This is to certify that the

thesis entitled

EVIDENCE FOR POLYHYDROXYPROLINE IN

THE EXTRACELLULAR MATRIX OF VOLVOX

presented by

Barbara Ann Mitchell

has been accepted towards fulfillment
of the requirements for

M.S. degree in Botany and Plant Pathology

David T. A. Lamport
Major professor

Date May 8, 1980

EVIDENCE FOR POLYHYDROXYPROLINE IN THE EXTRACELLULAR MATRIX OF *VOLVOX*

By

Barbara Ann Mitchell

A THESIS

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

MASTER OF SCIENCE

Department of Botany and Plant Pathology

1980

ABSTRACT

EVIDENCE FOR POLYHYDROXYPROLINE IN
THE EXTRACELLULAR MATRIX OF *VOLVOX*

By

Barbara Ann Mitchell

Hydroxyproline rich structural glycoproteins exist throughout the plant and animal kingdoms: as extensin in plant cell walls and collagen in animals. *Volvox* is phylogenetically intermediate in position between plants and animals, thus its hydroxyproline rich extracellular matrix is ideal for studying the chemical relationships of extensin-like and collagen-like proteins.

The extracellular matrix was obtained by detergent extraction of *Volvox*. Removal of carbohydrate from the matrix glycoprotein permitted enzyme digestion and sequencing of the polypeptides by automatic Edman degradation.

The major polypeptide, as partially sequenced, contained 750 to 1000 residues of hydroxyproline with other amino acids present in minute amounts. The amino acid content resembles extensin; the peptide length approximates collagen. Polyhydroxyproline and extensin form a polyhydroxyproline type-II helix, as does collagen.

Although these properties are consistent with the intermediate position of Volvox, the matrix glycoprotein is not the archetypal protein directly relating collagen and extensin.

ACKNOWLEDGEMENTS

I wish to thank Dr. D. T. A. Lamport for his continued support and enthusiasm and for providing me with an opportunity to study Volvox.

I also wish to thank Dr. U. V. Mani, Yukio Akiyama, Sharon Mohrlok, Dr. Jim Catt and Dr. Michael Jermyn for their advice and contributions.

Most importantly I wish to thank my husband, Dr. Michael Zaroukian, for his insightful discussions and constant support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGPs	Arabinogalactan proteins
G-25, G-50, G-100	Sephadex gels
GLC	Gas liquid chromatography
HF	Hydrogen fluoride
HPLC	High pressure liquid chromatography
Hyp	Hydroxyproline
MW	Molecular weight
R.T.	Room temperature
S.D.	Standard deviation
S.D.S.	Sodium dodecylsulfate
Ser	Serine
SP-C-25	Sulfopropyl Sephadex C-25
TCA	Trichloroacetic acid

INTRODUCTION

Volvox is classified as a plant by Whittaker (46) but more currently as a protist by Margulis (28) and Kudo (17) (Table 1). It possesses characteristics of both the plant and animal kingdoms and it is this intermediate position which makes the chemical characteristics of its extracellular matrix of intrinsic interest (Figure 1).

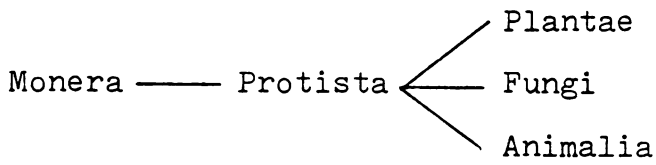
Volvox is a spherical, multicellular organism composed of two cell types: somatic chlamydomonad-like (19) cells and reproductive gonidia. Volvox is more appropriately described as a spheroid than as a colony (14, 45). In the vegetative organism the inner core of the spheroid is a gelatinous extracellular matrix, the outer mantle being comprised of several thousand (41) somatic cells each bounded by a thin cell wall and envelope. The cell wall and envelope are embedded in the peripheral regions of the gelatinous matrix. Eight to twelve gonidia are located in the matrix core. A highly structured sheath surrounds the entire organism and maintains its integrity (15, 19, 33, 40). It is the chemical composition of the gelatinous extracellular matrix that is the subject of this study.

During ontogeny, the Volvox spheroid undergoes a complete inversion reminiscent of gastrulation and

Table 1. Classification of Volvox (17, 34)

Kingdom:	Protista	Plantae
Phylum:	Protozoa	Chlorophyta
Subphylum:	Plasmodrama	Chlorophyceae
Class:	Mastigophora	- -
Subclass:	Phytomastigia	- -
Order:	Phytomonadid	Volvocales
Family:	Volvocidiae	Volvocaceae
Genus:	Volvox	

Figure 1. Classical Five Kingdom Scheme of Evolution (46)



neurulation in higher animals (43). Morphogenesis in Volvox carteri begins with an increase in size followed by cleavage of the gonidium (41). Cleavage and establishment of polarity is very specific and is described in detail by Darden 1966, Kochert 1968 and Starr 1970. Cytokinesis is incomplete during cleavage and the cytoplasmic bridges which form remain until the organism is mature. Shortly after completion of cleavage, inversion takes place as described by Darden 1966, Pickett-Heaps 1970, Viamontes and Kirk 1977 and Viamontes and Kirk 1979. The right-side-out organism begins to actively secrete matrix allowing for enlargement of the daughter spheroid, enhancing the separation of the vegetative cells. Although the vegetative cells increase in size, the bulk of the spheroid growth is a result of matrix deposition. Yates et al. (45) grew Volvox carteri synchronized on a 72 hour light - 36 hour dark cycle and measured the amount of protein and carbohydrate per milliter of culture medium at different growth stages. Protein content remained constant during spheroid growth (light period) but carbohydrate content increased (which suggests the matrix is composed largely of carbohydrate). During gonidial cleavage (light period) both the protein and carbohydrate content increased at a greater rate. Both became relatively constant in amount during daughter spheroid release (dark period). Mature daughter spheroids are released individually through ruptures in the parental somatic cell layer.

Electron microscopic studies of the sheath show it to be a tripartite structure (30). An electron dense fibrillar layer is located just internal to the tripartite structure. The somatic cells are surrounded by an electron dense fibrillar layer (the cell wall) which is continuous with the flagellar channel. A loose, widely spaced fibrillar matrix fills the rest of the spheroid. In addition the vegetative cell and gonidia are surrounded by separate fibrous envelopes, the cellular envelope and vesicle respectively; both are distinct from the rest of the matrix (19, 30, 40).

The extracellular matrix, composed of loose fibrillar material appears to be secreted in discrete layers by the golgi of the vegetative cells. Vesicles containing such fibrillar material derived from the golgi have been seen to fuse with the plasmalemma (40) and to be extruded as membrane bound vesicles which release their contents between the somatic cell and sheath (4).

The sheath, which has been determined to be at least partially composed of acid mucopolysaccharide and/or polysaccharide (30) may be important as a semipermeable barrier and may contain specific recognition sites for receptor binding and other membrane mediated phenomena (7, 15, 41). McCracken detected a sulfated mucopolysaccharide layer external to the tripