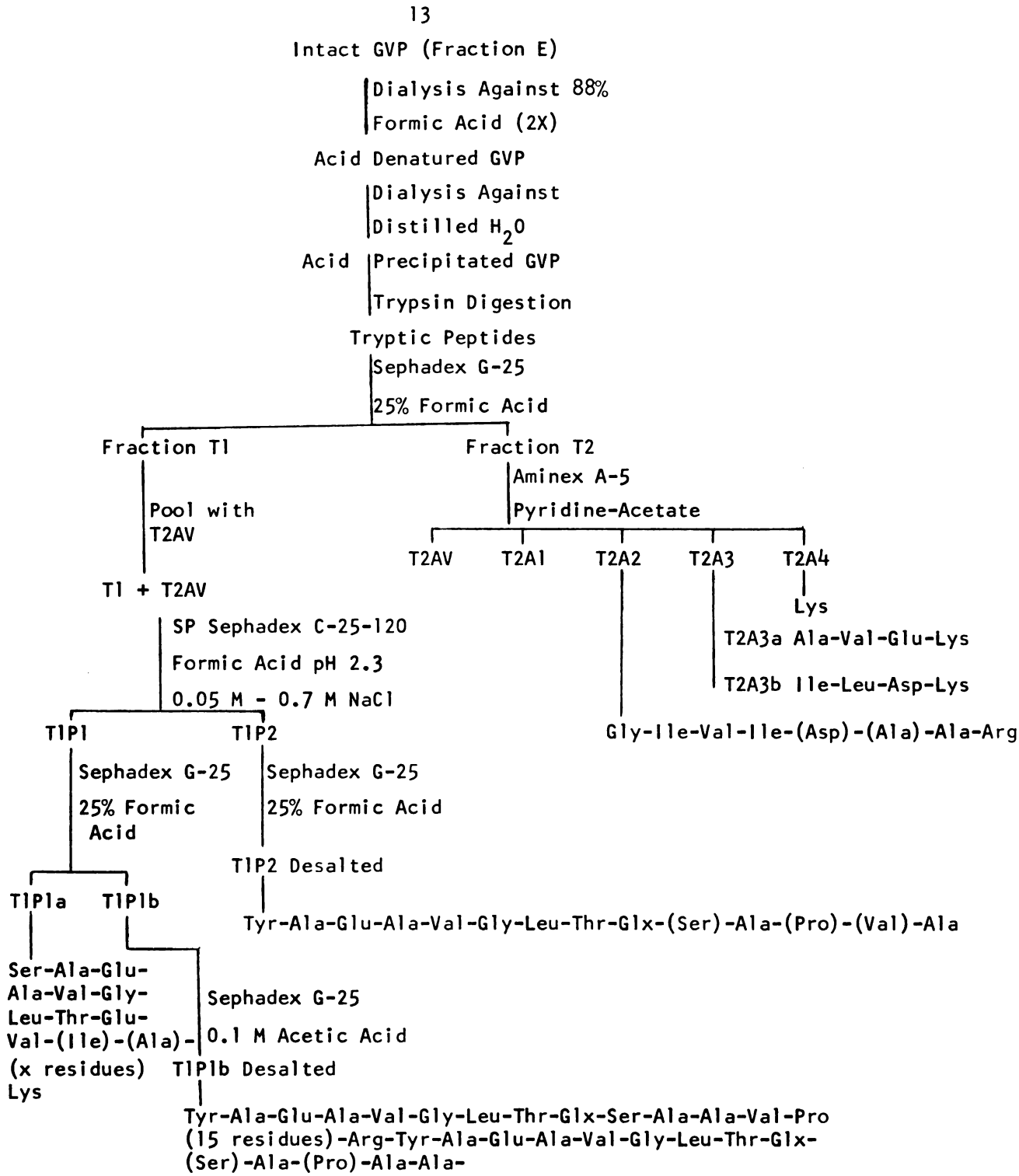


Figure 2. Fractionation Scheme of the Preparation and Purification of the Tryptic Peptides from Gas Vacuole Protein (GVP)

acid hydrolyzed, and electrophoresed at pH 1.9 (formic-acetic acid) at 5 kV for 135 minutes. The undansylated material was suspended in 2 ml of 25% formic acid and fractionated on a column (2 x 100 cm) of Sephadex G-25 (fine) in 25% formic acid (20). The different fractions were monitored for the presence of protein either at 280 nm or with ninhydrin at 570 nm. The latter assay was done with 15 µl aliquots in a Technicon Autoanalyzer. The major fractions from the Sephadex column were designated T1 and T2.

Fraction T1 plus the fraction T2AV (See Fig. 2) were pooled and purified on a column (2 x 25 cm) of Sulpho Propyl Sephadex C-25-120 (SP Sephadex fine) using a biphasic 0.05-0.7 M NaCl gradient in a formic acid-water mixture at pH 2.3.

T2 was subsequently separated on a column (1 x 25 cm) of Aminex A-5 (Bio-Rad) using a pyridine-acetate buffer system (66). The chambers of the gradient maker were loaded as follows:

	pyridine	acetic acid to pH	vol. in chamber (ml)
Chamber 1 had	0.068 M	2.7	70
2	0.068 M	2.7	70
3	0.177 M	3.1	70
4	0.177 M	3.1	70
5	0.177 M	3.1	70
6	0.177 M	3.1	70
7	2.0 M	5.0	70
8	2.0 M	5.0	70

The eluant profiles from both the Aminex A-5 and the SP Sephadex columns were monitored for the presence of protein as described above. Fractions from the SP Sephadex column were desalted over Sephadex G-25 (2 x 100 cm) in either 25% formic acid or 0.1 M acetic acid at pH 2.87.

2.5 Maleylated Tryptic Peptides. The procedure used is outlined in Figure 3 under A. Fraction E was acid denatured as described above. The ε-amino groups of the lysines of the water-suspended GVP were maleylated according to Butler et al. (67) (See Figure 4) and then trypsinized as described above. Subsequent chromatography was over a Sephadex G-25 column (2.5 x 100 cm) in 0.1 M acetic acid at pH 2.87 which effectively demaleylated the peptides. The different fractions were monitored for the presence of protein with ninhydrin at 570 nm and the T1 and T2 fractions were pooled and further separated on Aminex A-5 and/or SP Sephadex using the previously described gradients for separation of the tryptic peptides.

An alternative method of preparing maleylated tryptic peptides was that of irreversibly denaturing the protein as shown in Figure 3 under B. This was accomplished by heating the acid denatured protein to 105° C for 30 minutes and maleylating the material as described above. Trypsin treatment was performed as described above and the digest was chromatographed over Sephadex G-25, G-50, and G-75 in 0.01 M Tris-HCl buffer at pH 8.5. Fractions brought to pH 5 were analyzed with ninhydrin at 570 nm.

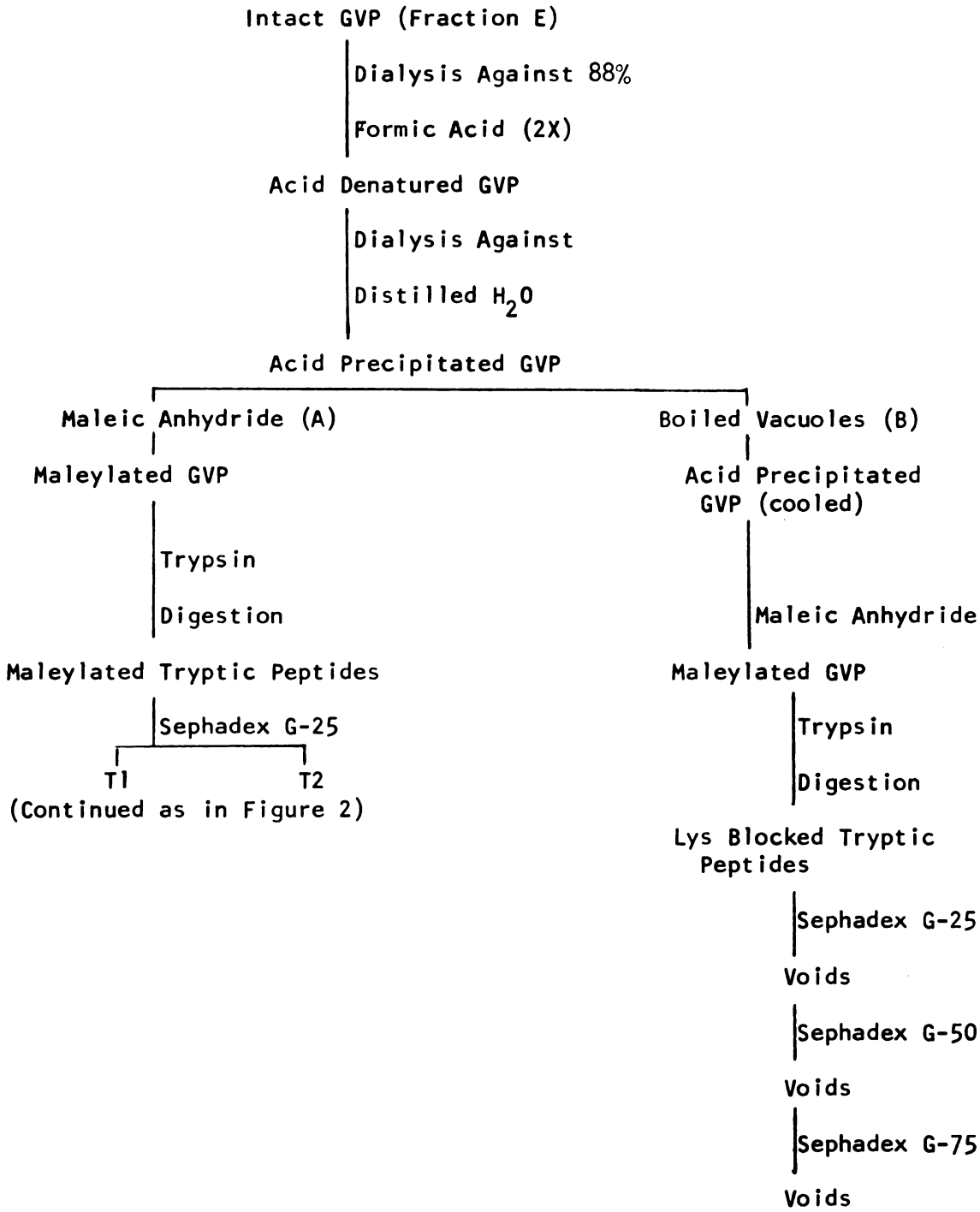


Figure 3. Fractionation Scheme of the Preparation and Purification of the Lysine Blocked Tryptic Peptides of GVP.

Figure 4. Maleylation.

Maleylation

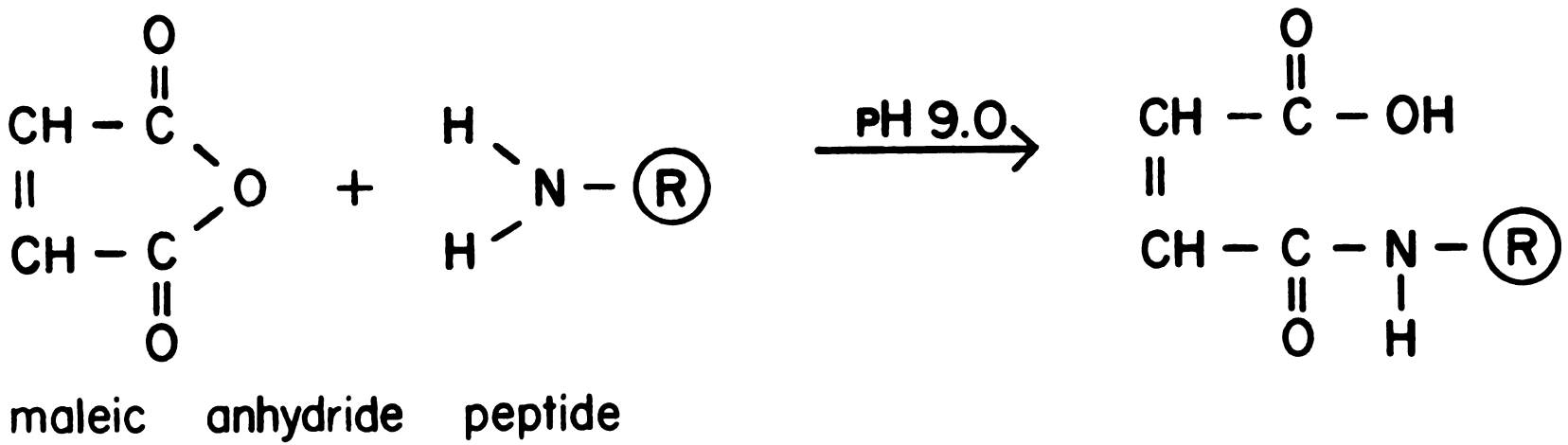


Figure 4.

2.6 N-Bromosuccinimide Peptides. The procedure used is outlined in Figure 5. One hundred mg of Fraction E was flash evaporated to dryness, resuspended in 88% formic acid, and stirred for 30 minutes. The acid denatured protein was again dried as above and then resuspended by cavitation in 27 ml of 50% glacial acetic acid. To this suspension was added 675 μM N-Bromosuccinimide (NBS) in 27 ml of 100% glacial acetic acid (68). After stirring at room temperature overnight, the reaction mixture was flash evaporated and the entire NBS procedure was repeated with a reaction time of only 4 hours. It was then flash evaporated and applied to a column (2.5 x 100 cm) of Sephadex G-25 in 0.1 M acetic acid, pH 2.87. The eluant profile was determined at 254 nm. NBS elutes with the monomer volume, Vm.

The soluble NBS peptides were divided into three separate fractions: N1, N2, and N3; these were rechromatographed over Sephadex G-25 in 25% formic acid and analyzed for protein at 254 nm.

The insoluble fraction NPT was eluted from the top 2 cm of the Sephadex G-25 column by successively washing the beads on a scintered glass filter (M) with both 88% and 25% formic acid. The filtrate which appeared flocculent was flash evaporated and again treated with NBS (as above) for 4 hours. The NPT peptide was effectively separated from the soluble peptides on Sephadex G-25, as before, eluted from the beads, dried and reacted with NBS for a third time.

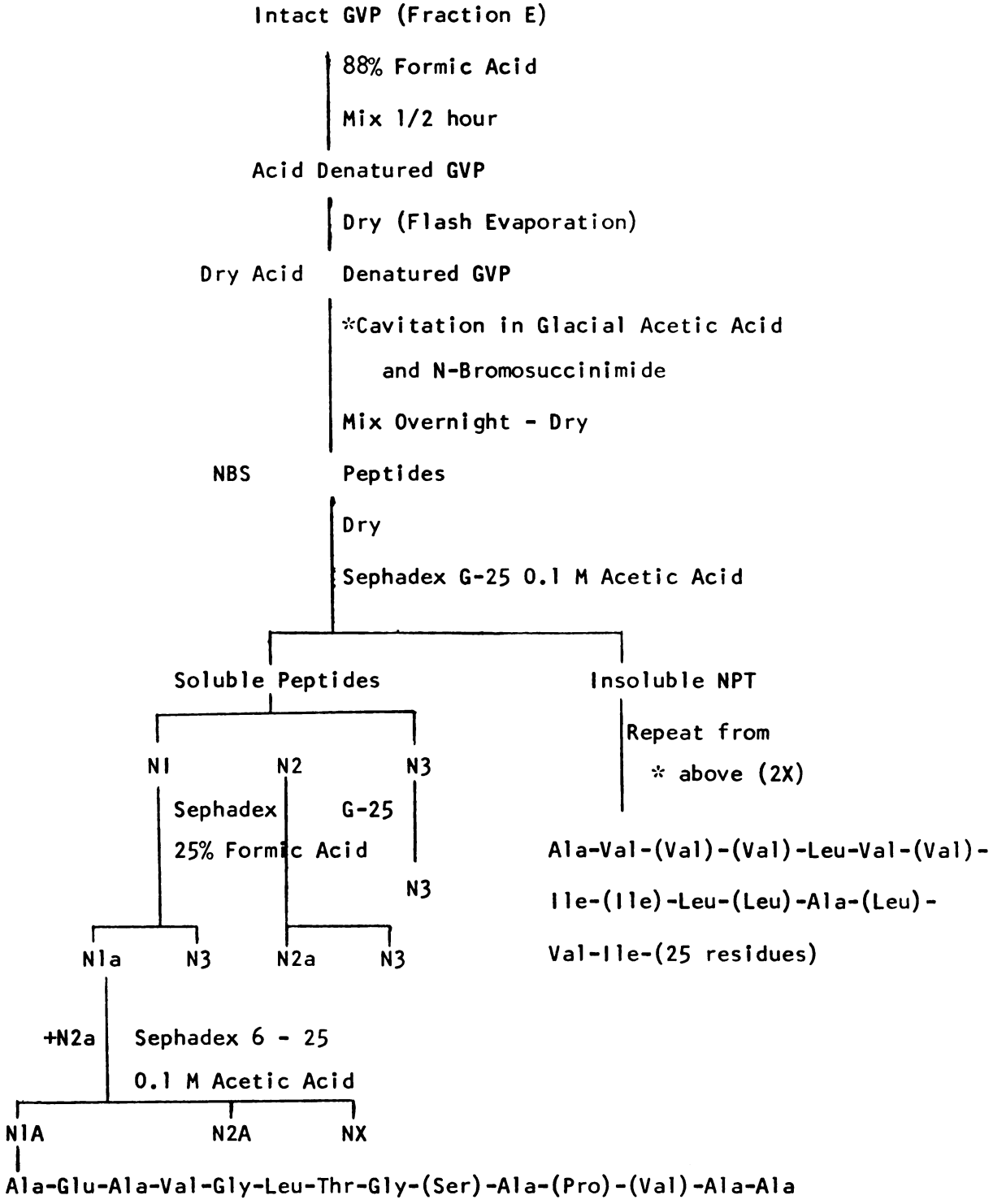


Figure 5. Fractionation Scheme of the Preparation and Purification of the N-Bromosuccinimide (NBS) Peptides of GVP.

2.7 Digestion of the NPT Peptide by Thermolysin. The incubation mixture consisted of 100 nm of NPT peptide, 100 nm of norleucine as an internal standard, and 100 µl of Tris-HCl buffer 0.05 M, pH 8.0, containing 0.002 M $CaCl_2$. Incubation was at 40° C after addition of 0.05 mg thermolysin (Calbiochem). Aliquots were removed at 0, 1, 3, 5, and 12 hours and the reaction stopped by the addition of glacial acetic acid to a pH below 2. All aliquots and standard amino acids were simultaneously subjected to electrophoresis at pH 1.9 in a formic acid-acetic acid buffer system (69) in order to measure the extent of digestion. Subsequently a large sample of NPT (3 µmoles) was digested with equivalent increases in reagents for 10 hours. The reaction mixture was chromatographed on a column of Sephadex G-25 in 25% formic acid. The insoluble NPT was removed from the column as described in the preparation of the NBS peptides.

The soluble fraction produced an eluant profile which was analyzed with ninhydrin at 570 nm after NaOH hydrolysis (70). Polypropylene tubes were used as glass releases silicate which interfered with the analysis. To 1-2% of the volume of each fraction, ordinarily corresponding to 15-30 µl, was added 200 µl of 5 N NaOH. All samples were autoclaved at 121°C for 1 hour, and then neutralized with 200 µl of 5 N HCl, and then brought to 1 ml with pH 6.1 citrate buffer 0.04 M Ninhydrin analysis at 570 nm followed.

2.8 Partial Acid Hydrolysis of the GVP and Peptides. Partial cleavage of fraction E of the GVP was attempted with 0.03 N HCl at 105° C for 48 hours in a nitrogen atmosphere (71). The rate of the reaction was followed by monitoring the release of free aspartic acid. The same hydrolysis procedure and assay method were used on peptide NPT and the hydrolysate chromatographed over Sephadex G-25 in 0.1 M acetic acid. Fractions were monitored for protein at 254 nm.

To prepare fragments of the TIP2 peptide, 2 μmoles of the peptide were hydrolysed with 200 μl of 6 N HCl at 105° C for 5 minutes. The hydrolysate was rapidly dried under a stream of nitrogen at 25° C and electrophoresed at pH 1.9. Peptides were eluted with 0.1 N acetic acid, dried, and stored in 25% formic acid at -20° C.

2.9 Recovery of Peptides (Balance Sheet). The percentage of recovery of peptides in my work is based on the amount of lysine and arginine applied to the column vs. the amount recovered in the various fractions. Aliquots from the pooled fractions were assayed for lysine and arginine on a Technicon Amino Acid Analyzer.

2.10 Purity of Peptides. High voltage electrophoresis at pH 1.9 and/or pH 6.5 (69) was used in conjunction with the cyanoethylation procedure of Fletcher (72) (Figure 6) and amino acid analysis in order to determine the homogeneity of the purified peptides. Both procedures were necessary as often very large peptides and peptides of

Figure 6. Cyanoethylation.

Cyanoethylation

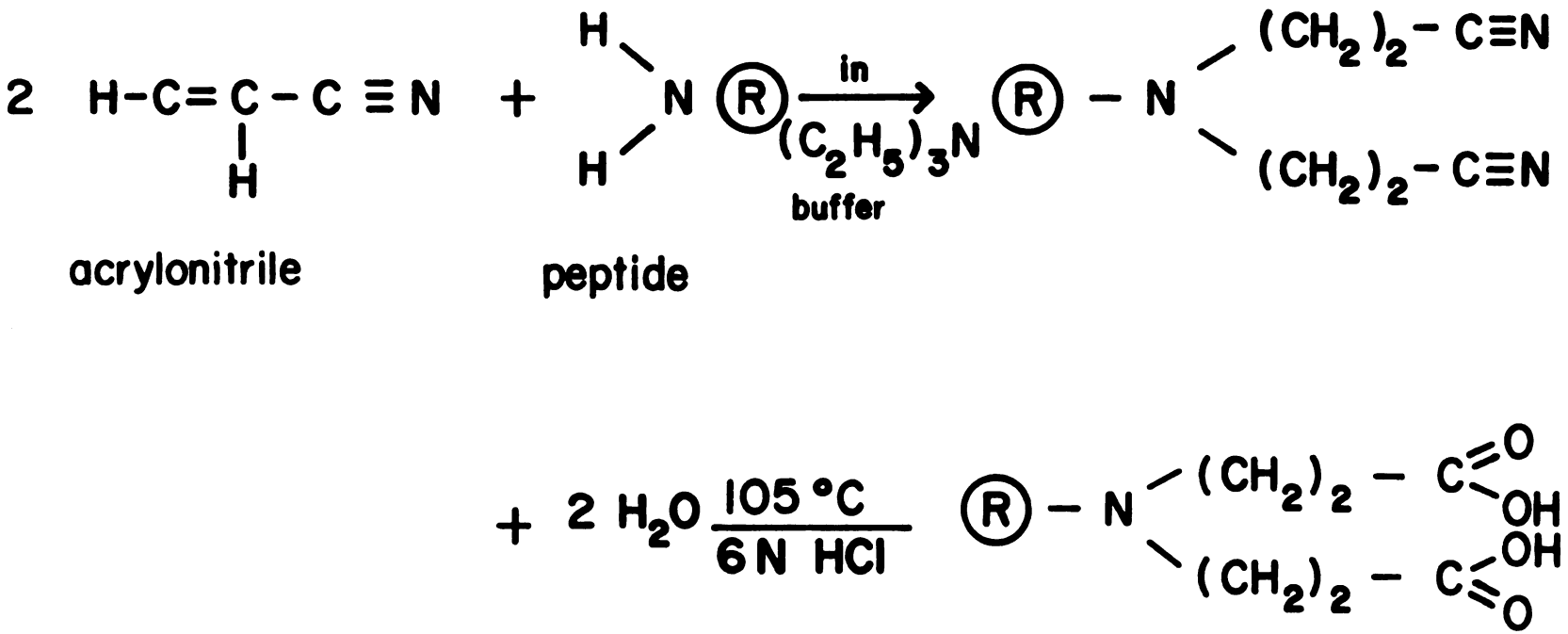


Figure 6.

very low yield could not be visualized by a ninhydrin assay after electrophoresis. Cyanoethylation was sensitive to a minimum of 5 nanomoles of peptide.

2.11 Sequence Analysis of Peptides by Carboxypeptidase C. Certain peptides were sequenced from their C-termini using carboxypeptidase C (Henley and Co.) using the method of Tschesche and Kupper (53). Norleucine was the internal standard. The reaction was stopped by adding 0.1 N HCl to pH 2 and then the mixture was heated to 80° C for 1 minute. The samples were then centrifuged to remove the enzyme. The supernatant liquid was removed and analyzed on a Technicon Amino Acid Analyzer.

2.12 Amino Acid Analysis. Samples were hydrolyzed in evacuated 1 ml microvials (Precision Sampling) containing 200 µl of "6 N" (constant boiling) HCl at 105° C for 18 hours. The solution was then brought to dryness at 50° C under a stream of nitrogen and the residue resuspended in citrate buffer pH 2.875 according to Moore et al. (73). Amino acid analyses were performed on an automated Technicon Amino Acid Analyzer by the accelerated method using Chromobeads C-2 (74). The chromatograms were integrated by an Autolab System IV Computing Integrator. Based on the retention times relative to the internal standard, norleucine, each amino acid was identified and the corresponding correction factors were applied automatically for the individual color reaction of each amino acid with ninhydrin.

2.13 Automated Sequencing Methods. Peptides and proteins were sequenced on a Beckman Sequencer Model 890C, except where otherwise

noted. The various methods used to sequence the intact gas vacuole protein via automated Edman degradation (Figure 7) are given in Table 1. All peptides were sequenced using the dimethylallylamine-trifluoroacetic acid buffer in a pyridine-water mixture (DMAA) with peptide program #021572 as diagramatically represented in Figure 8. When possible the peptide carboxyl terminal modification method of Braunitzer (59) (Figure 9) employing sulphophenyl isothiocyanate (SPITC) was used.

All other peptides were sequenced using my improved version of the method of Foster et al. (60) (Figure 10). This method was developed as indicated in Table 2 by using a synthetic tripeptide (DL-Leu-Gly-Gly).

Solution A: 300 nM peptide in
600 µl of water, adjusted to pH 4.0.

Solution B: 5 mg EDC (N-ethyl, N'-(-3-dimethylaminopropyl carbodiimide) HCl and 8 mg ANS (2-amino-1, 5 naphthalene disulfonic acid) in 1 ml water, adjusted to pH 4.0.

Solution C: 5 mg EDC in 1 ml water, adjusted to pH 4.0.

Add to Solution A 20 µl of Solution B. Mix for 3 hours, then add 20 µl of Solution C. Mix for 1 hour or longer at room temperature. Use 20 µl of Solution C repeatedly for poorly solubilized peptides. At the end of the reaction, add 50 µl glacial acetic acid. Apply to reaction cup on sequencer in amounts of less than 400 µl. Dry via the Beckman sequencer subroutine #02772.

Table 1

Methods and Results of Sequence Analyses

of Gas Vacuole Protein (Fraction E)

Treatment +	Treatment -	Acid denat.	Heat denat.	SPITC	DMAA #050771	Quadrol #02672	Sequence Results
A		+	-	-	-	+	Ala-Val-Glu
B		+	-	+	-	+	Gummy residues
C		+	-	+	+	-	Ala-Val
D		+	+	+	+	-	Ala-Val
E		+	+	-	+	-	Ala-Val
F		+	+	+	-	+	Ala-Val
G		+	+	-	-	+	Ala-Val

Figure 7. The Edman Degradation for Sequence Analysis of Polypeptides.

Edman Degradation

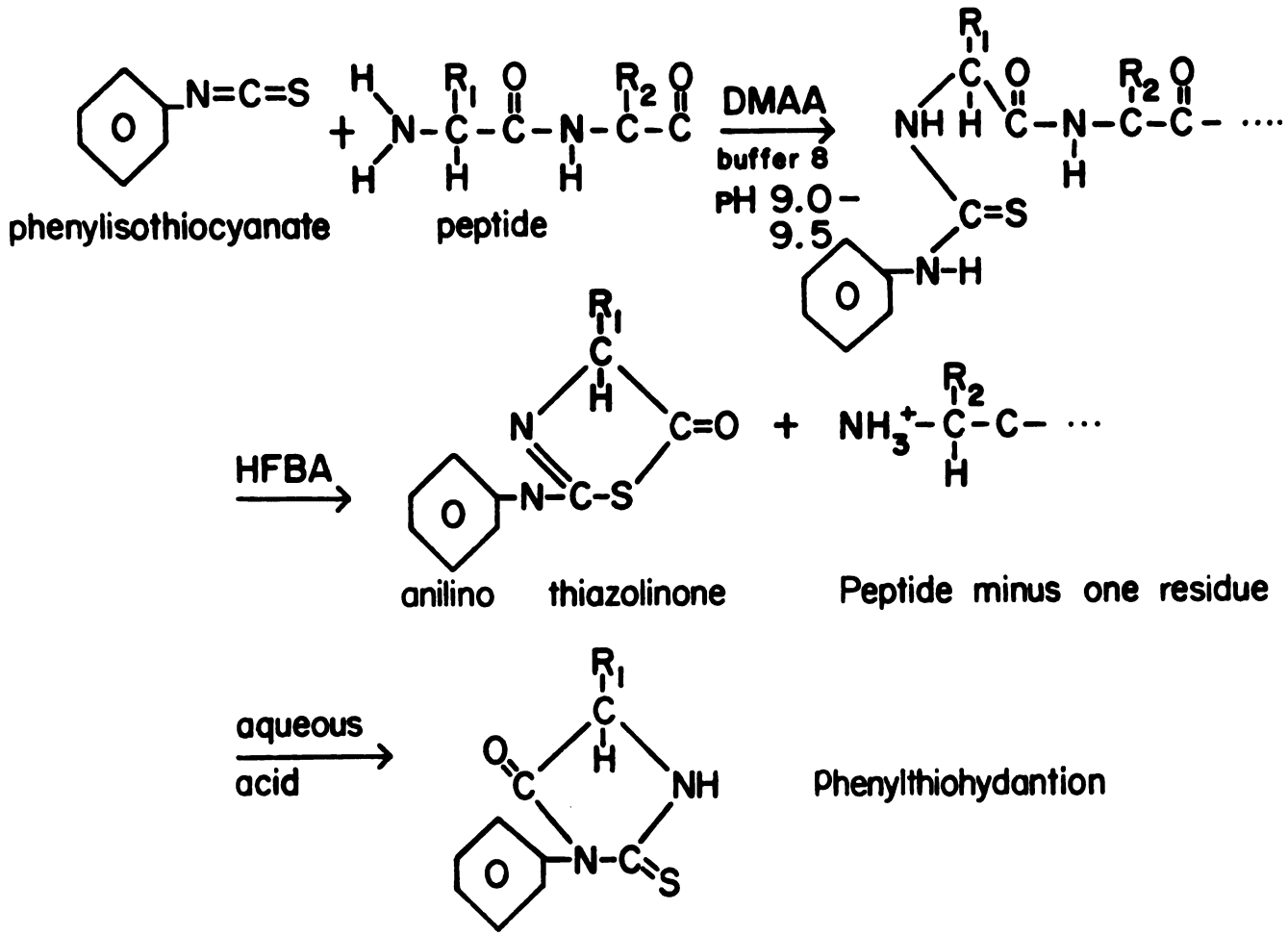


Figure 7.

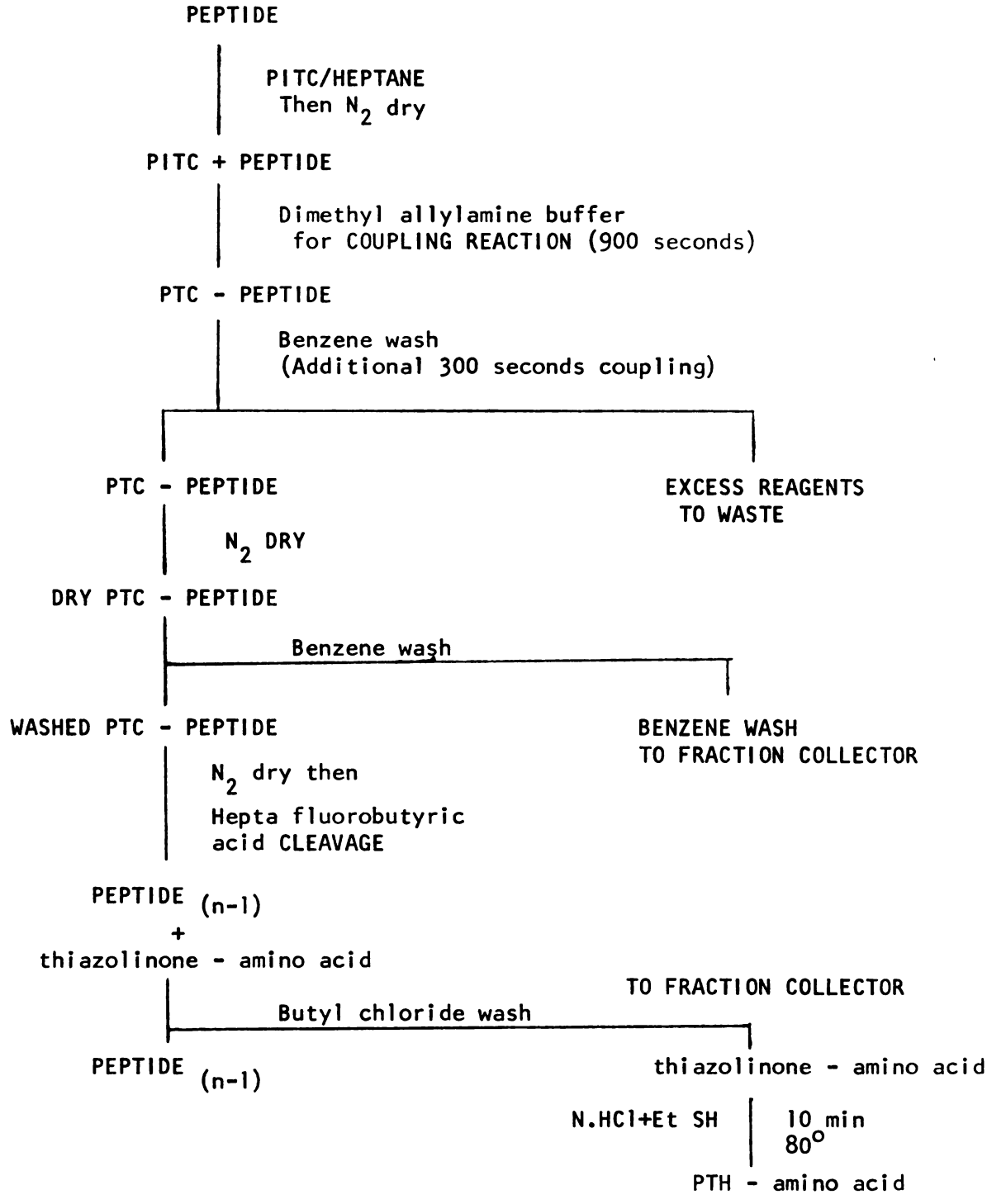


Figure 8. Operational Scheme for Peptide Program #021572 for Beckman Sequencer Model 890 C.

Figure 9. Peptide Modification Using Sulphophenyl Isothiocyanate (SPITC).

Braunitzer Method of Peptide Modification

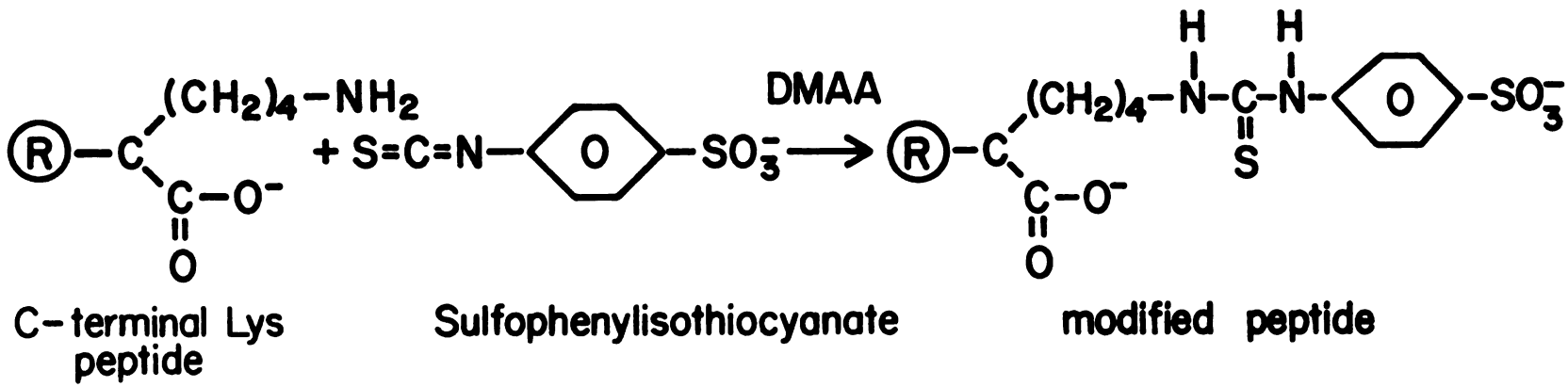
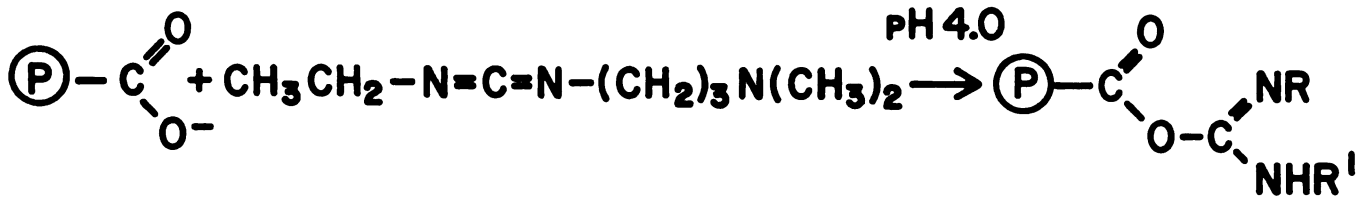


Figure 9.

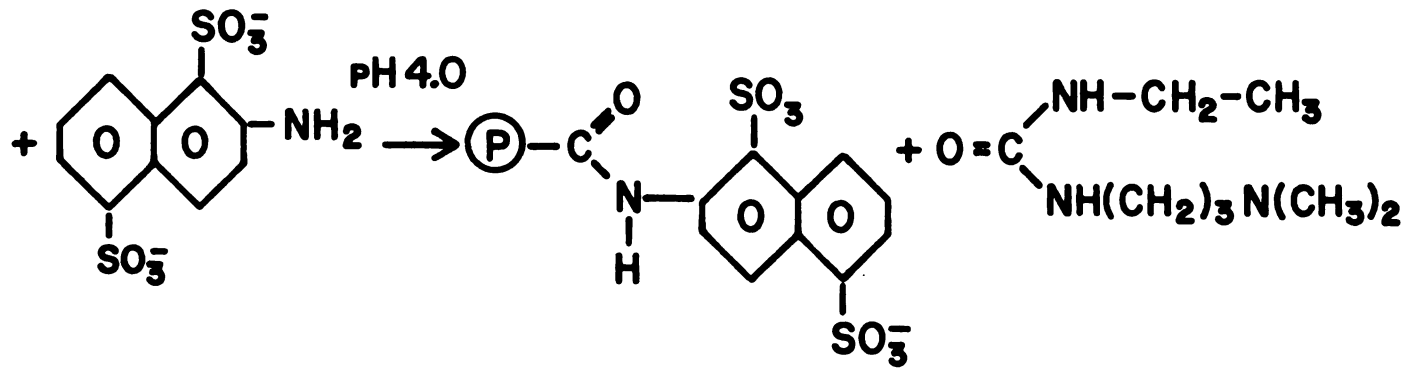
Figure 10. Peptide Modification by Carbodiimide and Amino Naphthalene Sulfonic Acid.

Foster Method of Peptide Modification



Peptide

N-ethyl, N'-(3-dimethylaminopropyl) Carbodiimide HCl



2-amino-1,5 naphthalene disulfonic acid

Modified peptide

Figure 10.

Table 2

Improved Peptide Modification Procedure

Treatment Number	Peptide Leu-Gly-Gly	ANS		EDC		pH	Additions	Treatment Fig. 27
		Amount	Time of Addition	Amount	Time of Addition			
	μM	μM	hr[a]	μM	hr[a]			
1	.3	0	0	0	0	4	-	a
2[b]	.3	.6	0	.6	0	4	-	e[b]
3	.3	.6	0	.6	0	3	-	d
4	.3	1.5	0	.6	0	4	-	b
5	.3	.3	0	1.5	0	4	-	f
6	.3	.3 .3	0 3	1.5 1.5	0 3	4 4	- -	g
7	.3	.6	0	.6	0	4	0.1 M NaCl	c

[a] The total reaction time was 4 hrs. The volume of the reaction mixture was between 600 and 1000 μl.

[b] The method of Foster et al. (60).

2.14 Identification of Amino Acid Derivatives from Sequence Analysis. All PTH amino acids were formed from the thiazolinone as follows (75):

> Add 0.2 ml of 1.0 N HCl containing 1 µl/ml ethanethiol to dried fractions. Blow nitrogen over solution for 30 seconds and seal with silicone rubber stoppers. Mix by cavitation. Heat to 80° C for 10 minutes. Cool quickly. Extract PTH with 2 volumes each of 0.8 ml of ethyl acetate.

With the exception of histidine and arginine, which is destroyed by gas-liquid chromatography (GLC), all PTH amino acids were analyzed using GLC employing the SPOV column matrix of Pisano et al. (76). The temperature program used was:

> Initial temperature: 170° C.
>
> Delay: 2 minutes.
>
> Rate: 6°/minute.
>
> Final temperature: 280° C, hold at this temperature for 12 min.

The on-column method of PTH-TMS derivatization was performed with N,O-bis-trimethyl-silyl-acetamide (BSA) in order to identify the more difficult PTH amino acids. However, due to difficulties concerning very low yields in vacuole peptides, Asn, Gln, Ser, and Thr were identified after hydriodic acid hydrolysis (120° C, 18 hours, N_2 atmosphere) according to Smithies et al. (77) and Inglis et al. (78). The hydrolysates were then analyzed for amino acids using the previously described citrate buffer system.

RESULTS

3.1 Purity of Gas Vacuole Membranes. The purification of the gas vacuoles (2.2) was monitored by growing cells in the presence of ^{35}S and by assaying the cell free fractions (Figure 1) for their specific activity (2.3) based on the fact that gas vacuoles do not have any sulfur containing amino acids. Table 3 summarizes the results and shows that Fraction E, which was used for peptide preparations, has a relative purity of 99%.

3.2 Amino Acid Composition of Gas Vacuole Protein. The amino acid composition of gas vacuole protein (GVP) has been reported before (39). As an improved analysis technique became available, I repeated these determinations, the results of which are summarized in Table 4. The relative amount of some amino acids changed slightly; noteworthy are alanine, valine, isoleucine, and leucine and the occurrence of proline.

3.3 Sequence Analysis of GVP. A common procedure in sequence analysis using the Edman degradation is to obtain an amino terminal sequence of the intact protein: sixty consecutive residues were obtained from apomyoglobin D of the humpback whale (75). I tried such an approach with GVP as shown in Table 1: however, a maximum of 3 residues Ala-Val-Glu-...was obtained. The possibility of protein loss from the reaction cup due to washout, was ruled out by amino acid analysis of the cup contents.

Table 3

Purity of Cell Free Fractions

Containing Gas Vacuole Membranes

Fraction assayed (Fig. 1)	cpm per 0.2ml[a]	Protein µg/0.2ml[a]	Specific Activity cpm/ µg	Relative Contamination
A	225092	550		
B	317424	430	738	100%
C	90865	150	605	81%
D	5761	60	96	13%
E	416	50	8	1%
F	77	35	2	0.2%

[a] Average of two determinations.

Table 4

Amino Acid Composition of GVP

Amino Acid	mole%[a]	mole%[b]	mole%[c]	mole%[d]
Aspartic Acid	6.3	6.4	5.6	11.6
Threonine	4.5	5.2	4.9	6.5
Serine	9.3	9.3	9.9	8.1
Glutamic Acid	9.2	11.6	12.4	8.6
Glycine	4.5	3.3	4.3	9.4
Alanine	15.2	18.6	15.9	14.7
Valine	15.0	11.6	12.6	8.8
Isoleucine	10.3	7.3	10.1	9.1
Leucine	9.3	11.2	10.0	0.9
Tyrosine	2.4	3.3	2.8	2.3
Phenylalanine	0.4	0.4	0.6	2.5
Lysine	4.2	5.4	4.7	4.7
Histidine	Trace	Trace	0	2.5
Arginine	4.0	4.8	4.2	5.5
Trytophan	n.d.	1.5	0.7	n.d.
Methionine	0	0	0	0
Cysteine	0	0	0	0
Proline	5.6	Trace	1.4	4.2

[a] Current data. These values are corrected for the decomposition of Asp, Thr, and Ser and the release of Val and Ile.

[b] Data from Jones and Jost (39).

[c] Composition of gas vacuoles from Anabaena flos-aquae (47).

[d] Composition of gas vacuoles from Halobacterium halobium (48).

3.4 Estimation of the Number of Tryptic Peptides. Jones (45) fingerprinted the tryptic peptides of the GVP with ninhydrin. However, incomplete trypsinization resulted in a variation of the pattern (See 3.6). I approached the problem by dansylating amino terminal amino acids of the tryptic peptides. After hydrolysis of the peptides the dansylated amino acids were separated and visualized. Figure 11 shows a minimum of 9 separate spots, two of which do not correspond to dansylated amino acids. Thus, a minimum of 7 peptides is indicated.

3.5 Separation of the Tryptic Peptides. Figure 12 represents a typical profile of tryptic peptides separated on Sephadex G-25. The profile as determined by ninhydrin analysis was always very reproducible, whereas the separation of the S_1, S_2, S_3 peaks as monitored in the ultraviolet, was inconsistent. Thus, the ninhydrin fractions, T1 and T2 were used. Further separation of the fractions T1 and T2 (Figure 4) is shown in Figures 13 and 14.

The first peak, T1P1, can be separated on Sephadex G-25 in 25% formic acid into 3 major fractions (Figure 15). T1P1a is a mixture of an arginyl peptide, a lysyl peptide, and some T1P1b peptide, as determined by amino acid analysis (Table 5) and by carboxypeptidase C digestion (2.11). Peak T1P1b is a large fragment of about 45 residues (Table 5). This large fragment results from incomplete trypsinization. The third major peak in Figure 15 contains some overlap from T1P1b, NH_4^+, and NaCl.

Figure 11. Minimum Estimation of the Number of Tryptic Peptides. Tryptic peptides (2.4) were dansylated and hydrolysed according to Gray (65) and separated electrophoretically at pH 1.9, 5kV, 135 minutes. Standard dansyl (DNS) amino acids were applied as indicated.

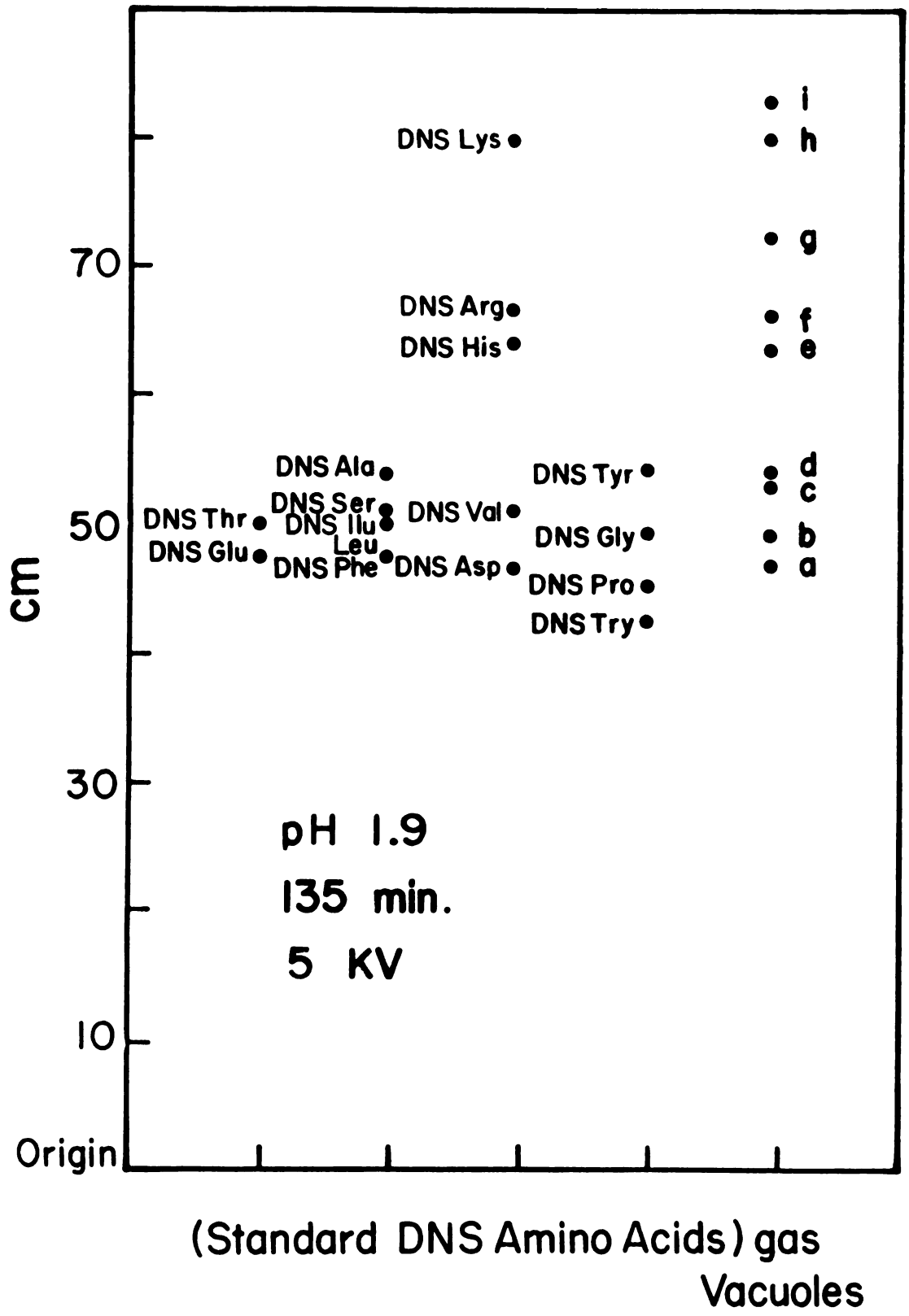

DNS Lys
DNS Arg
DNS His
DNS Ala
DNS Ser
DNS Thr
DNS Ilu
DNS Val
DNS Glu
Leu
DNS Phe
DNS Asp
DNS Tyr
DNS Gly
DNS Pro
DNS Try
i
h
g
f
e
d
c
b
a
70
50
30
10
Origin
cm
pH 1.9
135 min.
5 KV
(Standard DNS Amino Acids) gas
Vacuoles

Figure 11.

Figure 12. Separation of the Tryptic Peptides on Sephadex G-25 in 25% Formic Acid. T1, T2: fractions identified by ninhydrin and pooled for further analysis. S_1, S_2, S_3: fractions identified by U. V. absorption at 280 nm.

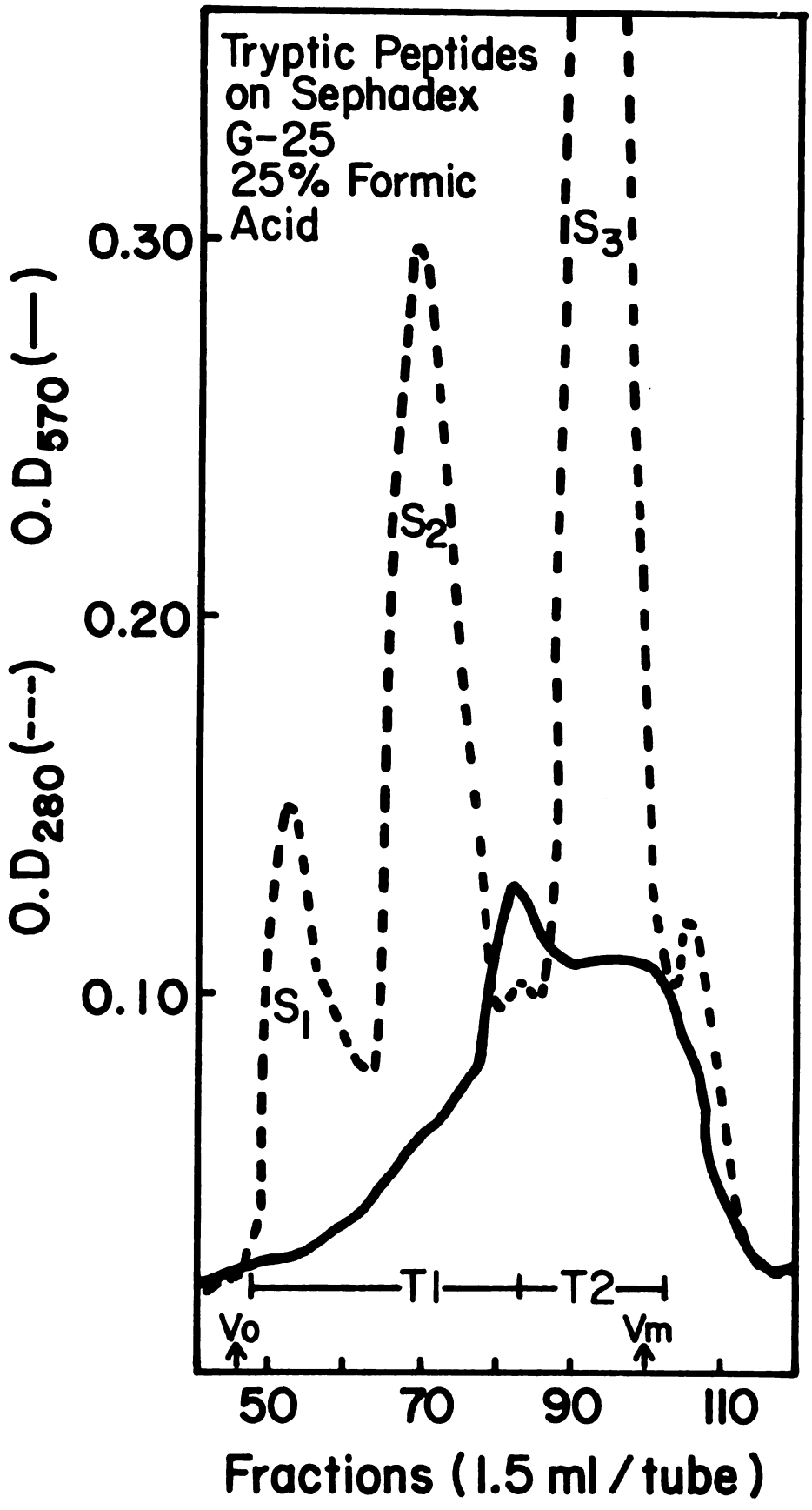

Tryptic Peptides
on Sephadex
G-25
25% Formic
Acid
0.30
0.20
0.10
O.D280 (---)
O.D570 (—)
S1
S2
S3
T1
T2
Vo
Vm
50
70
90
110
Fractions (1.5 ml / tube)

Figure 12.

Figure 13. Separation of Fractions T1 and T2AV[1] on SP Sephadex C-25-120. Elution was with a biphasic NaCl gradient (0.05 M - 0.7 M) in pH 2.3 formic acid and water (2.4).

[1]See Figure 14.

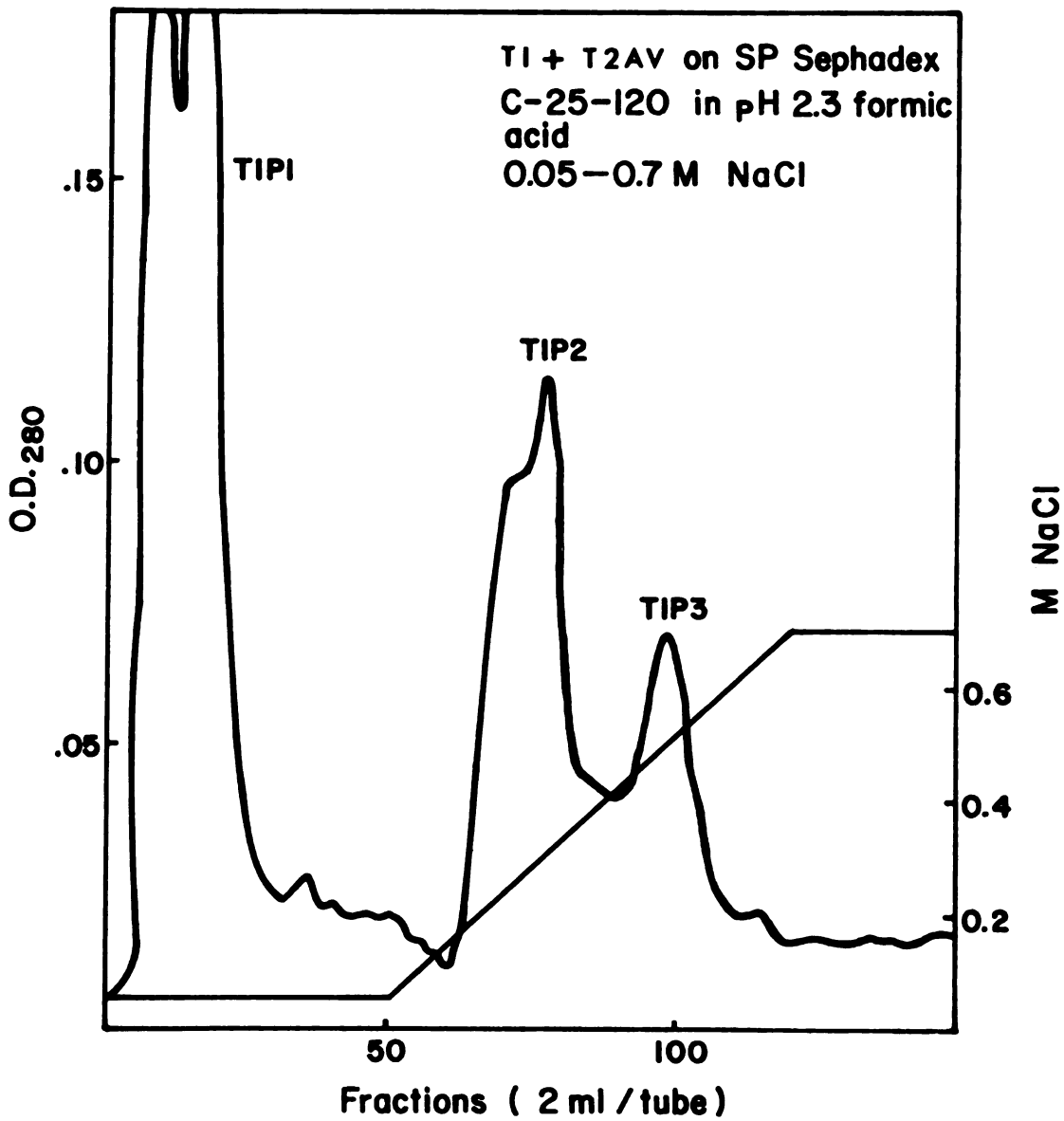

T1 + T2AV on SP Sephadex
C-25-120 in pH 2.3 formic acid
0.05–0.7 M NaCl
TIP1
TIP2
TIP3
O.D.280
.15
.10
.05
0.6
0.4
0.2
M NaCl
50
100
Fractions (2 ml / tube)

Figure 13.

Figure 14. Separation of Tryptic Fraction T2 on Aminex A-5. Elution was with a pyridine-acetic acid pH gradient (2.4).

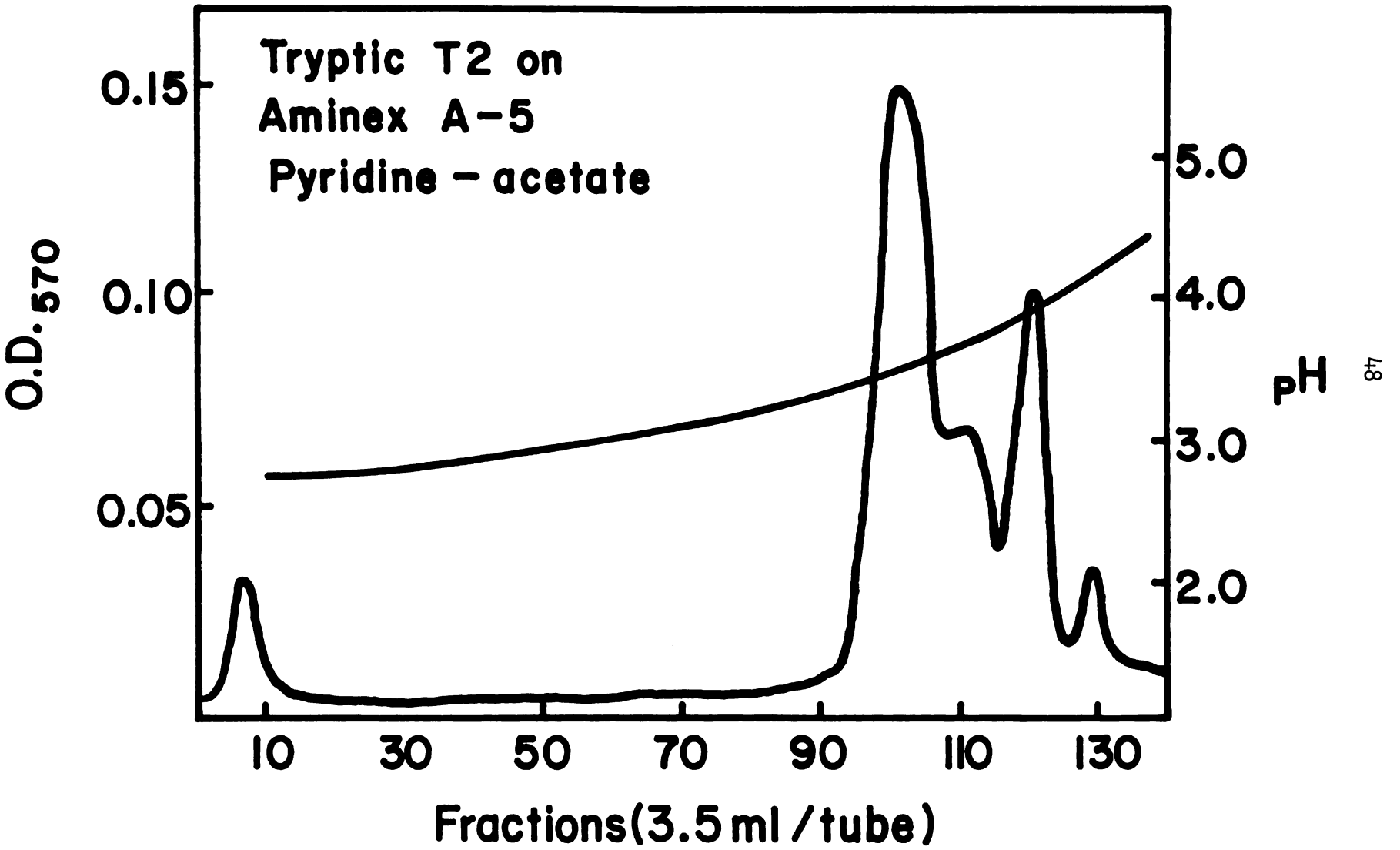

Tryptic T2 on
Aminex A-5
Pyridine – acetate
0.15
0.10
0.05
O.D.570
5.0
4.0
3.0
2.0
pH
10
30
50
70
90
110
130
Fractions(3.5 ml / tube)

Figure 14.

Figure 15. Elution Profile of Tryptic Fragment T1P1 on Sephadex G-25 in 25% Formic Acid. Fractions T1P1a, T1P1b are designated in the figure by the bars.

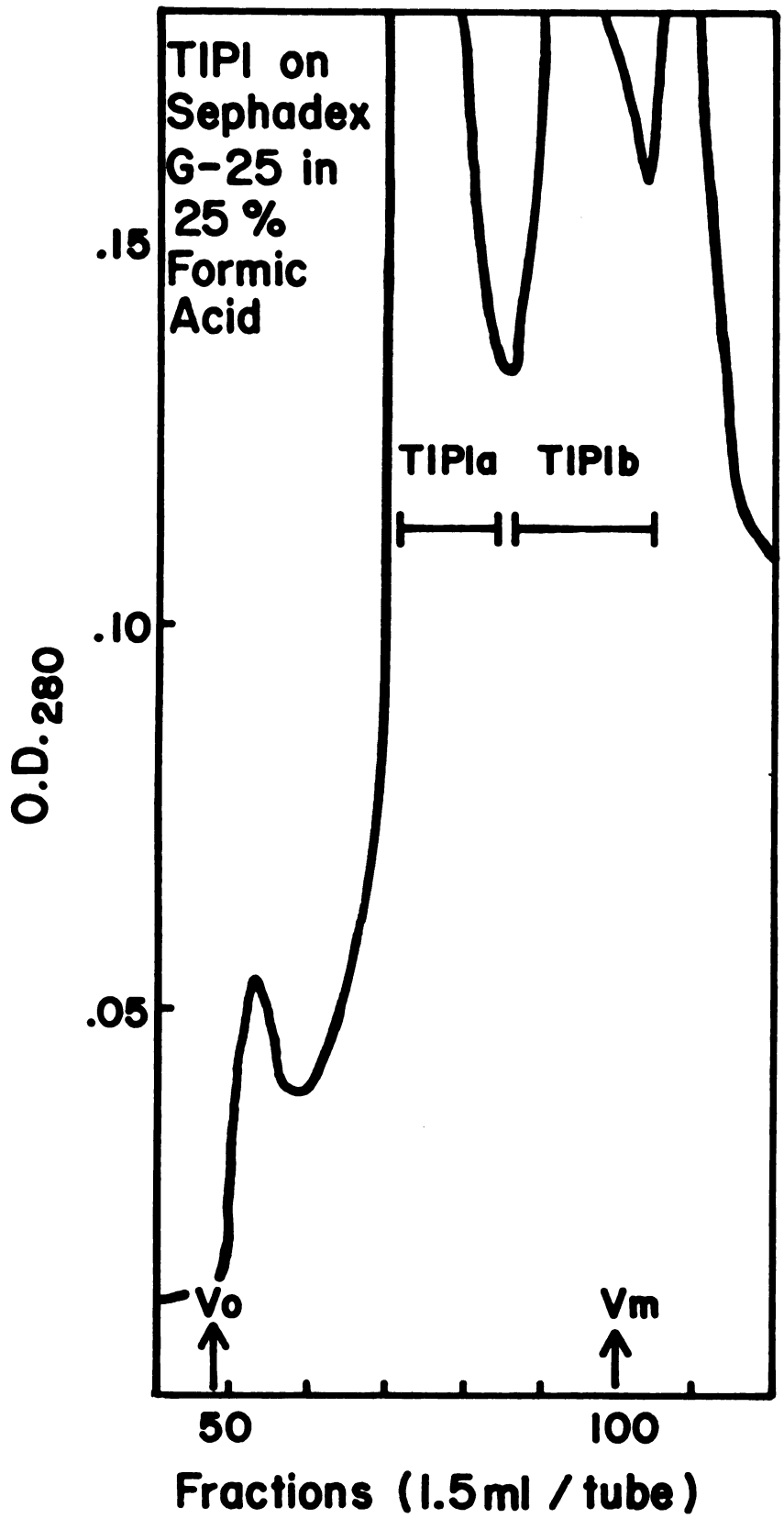

TIPI on
Sephadex
G-25 in
25 %
Formic
Acid
.15
.10
.05
O.D.280
TIPIa
TIPIb
Vo
Vm
50
100
Fractions (1.5ml / tube)

Figure 15

Table 5

Amino Acid Composition

of the Purified Tryptic Peptides[a]

Amino Acid	Peptides						
	T1P1a[b]	T1P1b	T1P2	T2A2	T2A3	T2A4	T1P2[c]
Asp	3.5	3.2		1.0	0.3		
Thr	2.5*	2.8	1.0				
Ser	5.9*	5.8	1.2	0.2			
Glu	5.5	5.8	2.1	0.2	0.7		
Pro	3.0	2.8	0.9				0.8
Gly	3.0	3.6	1.0	0.9*			
Ala	7.3¢	8.9¢	4.9¢	1.9	0.7¢		1.8¢*
Val	3.9	4.4	2.0	0.3	0.7		0.9
Ile	1.3*	1.1		1.1	0.3*		
Leu	3.0	3.3	1.1		0.3		
Tyr	1.2*	1.4*	0.8*				
Phe	0.6						
Lys	0.9¢				1.0¢	1.0¢	
Arg	1.0¢	1.0¢		1.0¢			
Total	42.6	44.1	15	6.6	4.0	1.0	3.5

[a] Values are expressed as residues per peptide.

[b] This peptide fraction is heterogeneous. The lysyl peptide was selectively modified and subsequently sequenced (3.6).

[c] This peptide is a fragment of T1P2.

* Amino terminal amino acids.

¢ Carboxyl terminal amino acids.

If peak T1P1 is separated over Sephadex G-25 in 0.1 M acetic acid (Figure 16), the peptide T1P1b elutes with the void as expected for a 45 residue fragment. The drastic difference in separation of this peptide on the same gel matrix with different solvents suggests that the highly protic 25% formic acid may be inducing ion-exchange qualities in Sephadex G-25.

The shouldered peak T1P2 contains the carboxyl terminus, peptide T1P2 (Table 5), of the GVP and a small amount of contaminating peptides, perhaps some overlap from fraction T2.

T1P3 is an agglomeration of several peptides which partially originate from T2.

The large T2A1 peak as seen in Figure 14, consists mainly of ammonia and at least one unidentified peptide as based on the presence of both arginine and lysine. The T2A2 shoulder contains primarily a unique arginine peptide (Table 5) with some overlap from T2A1 and T2A3. T2A3 is a mixture of two lysyl tetrapeptides, T2A3a and T2A3b, which consistently occur in unequal proportions (Table 5). Attempts to separate T2A3a from T2A3b by electrophoresis at pH 6.5 failed because the two peptides have the same size and charge.

Fraction T2A4 is free lysine (Table 5).

The T2AV fraction elutes with the void volume of the Aminex A-5 column and was combined with fraction T1. T2AV contains the overlap of T1 into T2.

The recovery during separation of the tryptic peptides (2.9) is given in Table 6. More arginyl peptides than lysyl peptides are

Figure 16. Elution Profile of Tryptic Fragment T1P1 on Sephadex G-25 in 0.1 M Acetic Acid.

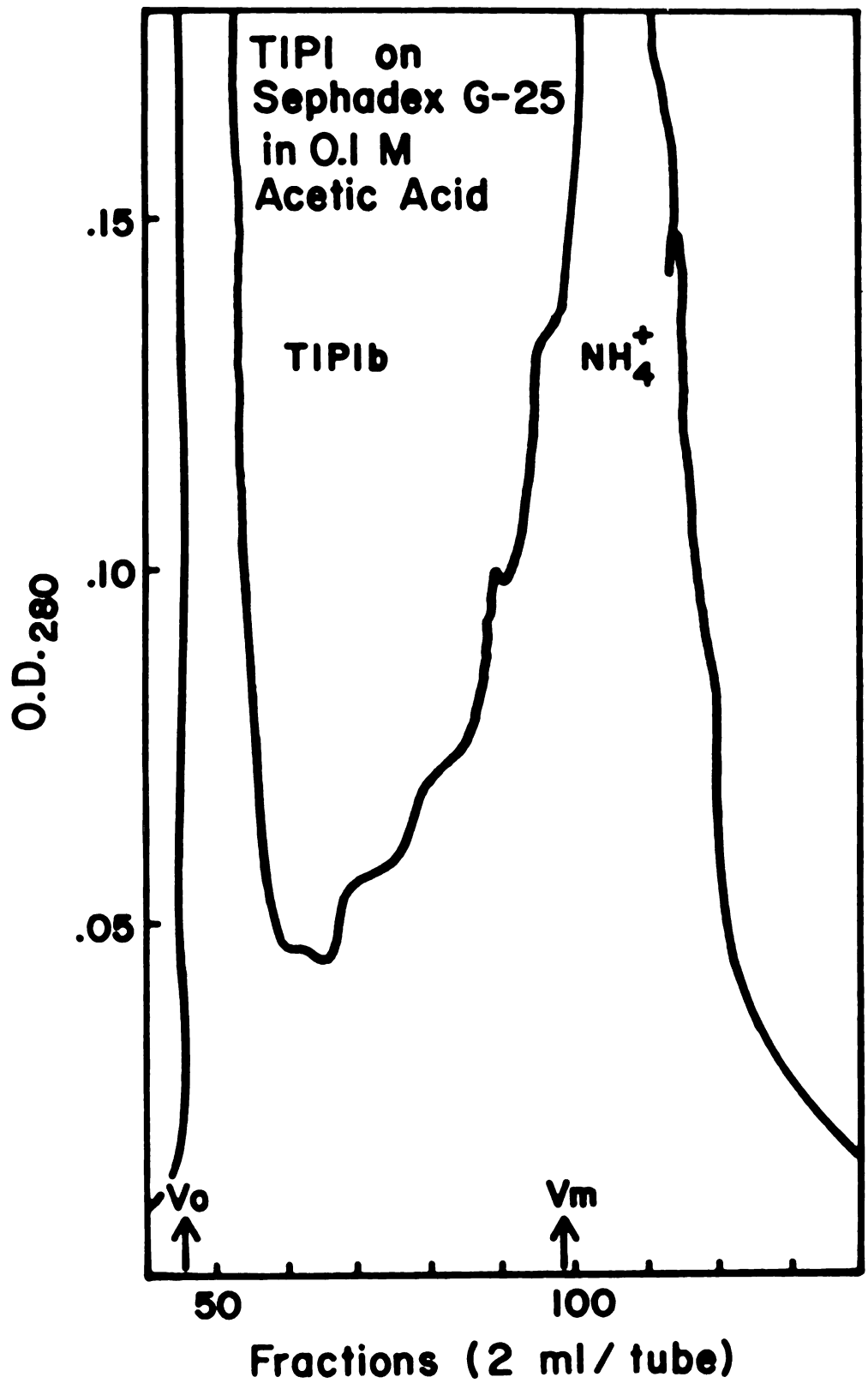

TIPI on
Sephadex G-25
in 0.1 M
Acetic Acid
TIPIb
NH_4^+
.15
.10
.05
O.D.280
Vo
Vm
50
100
Fractions (2 ml / tube)

Figure 16.

Table 6

Recovery of Tryptic Peptides

Fractionation Step[a]	μ Moles		% Yield	
	Lys	Arg	Lys	Arg
Fraction E	33.0	31.0	100	100
Acid Precip. GVP	33.0	31.0	100	100
T1	8.4	6.4		
T2	22.5	14.4		
Column residue	.7	1.1		
Total	31.6	21.9	96	70
T2AV	1.8	2.2		
T2A1, T2A2	8.5	1.7		
T2A3, T2A4				
Total	10.3	3.9	31	12
T1P1	1.5	1.8		
T1P2, T1P3	1.7	.5		
Total	3.2	2.3	9.8	7.5

[a] See Figure 2.

lost. Losses are due primarily to irreversible binding of various peptides to both Aminex and Sephadex matrices. Sephades G-25 e.g. retains 30% of the arginyl peptides of the fractions T1 and T2 (Figure 2), whereas virtually no lysyl peptides are lost. I also observed that a new batch of SP Sephadex will irreversibly bind more than 95% of the peptides applied to the column. With each succeeding application, this loss decreases. This also seems to be the case with Aminex A-5 and Aminex AG50W-x2.

Complete amino acid analyses of the tryptic peptides are presented in Table 5. The amino and carboxyl termini were identified by cyanoethylation (72), and when necessary, by carboxypeptidase C (53). Lysine and argine, when present, are considered in Table 5 as carboxyl termini on the basis of the specificity of trypsin. The peptide T1Pla is heterogeneous, but the peptide modification method of Braunitzer (59) allows selective sequencing of the lysine containing peptide.

3.6 Sequence Analysis of the Tryptic Peptides. This was done automatically by carboxypeptidase C and/or by automated Edman degradation. Carboxypeptidase C is an exopeptidase which removes all protein amino acids including lysine, arginine, and proline from the carboxyl terminus of a polypeptide. Peptides must be rather pure to obtain a reliable sequence. The carboxyl termini of tryptic peptides T2A2, T2A3, T2A4 are either arginine or lysine (Table 5).

The carboxyl terminal peptide T1P2 was treated with carboxypeptidase C. The data in Figure 17 show 2 alanines as the first amino acids liberated.[1]

After fragmentation of peptide T1P2 by partial acid hydrolysis (2.8), only one homogeneous peptide fragment, T1P2a, was isolated. The amino and carboxyl termini of the peptide are alanine (Table 5). Treatment with carboxypeptidase C suggested the sequence: Ala-Pro-Val-Ala-COOH but the yields were low and the amount of peptide limiting. Fragment T1P2a is considered to be adjacent to the carboxyl end of the peptide T1P2 based on the observation that partial acid hydrolysis would result in cleavage at seryl, or aspartyl residues and that Ala-Ala-COOH is the carboxyl terminal sequence of peptide T1P2. Thus, a sequence of Ala-(Pro, Val)-Ala-COOH could appear nowhere else in the peptide T1P2 (Table 7). Washout of the peptide T1P2 from the reaction cup prevented completion of the sequence. Chemical modification of the peptide increased the yields but did not allow a complete sequence determination either.

---

[1] The asterisked amino acids, Ser*, Glu* (Figure 17) and Ser*, Leu* (Figure 18), probably originate from the digested peptide which had 1 or 2 alanines removed by the enzyme and thus, are not part of the carboxyl sequence of either the peptides T1P2 or T1P1b (Table 7). The Technicon Amino Acid Analyzer also separates and displays small peptide fragments which may be eluted with the same relative retention time as amino acids. Consequently, the use of carboxypeptidase C is limited as a sequence tool since such spurious peaks appeared in all assays.

Figure 17. Release of Amino Acids from Peptide TlP2 by Carboxypeptidase C. Deduced Sequence: Ala-Ala. See Footnote 1 (3.6) for explanation of Ser* and Glu*.

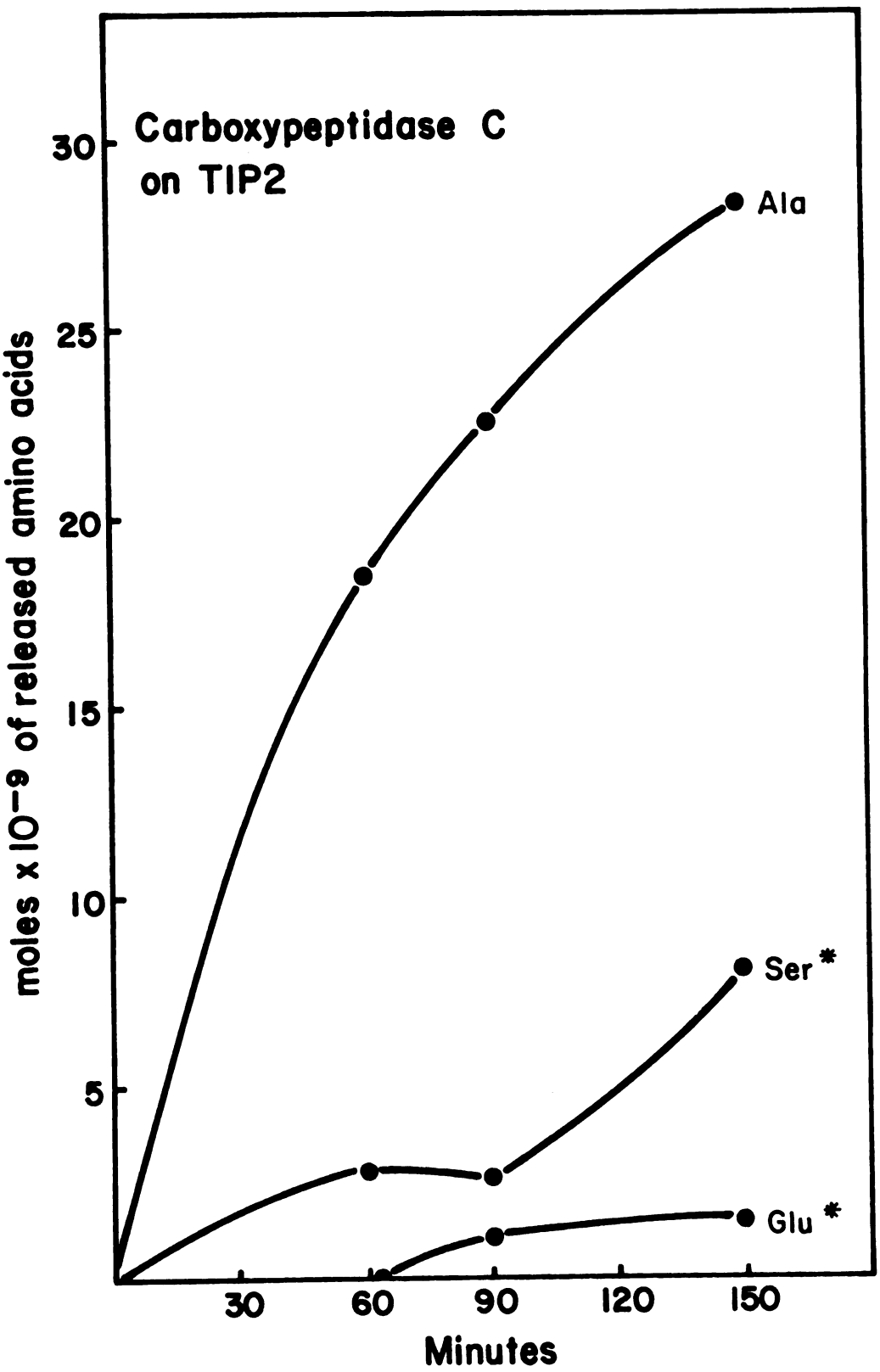

Carboxypeptidase C
on TIP2
30
25
20
15
10
5
moles x 10−9 of released amino acids
Ala
Ser *
Glu *
30
60
90
120
150
Minutes

Figure 17.

Table 7

Sequence Analyses of the Tryptic Peptides from GVP

Peptide \ Residue #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
T1P1a																													
Sequence	Ser	Ala	Glu	Ala	Val	Gly	Leu	Thr	Glu	Val	(Ile)	(Ala)	(X number of residues)					Lys											
Method	NH		H		H	H	H	H	H																				
% Yield[a]	60	40	25	21.8	12	11	18	8.3	10.6	10.2	11.5	9.9																	
T1P1b	Tyr	Ala	Glu	Ala	Val	Gly	Leu	Thr	Glu	Ser	Ala	Ala	Val	Pro	(15 residues)			Arg	Tyr	(T1P2)		Ala							
	H		H					H	H	H				H								C							
	9.2	34.4	22	16	15	11.6	18	11.3	9	9	10.8	12.5	11	9															
T1P2	Tyr	Ala	Glu	Ala	Val	(Gly,	Leu,	Thr,	Glu,	Ser,)	(Ala)	(Pro)	(Val)	Ala	Ala														
	N										PC	PC	PC	PC	C														
	39	71	52	73	65																								
T2A2	Gly	Ile	Val	Ile	(Asp)	(Ala)	Ala	Arg																					
	N				HC	C	C	C																					
	33	26	17	12	7																								
T2A3a	Ala	Val	Glu	Lys																									
	N		H																										
	-	46	20	-																									
T2A3b	Ile	Leu	Asp	Lys																									
	NH		H																										
	-	20	20	-																									
T2A4	Lys																												
	N																												
	-																												

The following code designates the method of Residue identification used in addition to GLC of the PTH/TMS: cyanoethylation, N; hydrolysis with HI, H; carboxypeptidase C, C; carboxypeptidase C digestion of peptides obtained by partial acid hydrolysis, P. [a]The percentage yield is based on the following procedure: PTH norleucine is added as an internal standard to the collection tubes and its percentage of recovery is used to normalize the amount of the individual amino acid released.

Figure 18. Release of Amino Acids from Peptide TlPlb by Carboxypeptidase C. Deduced Sequence: Ala-Ala. See Footnote 1 (3.6) for explanation of Leu* and Ser*.

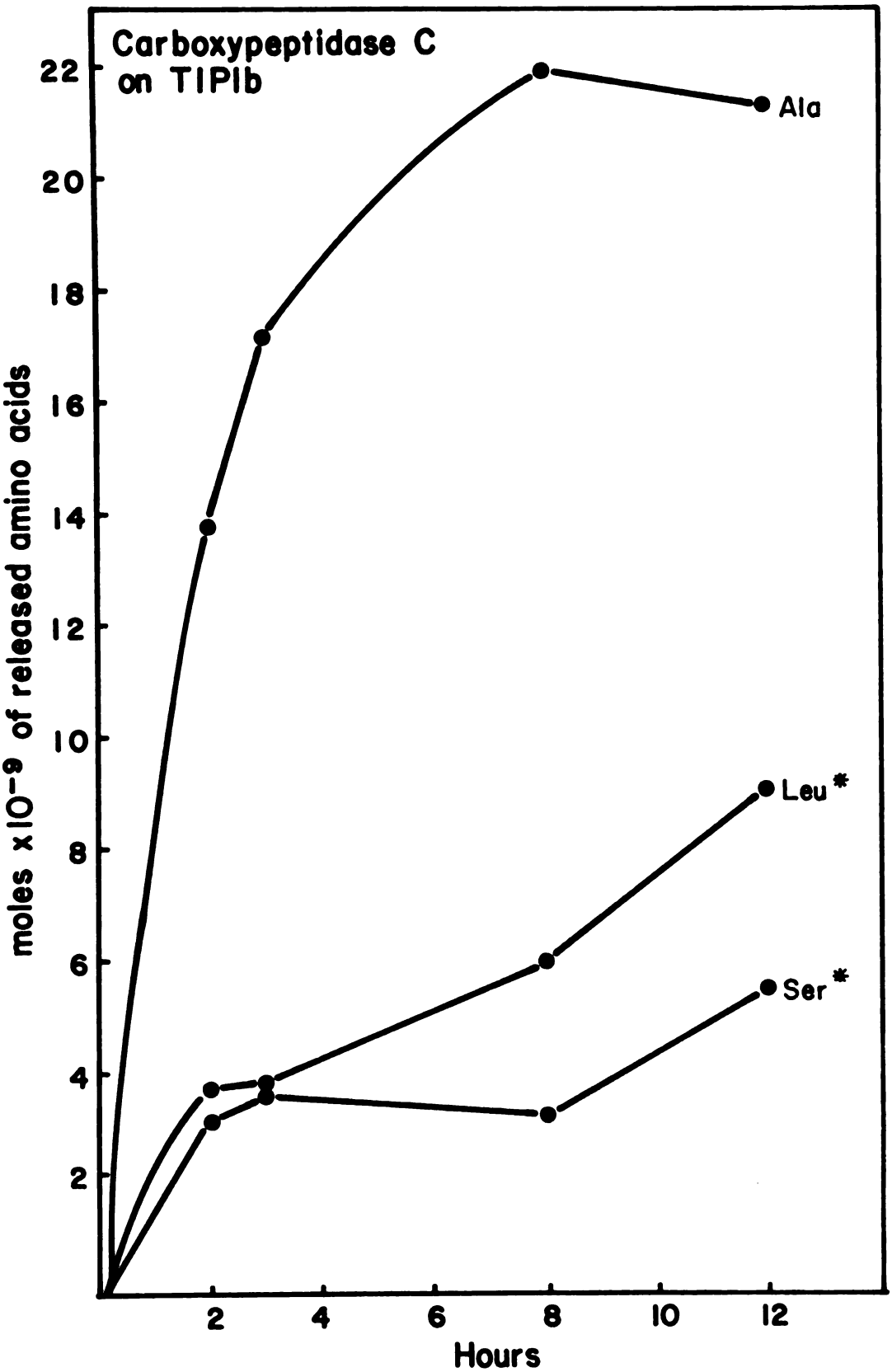

Carboxypeptidase C
on TIPIb
moles x10⁻⁹ of released amino acids
22
20
18
16
14
12
10
8
6
4
2
Ala
Leu*
Ser*
2
4
6
8
10
12
Hours

Figure 18.

Peptide T1P1b was only partially sequenced by automated Edman degradation (Table 7) despite modification of the carboxyl terminus (2.13). Although the peptide contains arginine, carboxypeptidase C releases 2 alanines (Figure 18) from peptide T1P1b. Thus the amino terminal portion of peptide T1P1b must be linked via arginine to the T1P2 peptide. The peptide T1P2b comprises the carboxyl terminus of GVP.

Peptide T2A2 was only partially sequenced by the automatic procedure due to peptide washout. Good information regarding the carboxyl terminus of the peptide T2A2 was obtained (Figure 19). The sequence is: (Asp)-(Ala)-Ala-Arg-COOH. There is no serine in peptide T2A2 and glycine is the amino terminus as determined by cyanoethylation (2.10, Table 5). Thus, Ser* and Gly* are not a portion of the carboxyl terminal sequence of T2A2. The apparent sequence of peptide T2A2 is given in Table 7.

Peptides T2A3a and T2A3b were sequenced by an unconventional method based on the presence of only 2 peptides occurring in very unequal amounts. A mixture of 2 peptides from tropoelastin was analyzed by a similar approach by Foster et al. (60). The yield ratio must be consistently unequal for the 2 peptides. The concentration of amino termini of peptides T2A3a and T2A3b differ proportionally i.e. the amount of isoleucine was less than that of alanine (Table 5). Sequencing of peptides T2A3a and T2A3b modified with SPITC proceeded as usual except that each cycle produced 2 residues in amounts as expected for the relative concentration of the

Figure 19. Release of Amino Acids from Peptide T2A2 by Carboxypeptidase C. Deduced Sequence: (Asp)-(Ala)-Ala-Arg. See text (3.6) for explanation of Ser* and Gly*.

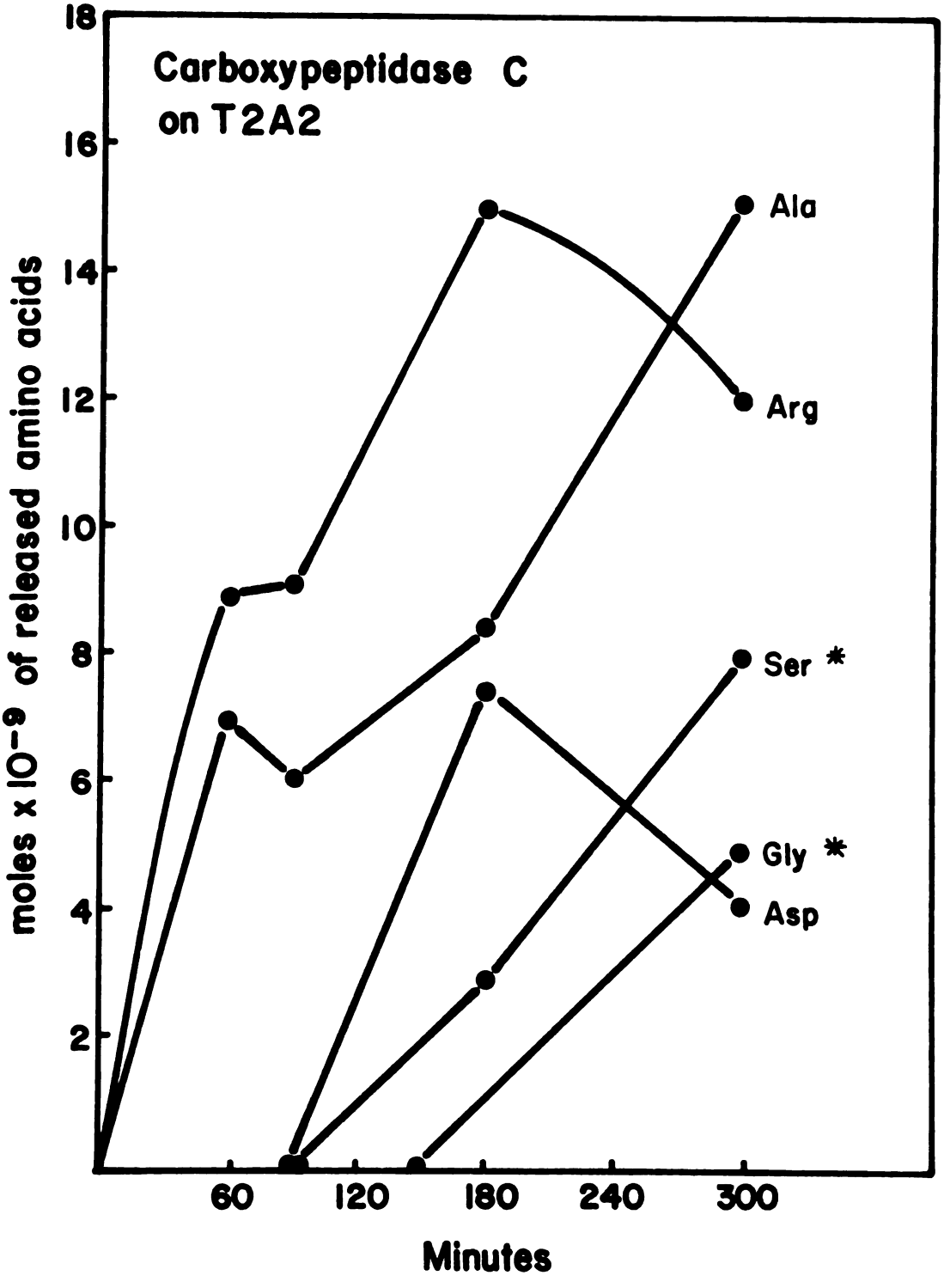

Carboxypeptidase C
on T2A2
moles x 10⁻⁹ of released amino acids
18
16
14
12
10
8
6
4
2
60
120
180
240
300
Minutes
Ala
Arg
Ser *
Gly *
Asp

Figure 19.

peptides. As seen in Table 7, the third residues were in equal yield so assignment of these amino acids to their respective peptides was based on other data. It was already known from amino acid analysis that the concentration of each of the amino acids (alanine, valine, glutamic acid, lysine) in peptide T2A3a was greater than the concentration of the amino acids (isoleucine, leucine, aspartic acid, lysine) in peptide T2A3b as shown in Table 5 so the conclusion that glutamic acid belonged to T2A3a and aspartic acid to T2A3b was made. Lysine was placed at the carboxyl terminus of each peptide on the basis of the specificity of trypsin.

A partial sequence of fraction T1P1a was also obtained by an unconventional method. The amino acid composition (Table 5) shows that peptide T1P1a contains a mixture of several peptides having different carboxyl and amino termini. SPITC reacts with only the free amino groups in peptides (59). As all unmodified GVP peptides readily wash out of the sequencer reaction cup within 2-3 cycles, the peptides terminating in lysine should remain. Based on the appearance of 9 consecutive unique residues (Table 7), the possibility of several peptides terminating with lysine was also discounted.

All other T1 and T2 peptides were not obtained in either sufficient yield or purity to warrant sequence analysis.

3.7 Maleylated Tryptic Peptides. Overlapping peptide sequences are needed, in order to align peptide fragments. Maleylation, partial acid hydrolysis, and N-Bromosuccinimide hydrolysis of GVP were tried. Blocking the ε-amino group of lysine often enhances solubility of proteins and prevents cleavage of lysine by trypsin. Therefore, I tried to obtain lysine-blocked tryptic peptides. The advantage of maleylation (67) is its reversibility below pH 6.5. The trypsinized, maleylated peptides were chromatographed and demaleylated over Sephadex G-25 with either 25% formic acid or 0.1 M acetic acid pH 2.87 (25). The elution profile of fraction T1, T2 is the same as in Figure 12. These fractions were then pooled and chromatographed (Figure 3). Elution profiles were identical to those for tryptic peptides (Figures 13 and 14). The amino acid composition of the peptides of T2A3 does not differ from that of unmaleylated protein. Apparently no maleylation of the lysines of the unboiled GVP occurred. On the other hand, heat denatured, maleylated GVP was found to have all of the ε-amino groups blocked (2.5). By cyanoethylating the protein before and after maleylation, the percentage of lysines blocked was calculated. This maleylation procedure was found to be 100% effective. After the tryptic treatment (Figure 3), the digest was chromatographed over Sephadex (2.5). The material voided all three gels. Since the GVP is very hydrophobic and poorly soluble, it is possible that either untrypsinized aggregates of the GVP form or that a few peptides have a tendency to associate and precipitate as the eluant at the void is rather turbid.

3.8 Dilute Acid Hydrolysis of GVP. Hydrolysis of protein with dilute acid (2.8) can produce peptides by cleavage and release of aspartyl residues (71). However, no free aspartate was found in the hydrolysate mixture with GVP.

3.9 N-Bromosuccinimide (NBS) Peptides. Tyrosyl and tryptophanyl residues in protein can be attacked and degraded by NBS (68). To achieve maximum cleavage of the GVP (2.6), 8 moles of NBS per mole of (tryptophan + tyrosine) were used. The extent of reaction of GVP with NBS was monitored by the appearance of peptide NPT (2.6) as measured by amino acid analysis. NBS was applied 3 times.

The NBS hydrolysate of GVP was chromatographed on Sephadex G-25 in 0.1 M acetic acid, pH 2.87 (Figure 20) and the three major fractions N1, N2, and N3 were separately fractionated on Sephadex G-25 in 25% formic acid. Figure 21 is a composite of these elution patterns. Fraction N1 separates into fractions N1a and N3. Fraction N2 separates into fractions N2a and N3. N3 elutes with the salt. The amino acid analyses of these different fractions are given in Table 8. Peptide N1A has alanine as its amino and carboxyl termini. The elution profile for N1a and N2a is given in Figure 22. As shown in the figure, more than 70% of the pooled fractions contain the single peptide N1A.

Recoveries (2.9) of NBS peptides during separation and purification are given in Table 9.

Cyanoethylation of the amino terminus of peptide NPT (Table 8) reveals that there are about 40 amino acid residues in the peptide.

Figure 20. Separation of the N-Bromosuccinimide (NBS) Peptides on Sephadex G-25 in 0.1 M Acetic Acid.

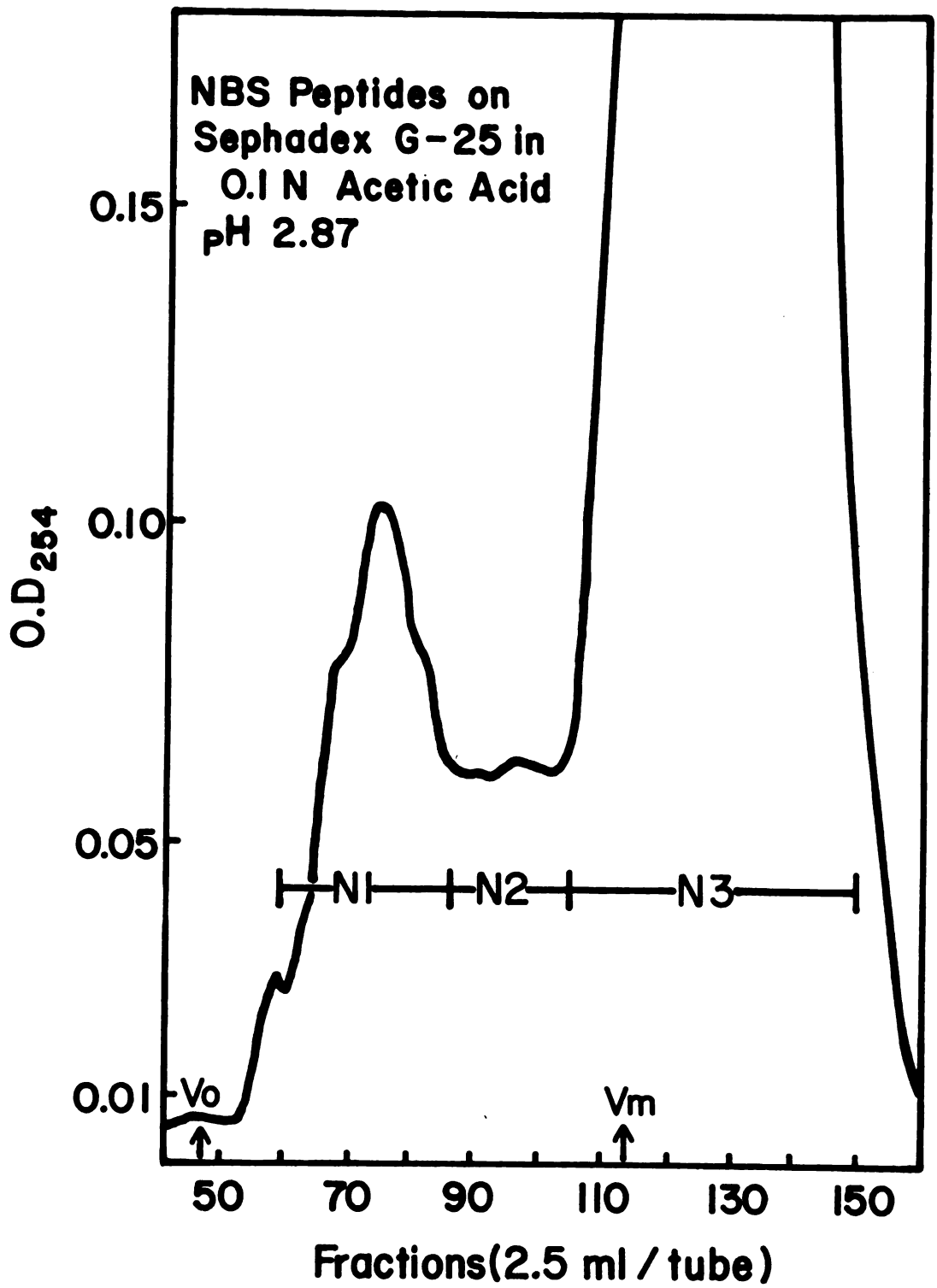

NBS Peptides on
Sephadex G-25 in
0.1 N Acetic Acid
pH 2.87
0.15
0.10
0.05
0.01
O.D254
N1
N2
N3
Vo
Vm
50
70
90
110
130
150
Fractions(2.5 ml / tube)

Figure 20.

Figure 21. Composite Elution Profiles of the NBS Fractions N1, N2, and N3 on Sephadex G-25 in 25% Formic Acid. (- - - -) is N1; (....) is N2; (——) is N3. Fractions are designated as N1a, N2a, and N3.

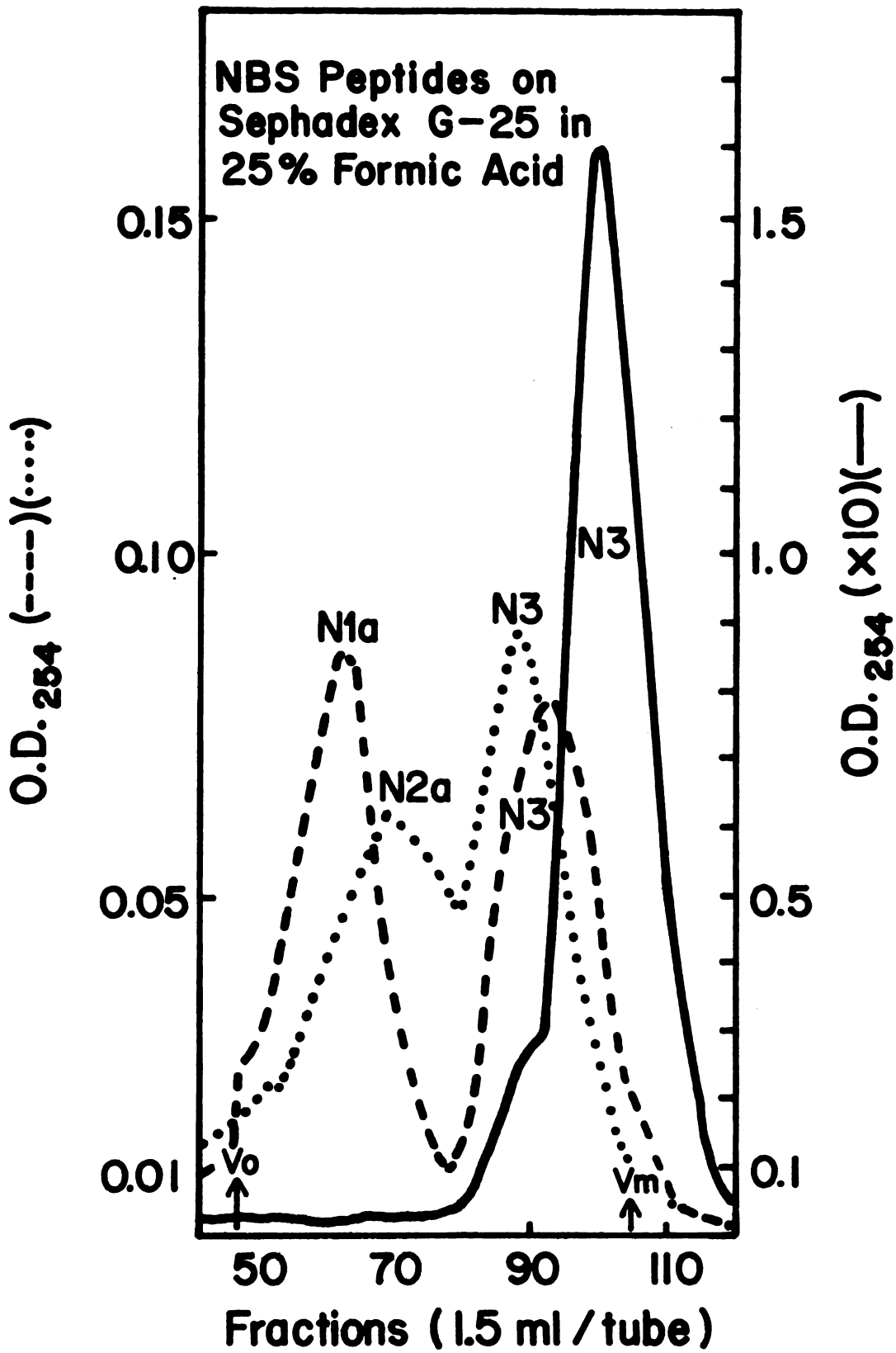

NBS Peptides on
Sephadex G-25 in
25% Formic Acid
0.15
0.10
0.05
0.01
1.5
1.0
0.5
0.1
O.D.254 (----)(....)
O.D.254 (×10)(—)
N1a
N2a
N3
N3
N3
Vo
Vm
50
70
90
110
Fractions (1.5 ml / tube)

Figure 21.

Table 8

Amino Acid Composition of the NBS Peptides[a]

Amino Acid	Peptides				
	NIA	N2A[c]	N3[c]	NPT	
Asp	(0.5)	1.6	1.7	1.5[b]	3.0[b]
Thr	1.2	1.3	1.0	1.0	2.0
Ser	1.1	1.9	1.3	2.4	4.8
Glu	2.3	4.3	2.8	2.3	4.6
Pro	1.1	0.4	0.9	0.7	1.4
Gly	1.1	1.9	1.5	1.1	2.2
Ala	5.5¢	4.0	3.2	3.5*	7.0*
Val	2.2	1.9	1.6	3.0	6.0
Ile	(0.2)	0.5	0.4	1.7	3.4
Leu	1.2	2.1	2.2	2.3	4.6
Tyr					
Phe			0.5		
Lys	(0.2)	0.6	1.8	0.9	1.8
Arg	(0.1)	1.0	1.0	1.0	2.0
Total	16.7	21.5	19.9	22.4	42.8

[a] Values are expressed as residues per peptide.

[b] The first set of values is based on the presence of a unit arginine. The second set of values is based on an amino terminal analysis of the peptide by cyanoethylation.

[c] These peptides may be heterogeneous.

( ) Residues in parentheses are fractional residues thought to be impurities.

* Amino terminus.

¢ Carboxyl terminus.

Figure 22. Separation of Fractions (N1a + N2a) on Sephadex G-25 in 0.1 M Acetic Acid. The fraction N1A is designated by the bar.

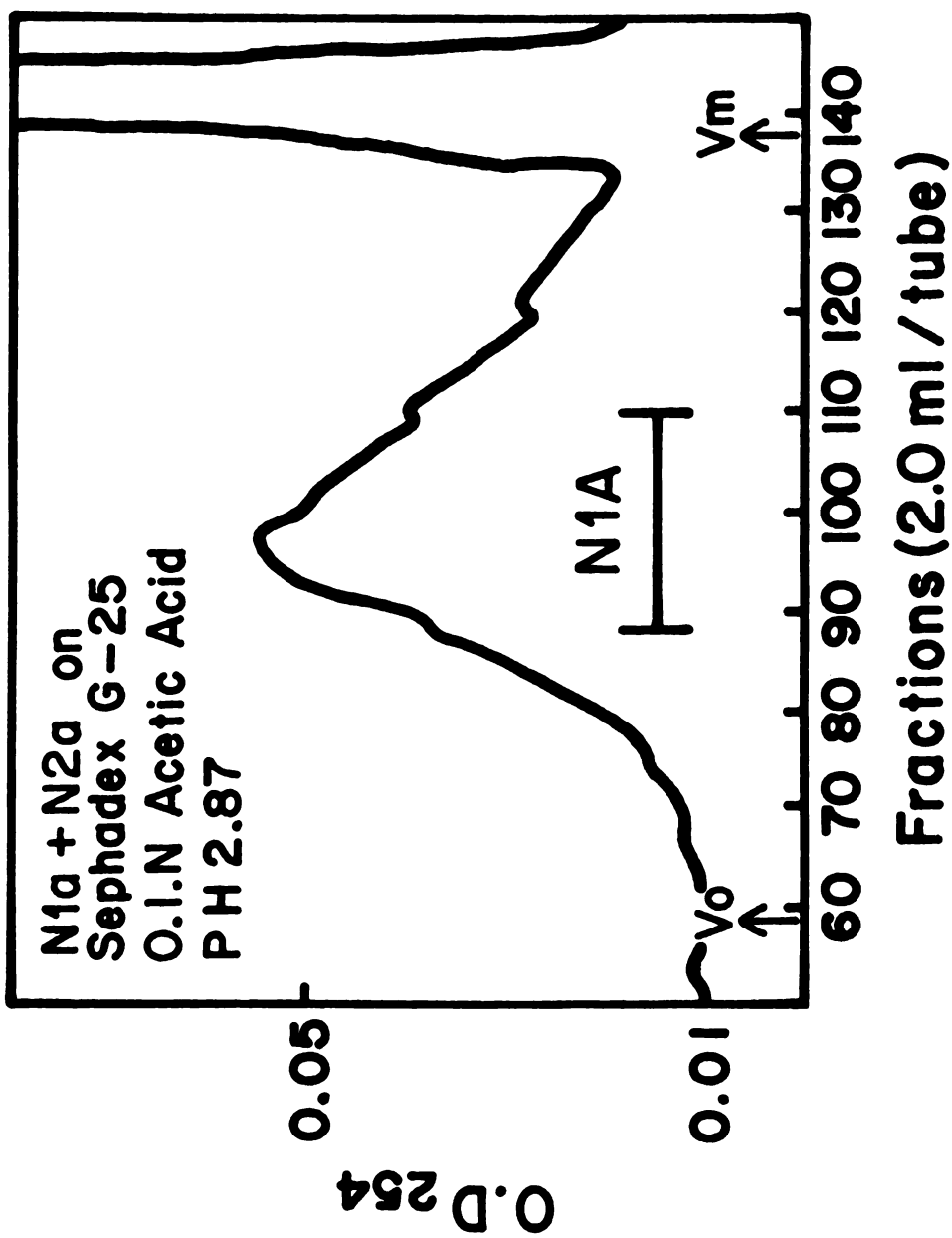

N1a + N2a on
Sephadex G-25
O.1.N Acetic Acid
P H 2.87
N1A
Vo
Vm
0.05
0.01
O.D 254
60 70 80 90 100 110 120 130 140
Fractions (2.0 ml / tube)

Figure 22.

Table 9

Recovery of NBS Peptides

Fractionation Step[a]	μ Moles		%Yield	
	Lys	Arg	Lys	Arg
Fraction E	28.0	27.0	100	100
N1, N2, N3	6.4	5.0	23	19
N1a	1.1	.5		
N2a	.3	.5		
all N3's	1.5	.3		
Total	2.9	1.3	10.6	5
NPT	6.0	7.0	21	25

[a] See Figure 5.

[b] This yield is low as the large amount of the carboxyl terminal peptide has no arginine or lysine.

This estimate of peptide size is not completely reliable, as the large number of aliphatic amino acid residues are linked to one another so as to be poorly acid hydrolysed. A carboxyl terminal amino acid in peptide NPT was not found even after extensive treatment with carboxypeptidase C. This may be due to the extreme insolubility of peptide NPT.

3.10 Sequence Analysis of the NBS Peptides. Peptide NIA has as a carboxyl terminal sequence, Ala-Ala-COOH, when treated with carboxypeptidase C (Figure 23). Small peptide contaminants in the peptide NIA preparation do not permit further sequence analysis.[2]

The best, although not complete sequence analysis of NIA was accomplished through use of the automated sequencer on EDC-ANS modified peptide. By comparing the sequence of peptides TIP2 and fragment TIP2a, a tentative sequence of NIA is proposed in Table 10.

In order to improve the sequence analysis of peptide NPT, modification with SPITC and/or EDC-ANS was used (2.13). Although

---

[2] The asterisked amino acids, Leu* and Thr* (Figure 23) probably originate from either the undigested portion of peptide NIA or the contaminating peptides found in the NIA preparation based on the reasons given in Footnote 1 (3.6). Automated sequence analysis has placed leucine and threonine as the sixth and seventh residues in the peptide, so they could not be part of the carboxyl sequence of peptide NIA. The other amino acids shown, proline, valine, serine, and glycine (Figure 23) are also subject to the same criticism and therefore can not be positioned into the sequence. Small amounts of lysine and arginine were also detected. Their release is not shown as they do not occur in peptide NIA.

Figure 23. Release of Amino Acids from Peptide NIA by Carboxypeptidase C. Deduced Sequence: (Val)-(Pro)-Ala-Ala. See Footnote 2 (3.10) for explanation of Leu*, Thr*, Gly* and Pro, Val, Ser.

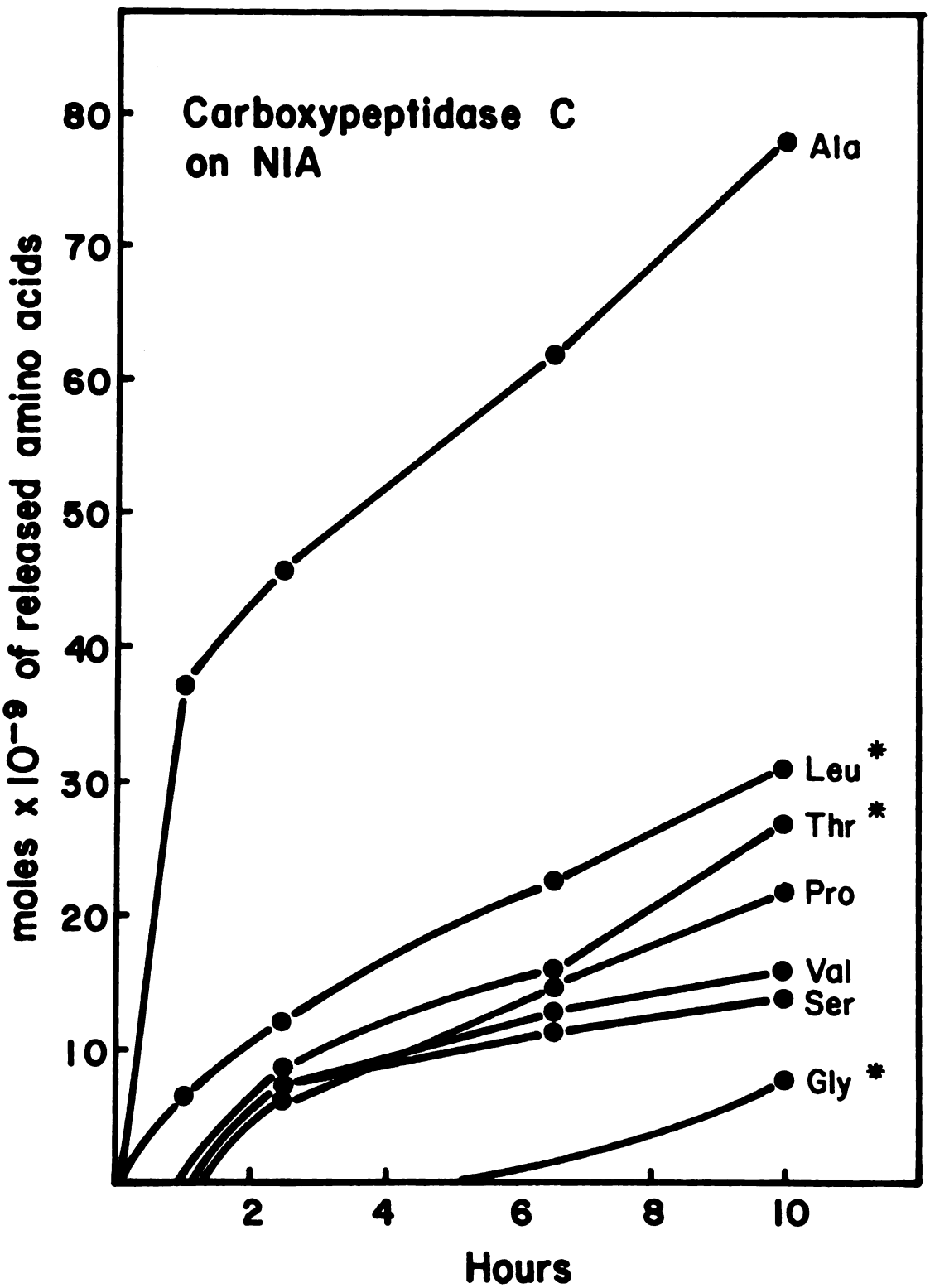

Carboxypeptidase C
on NIA
moles x 10⁻⁹ of released amino acids
80
70
60
50
40
30
20
10
2
4
6
8
10
Hours
Ala
Leu *
Thr *
Pro
Val
Ser
Gly *

Figure 23.

Table 10

Sequence Analyses of the NBS Peptides from GVP

Peptide \ Residue #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
NPT																													
Sequence	Ala	Val	(Val)	(Val)	Leu	Val	(Val)	Ile	(Ile)	Leu	(Leu)	Ala	(Leu)	(Val)	(Ile)	(X number of residues)													
Method	N		H	H			H						H																
% Yield[a]	32	23	20	10	3.9	12	6.8	9.2	10.4	11	11	5.8	5.0	2.8	8.8														
DNb	Ile	Val	Glu	(b)	Thr	(X number of residues)																							
	NH		H	H	H																								
	11	34	13.5	-	6.5																								
NIA	Ala	Glu	Ala	Val	Gly	Leu	Thr	Glu	(Ser)	(Ala)	(Pro)	(Val)	Ala	Ala															
		H			H	H	H	H	H			H	C	C															
	90	21	10.7	8	6	6.7	8	4	3.5			3																	

For the method of residue identification see Table 7.

[b] Not determined, due to high background from peptides in HI hydrolysis.

addition of 90 μmoles ANS and 5 or 6 treatments with 60 μmoles of EDC was made over a period of 72 hours, the amino acid sequence of peptide NPT is still questionable (Table 10): overlap from numerous repeating, aliphatic residues makes the data difficult to interpret. There is in peptide NPT, for example, considerable overlap of residue number 5, leucine, into residue number 6, valine. Thus, such overlap makes the repeating sequences of residues number 3 and 4, and number 9 through 11 questionable. A calculation of percentage yields of these residues does not alleviate these ambiguities as the peptide is progressively lost with the organic washes. Hydriodic acid hydrolysis of the PTH-TMS derivatives of peptide NPT (2.3) gave only very low yields as the background washout was higher. This was also the case when the ethyl acetate fraction was used (2.14). However, the data were very reproducible suggesting that an unusual sequence exists. Improved modification of the carboxyl end of the peptide was not feasible due to insolubility of the peptide. Thus, attempts were made to fragment peptide NPT by either trypsin (2.4), thermolysin (2.7), or partial acid hydrolysis (2.8).

Peptide NPT contains both arginine and lysine. Figure 24 gives the elution profile observed after tryptic digestion. The recovery of the material separated is given in Table 11. Sequence analysis of peptide fractions TNPTa, TNPTb, TNPTc was prevented by mechanical difficulties and by the presence of several peptides.

Figure 24. Elution Profile of the Tryptic Peptides TNPTa-e of NPT. Digestion with trypsin was 72 hours. Elution was with 25% formic acid on Sephadex G-25.

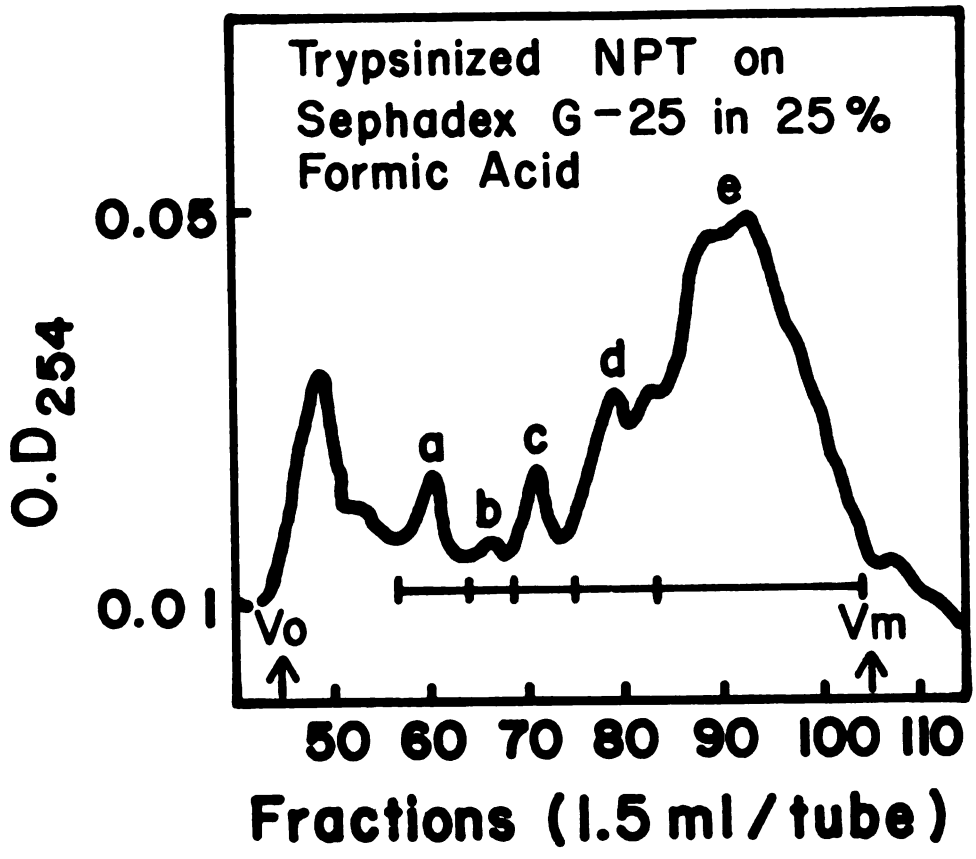

Trypsinized NPT on
Sephadex G-25 in 25%
Formic Acid
0.05
0.01
O.D254
a
b
c
d
e
Vo
Vm
50 60 70 80 90 100 110
Fractions (1.5 ml/tube)

Figure 24.

Table 11

Recovery of the Tryptic Fragments of NPT[a]

Fraction	10^{-9}M Asp	%Yield
TNPTa	640	16
TNPTb	450	11
TNPTc	530	13
TNPTd	420	10
TNPTe	300	7.5
NPT	4200	100

[a] Percent recovery is based on Asp present initially and in the subsequent fractions a - e.

Five hour incubations of peptide NPT with thermolysin produced 3 major fractions as observed after electrophoresis. However, their yield was too low for sequence analysis. Hydrolysis for 10 hours produced a large number of short di- and tripeptides as indicated by the Sephadex G-25 elution profile (Figure 25). The overlap of the fractions and the small size of peptides precluded sequencing. The large peaks seen early in the elution profile (Figure 25) may be ammonia.

Peptide NPT contains 3 aspartyl residues, thus, 3 or 4 peptides are expected after dilute acid hydrolysis. After 24 hours, more than 70% of the aspartyl residues were released. Chromatography of the hydrolysate is illustrated in Figure 26. Major fractions resolved are: DNa, b, c, and d. The recovery of these peptide fragments is given in Table 12. Amino acid analyses (2.12) of these fractions and NPT and fraction DNPT are given in Table 13. Cyanoethylation analyses (2.10) of fraction DNb indicated the presence of more than one peptide as shown in Table 13. Free aspartate was the primary constituent of fraction DNc, representing over 37% of the initial aspartyl residues. DNd included overlap of free aspartate from fraction DNc and ammonia.

Since fragment DNb contained lysine, selective modification using SPITC was used (2.13). Results are in Table 10. More extensive hydrolysis with either dilute or concentrated acid may permit additional sequencing. As the peptide NPT contains so many

Figure 25. Separation of NPT Peptides Released by Thermolysin. Elution was with 25% formic acid on Sephadex G-25.

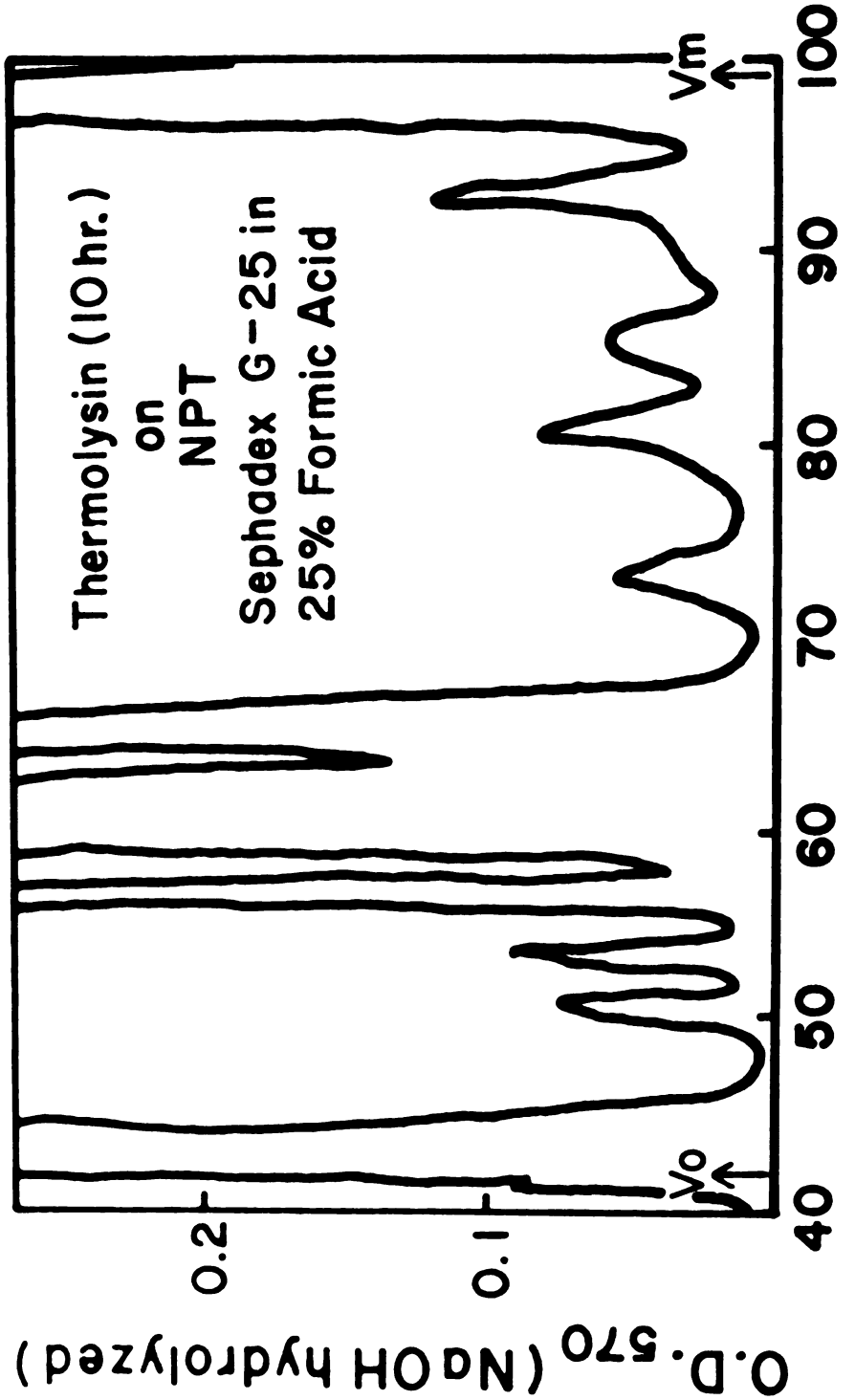

Thermolysin (10 hr.)
on
NPT
Sephadex G-25 in
25% Formic Acid
Vo
Vm
0.2
0.1
40
50
60
70
80
90
100
Fractions (1.5 ml / tube)
O.D.570 (NaOH hydrolyzed)

Figure 25.

Figure 26. Separation of the Peptides DNa-d Released by Dilute Acid Hydrolysis of Peptide NPT. Hydrolysis conditions: 0.03 N HCl, 105° C., 48 hours.

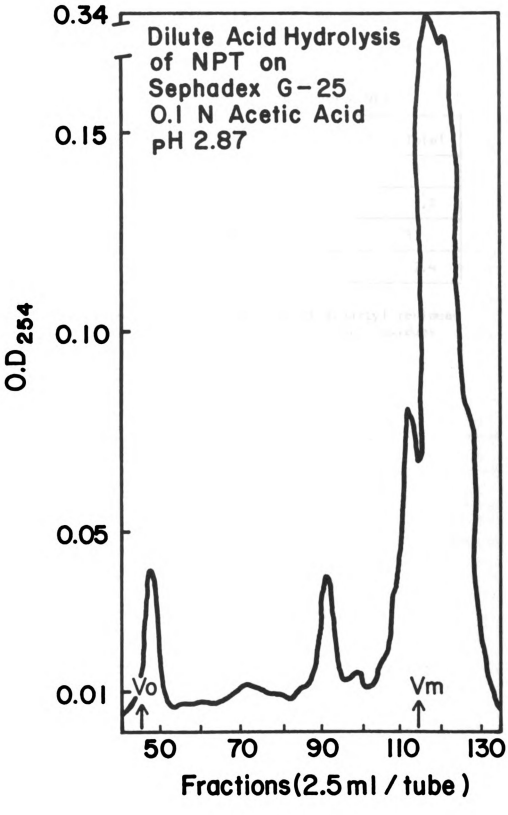

Dilute Acid Hydrolysis
of NPT on
Sephadex G-25
0.1 N Acetic Acid
pH 2.87
0.34
0.15
0.10
0.05
0.01
O.D254
Vo
Vm
50
70
90
110
130
Fractions(2.5 ml / tube)

Figure 26.

Table 12

Recovery of Fragments from Peptide NPT

Fraction	10^{-9}M Asp	% Total
DNa	0	0
DNb	325	7.2
DNc	1670	37
DNd	335	7.4

Percent yields are based on the amount of aspartyl residues in each fraction relative to the total aspartyl residues found in NPT prior to hydrolysis.

Table 13

Amino Acid Analyses of the Peptides

Observed after Dilute Acid Hydrolysis of NPT[a]

Amino Acid	Peptides				
	NPT[c]	DNPT[d]	DNa	DNb	DNc
Asp	3.0	0.8	0	0.9	3.9[b]
Thr	2.0	1.6	0	1.2	0.4
Ser	3.0	3.2	0	2.8*	0.8
Glu	4.6	4.4	0	2.0	0.6
Pro	1.4		0	0.9	1.1
Gly	2.2	2.2	0	2.0	1.2
Ala	7.0	7.0	0	3.0	2.1
Val	6.0	5.6	0	2.0	0.5
Ile	3.4	3.6	0	1.6*	1.0
Leu	4.6	5.2	0	1.2	1.0
Lys	1.8	0.6	0	2.0	0.5
Arg	2.0	2.0	0		0.8
Total	41.0	36.2	0	102.9	13.9

[a] Values are expressed as residues per peptide.

[b] This value represents primarily free Asp.

[c] NPT residues/peptide before dilute acid hydrolysis.

[d] DNPT residues/peptide after dilute acid hydrolysis. This material was recovered from the top of the column.

* Amino terminus.

hydrophobic residues at the amino terminal portion of the peptide, only the carboxyl end of the peptide containing the seryl and aspartyl residues should be degraded.

3.11 Improved Peptide Modification Method for Sequencing. In order to keep the very hydrophobic GVP peptides in the sequencer reaction cup, the EDC-ANS method (2.13) was used to modify the carboxyl terminus. Figure 29-f gives the yields in logarithmic plot obtained from sequencing the synthetic tripeptide DL-Leu-Gly-Gly. Sequencing of the unmodified peptide results in total loss of the synthetic peptide to washout (Figure 29-a). The yield obtained after using the method of Foster (60) is shown in Figure 29-e. Successive treatments with EDC produces the best retention of the peptide in the reaction cup (Figure 29-g). This procedure is especially effective for insoluble peptides like NPT if 2 or more doses of EDC are given over a long time (18-48 hours).

Figure 27. Peptide Modification for Improved Sequencing. The different treatments are: unmodified peptide, a; 1.5 μM ANS + 0.6 μM EDC, pH 4.0, b; 0.6 μM ANS + 0.6 μM EDC + 0.1 M NaCl, pH 4.0, c; 0.6 μM ANS + 0.6 μM EDC, pH 3.0, d; 0.6 μM ANS + 0.6 μM EDC, pH 4.0, e; 0.3 μM ANS + 1.5 μM EDC, pH 4.0, f; 0.3 μM ANS + 1.5 μM EDC) x 2, pH 4.0, g; 0.3 μM peptide DL-Leu-Gly-Gly- used in all samples. Total reaction time for all samples = 4 hours.

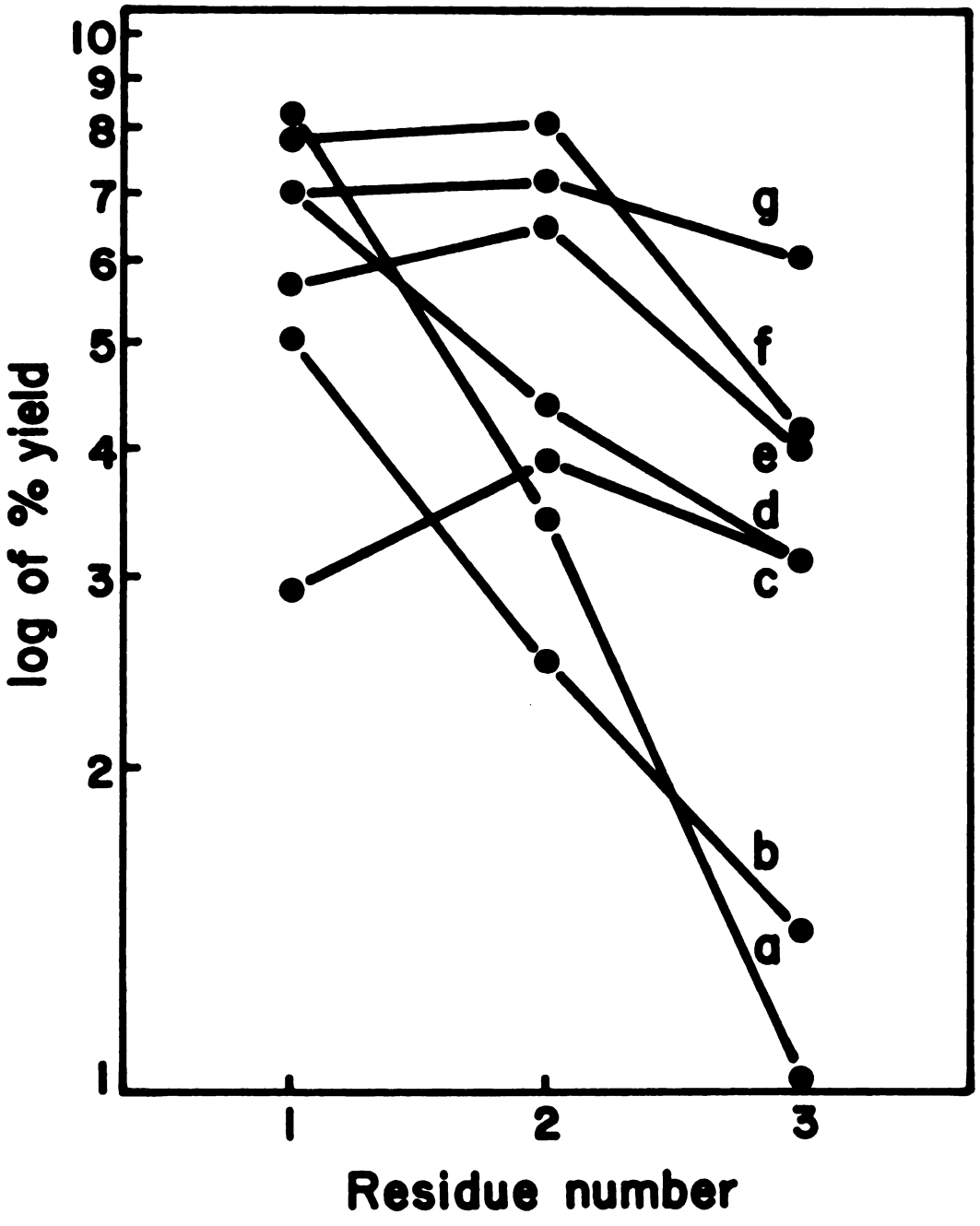


Figure 27.

DISCUSSION

4.1 General Problems Encountered in Purification, Separation, and Sequencing of GVP Peptides. Numerous difficulties were encountered in working with the GVP. Foremost among these was the problem of solubility, a feature common to both integral membrane proteins and proteins like collagen and elastin. Although attempts were made to modify the protein by maleylation, succinylation, etc., these failed. The lysine residues react with maleic anhydride only after boiling the GVP, (3.7). Lysine residues, besides being buried in the interior of the protein, as indicated by these results, may be cross-linked to other portions of the protein. However, trypsinization of lysine residues occurs and this makes cross-linkage rather unlikely. In addition, Jones et al. (79) demonstrated that in intact gas vacuoles, lysines are accessible to spin labels. Thus, more extensive maleylation of unboiled GVP may be necessary in order to achieve modification of the lysine residues.

Problems were also encountered in the area of the production of peptides. As seen above (3.7, 3.8), several methods of generating peptide fragments did not work with the GVP and those which did, as tryptic digestion and NBS cleavage, produced only a fair yield of peptides. Without doubt, these problems are directly related to the insolubility of GVP. Since others (47, 80) have failed to generate peptides from GVP even these small yields are encouraging.

Once peptides were made, additional difficulties presented themselves in the isolation and purification procedures. The losses during purification of peptides (Tables 7 and 10) are large. Efforts to reduce these losses have failed and it became apparent that peptides were binding irreversibly to not only ion change matrices (3.5), but also to Sephadex gels which is rather unusual (Table 7). This appears to be due to the fact that as more material is run over the same column bed, fewer sites are available for irreversibly binding to the peptides (3.5). Despite the use of many modification methods, segments of the protein remain yet to be sequenced. However, up to now this is not unusual as only one other hydrophobic protein has ever been sequenced, glycophorin (30). Moreover, the properties of glycophorin do not compare to those of GVP. Even the hydrophobic domain of glycophorin does not present the extreme problems of residue overlap as does the amino terminal region of the peptide NPT (3.10).

As indicated in Table 8, only two NBS peptides were obtained in yields sufficient to warrant sequence analysis. It is possible that not all of the tyrosine and tryptophan residues in GVP reacted with NBS; incomplete cleavage fragments may yet be found. Cases of unreactive tyrosine and tryptophan have been reported (81, 82, 83).

4.2 An Improved Method for Reduction of Extractive Losses of Peptides. Since complete sequences of long peptides are seldom obtained in single runs using the sequencer, modification of the carboxyl terminal residues has become the standard method of

ending premature losses of residual peptide with the wash fluids. Current methods were discussed earlier (1.7), however, I have improved the carbodiimide method of Foster et al. (60) (Figure 27-e) by using excessive amounts of EDC in small but frequent doses over extended periods of time. This has improved yields by 5-20% for soluble tryptic and NBS peptides, especially for those with solubility problems as the peptide NPT. This extensive dose procedure seems necessary since EDC is slowly hydrolyzed by water to the corresponding urea. The coupling reaction should be driven to completion upon addition of excess ANS (84). However, upon testing, this failed, as seen in Figure 27-b. The low sequencer efficiency caused by increased ANS is probably due to undesirable side reactions with the sequencer reagents. Whereas, increasing only the EDC concentration by means of this improved method, enhances the yields.

4.3 Molecular Weight Determination of GVP. Evidence is discussed here which will clarify some of the discrepancy in the reported molecular weight of GVP (41, 43). Based on the amino acid composition of gas vacuoles from either M. aeruginosa or A. flos-aquae and assuming the protein contains one tryptophan per mole of protein, a molecular weight of about 7,300 is reached (mean residue molecular weight for GVP = 110). However, since a minimum of 89 nonoverlapping residues have been separated and partially sequenced, a molecular weight of at least 9,800 should be considered. Based

on this, there are probably 2 tryptophans instead of 1 which would make the molecular weight about 14,600. In addition, a sequence analysis of the protein showed that greater than 80% of the amino terminus, alanine, was retrieved, based on a molecular weight of 14,300. If the molecular weight of the protein were 7,300, a yield closer to 200% (160%) should have been obtained. If the molecular weight were 21,500, a maximum yield of 50% should have been obtained. These results indicate a molecular weight of about 14,300 provided that the amino terminus of GVP was fully accessible to the sequencer reagents and relatively complete cleavage was achieved. Considering all of this data, a minimum molecular weight of 14,600 is proposed for the GVP.

4.4 Sequence Analysis of Intact GVP and Alignment of Peptides. A customary procedure in beginning a sequence analysis of a protein is to attempt automated Edman degradation of the intact protein. Although it is unusual to sequence an entire protein in one attempt on an automated sequencer, large portions of the amino terminal sequence of many proteins have been reported. Such information greatly enhances the speed of sequence analysis. In addition to the primary purpose of accelerating the determination of the amino acid sequence, the sequencer can be used to test the purity of a particular protein preparation. Although this approach was not successful but for the first three residues of the GVP, these few residues eliminated the possibility that the GVP was not a single protein, unless a contaminating species had a blocked (e.g. acetylated) amino terminal. Despite obtaining this short sequence of

the intact GVP, the protein remained in the reaction cup i.e. it did not wash out (3.3).

Although the amino terminal sequence of the GVP agreed with that of the T2A3a peptide: Ala-Val-Glu-Lys, the possibility of formation of pyrrolidone carboxylic acid from glutamine was considered as a possible cause for the sudden drop in the percentage yield. However, attempts to shorten the anhydrous acid cleavage step which causes this cyclization of glutamine did not alter the results. Besides, if the identification of T2A3a is accepted as the amino terminal peptide of GVP, then formation of pyroglutamate could not possibly be the problem.

Since there are no unusual amino acids found in the GVP, it can only be suggested that there may be a tertiary protein interaction (e.g. via an interpeptide bond) which is not readily denatured either by treatment with 88% formic acid or heat. This seems possible since the protein is rather insoluble in most organic and aqueous solvents (39). Work currently in progress with Dr. J. Foster of Boston University School of Medicine should establish if any cross-links between polypeptide chains of GVP are found as in some structural proteins of eucaryotes, e.g. elastin (85).

The sequence analysis of the intact GVP enabled positive identification of the peptide T2A3a as the amino terminus of the protein. In addition, it definitely showed that the NPT peptide, although beginning with the same sequence (Ala-Val) was not just a lengthy amino terminal fragment since removal of the amino

terminal alanine and subsequent production of peptide NPT did not alter the amino acid sequence of NPT.

With the additional information from the NBS peptides that NAI was the carboxyl terminus of the protein, placement of NPT internal to the protein as seen in Figure 28 became obvious.

Based on the fortuitous occurrence of the T1P1b tryptic peptide, a result of incomplete tryptic cleavage, it was also possible to align the carboxyl end of the protein. Thus, as mentioned before (3.6), T1P2 is the carboxyl terminus of T1P1b. Alignment of the other tryptic fragments T2A3b, T2A3, T2A4, and T1P1a remains unsettled. Consequently, only a small portion of the peptides of the GVP can, as yet, be aligned.

Those tryptic peptides shown in parenthesis in Figure 30 may be incorporated, at least in part, in the NPT peptide. Complete analysis of peptide NPT may allow such conclusions.

4.5 Discussion of the Peptide NPT. Two aspects of the sequence analysis of the GVP are intriguing. These are the recovery and partial sequence of the very hydrophobic peptide NPT, and the occurrence of an octapeptide which is thrice repeated in the protein.

The NPT peptide, which according to its amino acid composition contains greater than 50% hydrophobic residues, has a unique amino terminal sequence: a stretch of 15 aliphatic amino acids. As shown in Table 10, some of this sequence is questionable, although the same sequence was obtained using different peptide modification methods and both GLC and hydriodic acid hydrolysis for residue

identification. In any sequence analysis dealing with consecutive aliphatic residues, as residues number 2 through 4 (Val-Val-Val) in NPT, as much as 50% overlap can and does exist (56, 70). Moreover, it is known that peptide bonds formed between two β-branched amino acids (e.g. Val-Ile, Ile-Ile) are extremely difficult to hydrolyse and probably form the most acid resistant type of peptide bond (86). For example, more than 50% of residue number 5, leucine, routinely appears with residue number 6 of peptide NPT. However, since valine appeared at position number 6, leucine was observed as an overlap and not as the seventh residue. Similar but as yet unresolved problems occur for residues number 3, 4, 7, 9, and 11. All attempts at resolving the identity of these residues using hydriodic acid have met with little or no success as the background appearance of NPT peptide washout is high enough to obscure the appearance of these low-yield, PTH residues. Thus, the identification of all of the NPT residues number 1 through 15 has relied solely on GLC analysis.

If the sequence of the NPT amino terminal is as shown (Table 10), it is one of the most hydrophobic stretches of a protein yet described. The only other comparable sequences seem to be the intramembranous region of glycophorin (30) and the glycine-and alanine-rich peptides from elastin (85). These, however, do not contain such an aliphatic stretch as they do hydrophobic and neutral amino acids. Four other proteins are known which have either very hydrophobic regions or amino acid composition and which might have such aliphatic domains.

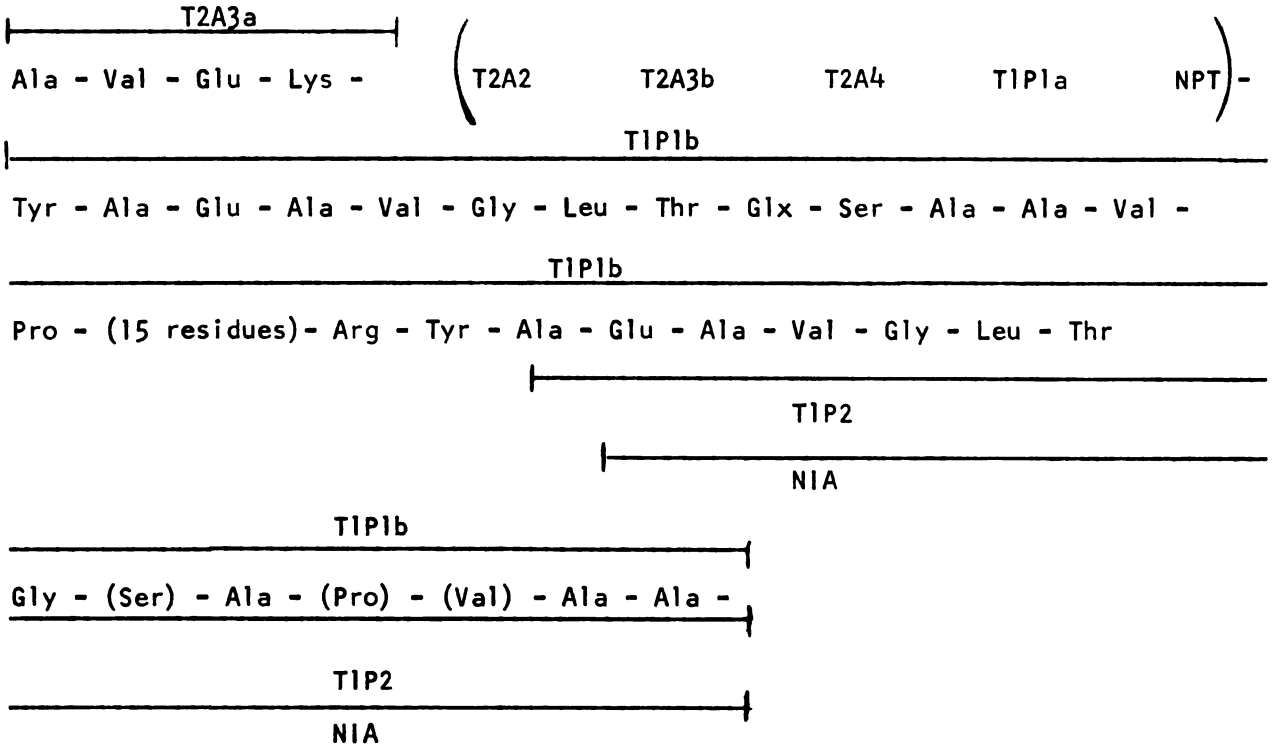


Figure 28. Alignment of the Tryptic and NBS Peptides in the GVP. Peptides within the brackets are not yet aligned.

They are the hydrophobic peptide from cytochrome b_5 (19), rhodopsin (87), the liver membrane protein (88), and C_{55} isoprenoid alcohol phosphokinase (89). Moreover, the last three proteins possess the highest known content of hydrophobic residues of any membrane-bound proteins (Table 14). However, no structural data are yet available for any of these polypeptides.

Since I have established that the NPT peptide resides in the intrapeptide region of the GVP, it suggests that the gas vacuole membrane protein is following the pattern observed in glycophorin and cytochrome b_5 which have their hydrophobic regions embedded in the lipid bilayer of the membrane. One may speculate that the gas vacuole membrane, having no lipid, compensates for this problem by creating its own hydrophobic environment which is interior to the vacuole surface, a feature suggested by Walsby (41). It is known that the hydrophobic portions of some proteins, notably cytochrome c, are maintained over long evolutionary periods, although the total amino acid composition has changed up to 50% throughout this time (90, 91). Perhaps, the gas vacuole membrane exists as a living fossil; a prototype for the more complex membranes. Since the high degree of hydrophobicity is apparently a common feature of all integral membrane proteins, this would provide a conservative mechanism for maintaining the integrity of membranes over long periods of time.

Segrest et al. (20, 30) have postulated that the hydrophobic domain of glycophorin, which has been shown to span the lipid

Table 14

Relative Content of Polar, Intermediate, and Apolar Amino Acids for Some Membrane-Bound Proteins

Membrane Protein	Apolar(%)	Polar(%)	Intermediate(%)	Ref.
Chromatium Sulfur Membrane (single protein)	41.7	34.2	24.0	94
Rhodapseudomonas spheroides-envelope	50.8	27.5	20.7	95
C_{55}-isoprenoid alcohol phosphokinase	63.0	18.8	18.2	89
Glycophorin (intramembranous region)	60.0	14.5	21.6	30
Cyt. b_5	43.3	32.5	24.2	19
Purple membrane *H. halobium*	55.7	23.1	22.0	48
Gas Vacuole Membrane				
M. aeruginosa	58.2	23.7	18.3	39
A. flos-aquae	54.1	26.9	19.1	47
H. halobium	52.0	28.5	19.4	80
Proteolipid protein, peripheral myelin	60.9	15.6	23.2	96
Proteolipid protein, heart	58.8	16.8	25.3	97
Proteolipid protein, central w. matter	56.8	17.1	26.1	98
Proteolipid protein, central myelin	55.4	17.1	26.0	99
Protein N-2, central myelin	56.1	18.1	28.2	100
Coat Protein				
Bacteriophage FD	63	12	24	50
Q Beta	40	32	27	50
TMV	42	28	29	50

bilayer of the erythrocyte membrane is helical. Other authors studying the CD spectra of isolated membranes, have also suggested that integral membrane proteins have a high helical content (15). X-ray (41), CD (44), and IR (46) analyses of the intact gas vacuole membrane protein indicate the presence of both helix and sheet structure. However, the location of these secondary structures is unknown. Based on the fact that the NPT peptide is very insoluble in aqueous solvents a sheet conformation was suspected. Moreover, the Chou and Fasman method of prediction of protein conformation (32), indicated that the NPT peptide sequence favors sheet formation. Where the relative strength of helix formers is Ha>ha>la>ia>ba>Ba and the relative strength of sheet formers is Hb>hb>Ib>ib>bb>Bb; $\langle Pa \rangle$ and $\langle Pb \rangle$ are the average amino acid residue strengths of formation of helix and sheet, respectively (32):

helix	Ha	ha	ha	ha	Ha	ha	ha	Ia	
sequence	Ala-	Val-	(Val)-	(Val)-	Leu-	Val-	(Val)-	Ile-	
sheet	Ib	Hb	Hb	Hb	hb	Hb	Hb	Hb	
	Ia	Ha	Ha	Ha	Ha	ha	Ia		1.20 = $\langle Pa \rangle$
continued	(Ile)-	Leu-	(Leu)-	Ala-	(Leu)-	(Val)-	(Ile)		
	Hb	hb	hb	Ib	hb	Hb	Hb		1.43 = $\langle Pb \rangle$

Although the hydrophobic domain of the GVP does not have a helical nature, this does not preclude the possibility of helix existing elsewhere in the protein.

4.6 A Repeating Octapeptide in the GVP. The second intriguing aspect of the primary structure of the GVP is the presence of a

repeating octapeptide as found in peptides T1P2 and N1A, T1P1a, and T1P1b (Tables 7 and 10):

Ala-Glu-Ala-Val-Gly-Leu-Thr-Glu

Since these peptides have been shown to have different elution patterns, different amino termini (serine or tyrosine), and different carboxyl termini (alanine, arginine, lysine), they are definitely separate peptides; that is, they are not merely different overlapping fragments of one sequence region.

Repetitive sequences are common to structural proteins such as collagen and elastin (50). Many investigators, among them Nolan and Margoliash (49), have suggested that the presence of repeating amino acid sequences in the primary structure of some proteins like immunoglobins might indicate that "partial gene duplications represent a common and important evolutionary mechanism for increasing the size of polypeptide chains". This may well be the case for immunoglobins, ferredoxins, and clupeines as they suggest. However, in structural proteins, sequence repetition appears to be more important as a structural building block rather than as a method for increasing chain length. For example, in elastin the short-range repetition of sequences $(Pro\text{-}Gly\text{-}Val\text{-}Gly\text{-}Val\text{-}Ala\text{-})_n$ and $(Pro\text{-}Gly\text{-}Val\text{-}Gly\text{-}Val)_n$ suggest a regular structure. Thus, the "oiled-coil" model was proposed for elastin in which these repeating units are theorized to account for the extensibility of the protein (85).

Perhaps, this periodicity of sequence in the GVP accounts for its unique rigidity. Based on the Chou and Fasman method (32) of

prediction of secondary structure, these repetitious regions should have a strong tendency to form helix as shown below:

helix Ha Ha Ha ha Ba Ha ia Ha = 1.22 ⟨Pa⟩

sequence Ala-Glu-Ala-Val-Gly-Leu-Thr-Glu

sheet Ib Bb Ib Hb ib hb hb Bb = 0.92 ⟨Pb⟩

Underlining denotes the predicted helical regions for each of these peptides:

TIP1b Tyr-Ala-Glu-Ala-Val-Gly-Leu-Thr-Glx-Ser-Ala-Ala-Val-Pro-...

TIP1a Ser-Ala-Glu-Ala-Val-Gly-Leu-Thr-Glx-Val-Ile-Ala-...

TIP2 Tyr-Ala-Glu-Ala-Val-Gly-Leu-Thr-Glx-Ser-Ala-Pro-Val-Ala-Ala.

If indeed these are periodic, helical regions in the protein, then what could lend greater strength to a structure than a series of coils or "cylinders", the geometry of which represents one of the strongest structures available to architectural design. Only the complete sequence of the GVP can show if more of these repetitive segments exist in the protein.

As stated above, predictions, based solely on primary structure, can be made regarding the helical, sheet, turn and coil regions of a particular polypeptide with greater than 85% certainty (32). Application of this method of prediction of secondary structure based only on peptide fragments tends to underpredict the amount of helix and sheet in a protein since it neglects to consider the tertiary folding of the protein which can provide stability to secondary structures (32). Nevertheless, the method can provide valuable information for a proposed structural model of the molecule. All of the other GVP

peptides which have been sequenced are too short to provide enough information for a prediction of secondary conformation with the exception of T2A2:

helix Ba Ia ha Ia ia Ha Ha ia = 1.04 ⟨Pa⟩

sequence Gly-Ile-Val-Ile-(Asp)-(Ala)-Ala-Arg

sheet ib Hb Hb Hb ib Ib Ib ib = 1.15 ⟨Pb⟩

This peptide can be evaluated as probably having either a coiled or helical configuration but not sheet structure as it contains charged residues in a sheet nucleation area (32) which makes sheet conformation less favorable.

4.7 Implications of the Amino Acid Composition of GVP and Its Relationship to Other Integral Membrane Proteins. Much information can be gleaned from a comparison of the amino acid composition of proteins which are considered to be related in function or structure. Not only can relative polarities (22, 92) be determined from the amino acid composition, but also calculations of relative amounts of secondary structure (93).

The results of the improved amino acid analysis (Table 4) show that valine, leucine, and isoleucine constitute 35% of the residues of the GVP of Microcystis aeruginosa, a value close to those found for GVP from Anabaena flos-aquae (32%) (47) and Halobacterium halobium (33%) (80). It is of interest that the gas vacuole membrane protein from bacteria and blue-green algae all contain over 50%

hydrophobic residues as based on the criteria of Tanford (92) and Capaldi and Vanderkooi (22).[3]

The relative percentages of hydrophobic, polar and intermediate groups in various integral membrane proteins fall into a distinct category (Table 14). If compared to corresponding values for non-membrane-bound or peripheral proteins, the polarity, as defined in Footnote 3, of most integral membrane proteins is between 29 and 40% for polar amino acids. The corresponding values for most peripheral proteins fall between 41 and 53% (22). From such data the relationship emerges that the less a particular protein is associated with a membrane structure, the more polar it becomes with a concomitant decrease in hydrophobicity. Such a relationship has been alluded to by others (17, 22) but, perhaps, the most fascinating insight is the similarity between those integral proteins from procaryotes and those from eucaryotes; they appear the same in relative polarity.

Similarities as those described above emphasize that the gas vacuole membrane protein is indeed a true membrane protein of the integral-type. Although the gas vacuole membrane may be said to be only a half-membrane in the morphological sense that it contains no lipid, it does have an integral-type membrane protein.

---

[3] This operational definition divides amino acids into three classes (92, 22): Polar (Asp, Asn, Glu, Gln, Lys, Arg); Intermediate (Ser, Thr, His, Gly); Nonpolar (Ala, Val, Ile, Leu, Cys, Met, Pro, Phe, Trp, Tyr).

Krigbaum et al. (93) have developed a procedure whereby the amount of secondary structure in a globular protein may be predicted from its amino acid composition. The average errors of this method are 8.2% for helix, 8.2% for sheet, 5.5% for turns, and 5.7% for coil regions. Regions of coil and turn are as defined in Krigbaum et al. (93). The error is even less if the acid and amide side chains can be distinguished. I applied this empirical approach to the GVP. The following sets of relations were used in predicting the percent of secondary structure in the GVP based on the amino acid composition as given in Table 4; where Ha^+ and HB^+ are helix formers; HA^- and HB^- are nonhelix formers; the same holds for sheet (B^+, B^-), turn (T^+, T^-), coil (C^+, C^-) (93):

helix: HA^+ = Ala + Leu + His + Ile

HA^- = Pro + Thr + Try + Met + Tyr

HB^+ = Ala + Leu + His + Tyr

HB^- = Pro + Thr + Ile + Val + Arg

sheet: B^+ = Asp + Thr + Arg + Pro + Val + Asn

B^- = Leu + Ala + Glu + Gln + Gly

turn: T^+ = Gly + Thr + Asp + Gln + Glu + Met + Asn

T^- = Ser + Ala + Phe + His + Cys

coil: C^+ = Cys + Tyr + Ala + Thr

C^- = His + Asp + Asn

$$\% \text{ helix} = 43.02 + 0.7.07\ HA^+ + 0.676\ HB^+ - 1.223\ HA^- - 0.865\ HB^-$$

$$\% \text{ sheet} = 19.13 + 1.633\ B^+ - 1.477\ B^-$$

$$\% \text{ turn} = 25.91 + 0.904\ T^{+} - 0.909\ T^{-}$$

$$\% \text{ coil} = 24.19 + 1.125\ C^{+} - 2.041\ C^{-}$$

Using this set of equations predicts the GVP to contain 35.5% helix, 20.2% sheet, 25.5% turn, and 15.5% coil. Whether or not these calculations are valid for a hydrophobic membrane protein depends on whether comparisons can be made between integral membrane proteins and globular proteins. X-ray data for many more proteins are necessary before such assertions can be completely validated. In addition, estimation of secondary structure in the GVP has not yet been fully evaluated by other physical methods such as X-ray, CD and IR analyses (See Introduction (1.5) for current available data.). Such information, ought to indicate whether or not this prediction method is valid for nonglobular proteins.

4.8 Tentative Molecular Model for the Function of the Gas Vacuole Membranes. Based on the above data, it is possible to propose a functional model for the gas vacuole membrane. Gas vacuoles apparently do not actively regulate the flow of gas through the vacuole membrane; that is, gases appear to pass through the membrane by diffusion (40, 101). Since many diatomic gases are apolar, passage through an apolar milieu would provide a path of least resistance. This would also be the case for other gases as Ar, CH_4, and CO_2 which have been shown by Walsby to freely permeate the membrane (41). Such an apolar milieu would exist in the aliphatic, amino terminal portion of the peptide NPT. Movement of gases through the membrane

is dependent on several factors. Initially, it is necessary to assume that a subunit substructure exists in the membrane as already proposed (41, 102). Secondly, in order to allow gas to move through the membrane without a conformational change in the GVP (for which there is no evidence), either the gas must pass through the intermolecular space of the protein, or through pores. The lining of either of these passageways would be the aliphatic portion of the NPT peptide. Polar molecules would, therefore, be excluded from passing through this region. Larger apolar molecules would also be restricted based on the rigidity and size of the space which must be at least 3 $\mathring{A}$ in order to accomodate molecules as large as N_2 and CH_4 for which the van der Waals radii are known to be about 3.0 and 2.5 $\mathring{A}$, respectively.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Lenard, J. and Singer, S. J. Proc. Nat. Acad. Sci. 56, 1828 (1966).

2. Davson, H. and Danielli, J. F. J. Cell Physiol. 5, 495 (1935).

3. Robertson, J. D. Progr. Biophys. Chem. 10, 343 (1960).

4. Korn, E. D. Science, 153 1491 (1966).

5. Kavenau, J. L. Fed. Proc. 25, 1096 (1966).

6. Stockenius, W. and Engelman, D. M. J. Cell Biol. 42. 613 (1969).

7. Engelman, D. M. J. Molec. Biol. 58, 153 (1971).

8. Singer, S. J. Ann. Rev. Biochem. 43, 866 (1974).

9. James, R. and Branton, D. Biochem. Biophys. Acta 323, 378 (1973).

10. Scott, R. E. and Carter, R. L. Nature New Biol. 233, 219 (1971).

11. Dupont, Y., Harrison, S. C. and Hasselbach, W. Nature 244, 555 (1973).

12. Wallach, D. F. and Gordon, A. Fed. Proc. 27, 1263 (1971).

13. Chuang, T. F., Awasthi, Y. C., Funk, L., Crane, F.L. Biochem, Biophys. Acta 211, 599 (1970).

14. Wallach, D. F. J. Gen. Physiol. 54, 35 (1969).

15. Lenard, J., and Singer, S. J., Proc. Nat. Acad. Sci. 56, 1828 (1966).

16. Wallach, D. F. and Zahler, P. H. Proc. Nat. Acad. Sci. 56, 1552 (1966).

17. Singer, S. J. and Nicolson, C. L. Science 175, 720 (1972).

18. Spatz, L. and Strittmatter, P. J. Biol. Chem. 248, 793 (1973).

19. Spatz, L. and Strittmatter, P. Proc. Nat. Acad. Sci. 68, 1042 (1971).

20. Segrest, J. P., Kahane, I., Jackson, R. L. and Marchesi, V. T. Arch. Biochem. Biophys. 155, 167 (1973).

21. Salton, M. R.: in "Biomembranes," (ed. Manson, A.). New York: Plenum Press. (1971) p. 40.

22. Capaldi, R. A. and Vanderkooi, G. Proc. Nat. Acad. Sci. 69, 930 (1972).

23. Crane, L. J. and Lampen, J. O. Arch. Biochem. Biophys. 160, 655 (1974).

24. Tanford, C.: in "The Hydrophobic Effect: Formation of Micelles and Biological Membranes," New York: John Wiley and Sons. (1973).

25. Scheraga, H., Nemethy, G. and Steinberg, I. J. Biol. Chem. 237, 2506 (1962).

26. Bigelow, C. J. Theoret. Biol. 16, 187 (1967).

27. Singleton, R. and Amelunxen, R. E. Bact. Rev. 37, 320 (1973)

28. Goldsack, D. E. Biopolymers 9, 247 (1970).

29. Bull, H. B. and Breese, K. Arch. Biochem. Biophys. 158, 681 (1973).

30. Segrest, J. P., Jackson, R. L. and Marchesi, V. T. Biochem. Biophys. Res. Commun. 49, 964 (1972).

31. Marchesi, V. T., Tillack, T. W., Jackson, R. L., Segrest, J. P. and Scott, R. E. Proc. Nat. Acad. Sci. 69, 1445 (1972).

32. Chou, P. Y. and Fasman, G. D. Biochem. 13, 222 (1974).

33. Kotelchuck, D. and Scheraga, H. A. Proc. Nat. Acad. Sci. 62, 14 (1969).

34. Lewis, P. N., Gō, N., Gō, M., Kotelchuck, D. and Scheraga, H. A. Proc. Nat. Acad. Sci. 65, 810 (1970).

35. Finkelstein, A. V. and Ptitsyn, O. B. J. Molec. Biol. 62, 613 (1971).

36. Chou, P. Y. and Fasman, G. D. J. Molec. Biol. 74, 263 (1973).

37. Robson, B. and Pain, R. H. J. Molec. Biol. 58, 237 (1971).

38. Puett, D. Biochem. Biophys. Acta 257, 537 (1972).

39. Jones, D. D. and Jost, M. Arch. Mikrobiol. 70, 43 (1970).

40. Walsby, A. E. Proc. Roy. Soc. - Ser. B 178, 301 (1971).

41. Walsby, A. E. Bact. Rev. 36, 1 (1972).

42. Waaland, J. R., Waaland, S. D. and Branton, D. J. Cell Biol. 48, 212 (1971).

43. Weathers, P. J. unpublished data.

44. Weathers, P. J., Lalitha, S. and Haug, A. unpublished data using CD showed 24% helix, 36% sheet, and 40% coil present in inflated gas vacuole membranes.

45. Jones, D. D. Ph.D. thesis, Michigan State University (1970).

46. Jones, D. D. and Jost, M. Planta 100, 277 (1971).

47. Falkenberg, P., Buckland, B. and Walsby, A. E. Arch. Mikrobiol. 85, 304 (1972).

48. Stockenius, W. and Kunau, W. H. J. Cell Biol. 38, 337 (1968).

49. Nolan, C. and Margoliash, E. Ann. Rev. Biochem. 37, 727 (1968).

50. Dayhoff, M. O.: in "Atlas of Protein Sequence and Structure," Washington, D. C.: National Bio Medical Research Foundation. (1968-1972).

51. Wu, T. T., Fitch, W.M. and Margoliash, E. Ann. Rev. Biochem. 43, 539 (1974).

52. Liljas, A. and Rossman, M. G. Ann. Rev. Biochem. 43, 475 (1974).

53. Tschesche, H. and Kupfer, S. Eur. J. Biochem. 26, 33 (1972).

54. Callahan, P. X., McDonald, J. K. and Ellis, S. Fed. Proc. 31, 1105 (1972).

55. Paukovits, W. R. J. Chrom. 85, 154 (1973).

56. Edman, P.: in "Molecular Biology, Biochemistry and Biophysics: Protein Sequence Determination," (ed. Needleman, S. D.) New York: Springer-Verlag (1970).

57. Edman, P. and Begg, G. Eur. J. Biochem. 1, 80 (1967).

58. Inman, J. K., Hannon, J. E. and Appella, E. Biochem. Biophysics. Res. Commun. 46, 2075 (1972).

59. Braunitzer, G. Schrank, B. and Ruhfus, A. Hoppe-Seyler's Z. Physiol. Chem. 351, 1589 (1970).

60. Foster, J. A., Bruenger, E., Hu, C. L., Albertson, K. and Franzblau, C. Biochem. Biophys. Res. Commun. 53, 70 (1973).

61. Wachter, E., Machlerdt, W., Hofner, H. and Otto, J. FEBS Lett. 35, 97 (1973).

62. Previero, A., Derancourt, J. and Coletti-Previero, M-A. manuscript in preparation.

63. Gorham, P. R., McLachlin, J., Hammer, U. T. and Kim, W. K. Verh. int. Ver. theor. angew. Limnol. 15, 796 (1964).

64. Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. J. Biol. Chem. 193, 265 (1951).

65. Gray, W. R.: in "Methods in Enzymology," (ed. Hirs, C. H. W.) Vol. 11 p. 139. New York: Academic Press (1967).

66. Schroeder, W. A., Jones, R. T., Cormick, J. and McCalla, K. Anal. Biochem. 34, 1571 (1962).

67. Butler, P. J. G., Harris, J. I., Hartley, B. S. and Leberman, R. Biochem. J. 112, 679 (1969).

68. Ramachandran, L. K. and Witkop, B.: in "Methods in Enzymology," (ed. Hirs, C. H. W.) Vol. 11 p. 290. New York: Academic Press (1967).

69. Offord, R. E. Nature 211, 591 (1966).

70. Foster, J. A. personal communication.

71. Schultz, J.: in "Methods in Enzymology," (ed. Hirs, C. H. W.) Vol. 11 p. 255 New York: Academic Press (1967).

72. Fletcher, J. C. Biochem. J. 88, 34 c (1966).

73. Moore, S., Spackman, D. H. and Stein, H. H. Anal. Chem. 30, 1185 (1958).

74. Schmidt, P. I.: in "Techniques in Amino Acid Analysis," Chertsey: Technicon Instruments Company Ltd. (1966).

75. _____ Beckman Model 890c Sequencer Instruction Manual, January 1972.

76. Pisano, J. J. and Bronzert, T. J. Anal. Biochem. 45, 43 (1972).

77. Smithies, O., Gibson, D., Franning, E. M., Goodfliesh, R. M., Gilman, J. D. and Ballantyne, D. L. Biochem. 10, 4912 (1971).

78. Inglis, A. S., Nicholls, P. W. and Roxburgh, C. M. Aust. J. Biol. 24, 1247 (1971).

79. Jones, D. D., Haug, A., Jost, M. and Graber, D. R. Arch. of Biochem, Biophys. 135, 296 (1969).

80. Krantz, M. J. and Ballou, C. E. J. of Bact. 114, 1058 (1973).

81. Khandwala, A. S. and Kasper, C. B. Biochem. Biophys. Acta 233, 348 (1971).

82. Gross, E. and Witkop, B. J. Biol. Chem. 237, 1856 (1962).

83. Smyth, D. G., Stein, W. H. and Moore, S. J. Biol. Chem. 237, 1845 (1962).

84. Carruway, K. L. and Koshland, D. E.: in "Methods in Enzymology," (ed. Hirs, C. H. W. and Timasheff, S. N.) Vol. 25 p. 616 New York: Academic Press. (1972).

85. Gray, W. R., Sandberg, L. B. and Foster, J. A. Nature 246, 461 (1973).

86. Kasper, C. B.: in "Molecular Biology Biochemistry and Biophysics: Protein Sequence Determination," New York: Springer-Verlag. p. 142 (1970).

87. Vanderkooi, G. and Sundarlingam, M. Proc. Nat. Acad. Sci. 67, 233 (1970.

88. Hinman, N. D. and Phillips, A. H. Fed. Proc. 31, Abs. 1086.

89. Sanderman, H., Jr. and Strominger, J. L. Proc. Nat. Acad. Sci. 68, 2441 (1971).

90. Perutz, M., Kendrew, J. and Watson, H. J. Mol. Biol. 13, 669 (1965).

91. Margoliash, E. and Schejter, A. Advan. Prot. Chem. 21, 113 (1966).

92. Tanford, C. J. Am. Chem. Soc. 84, 4240 (1962).

93. Krigbaum, W. R. and Knutton, S. R. Proc. Nat. Acad. Sci. 70, 2809 (1973).

94. Schmidt, G. L., Nicolson, G. L. and Kamen, M. D. J. Bact. 105, 1137 (1971).

95. Niederman, R. A., Segen, B. J. and Gibson, K. D. Arch. Biochem. Biophys. 152, 547 (1972).

96. Eng, L. F., Chao, F. C., Gerstl, B., Pratt, D. and Tavastsjerna M. G. Biochem. 7, 4455 (1968).

97. Eichberg, J. Biochem. Biophys. Acta 187, 533 (1969).

98. Tenembaum, D. and Folch-Pi, J. Biochem. Biophys. Acta 115, 141 (1966).

99. Wolfgram, F. and Rose, A. S. J. Neurochem. 8, 161 (1961).

100. Gagnan, J., Finch, P. R., Wood, D. D. and Moscarello, M. A. Biochem. 18, 4756 (1971).

101. Walsby, A. E. Proc. Roy. Soc. Ser. B 173, 235 (1969).

102. Jost, M. and Jones, D. D. Can. J. Microbiol. 16, 159 (1970).

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 03178 1622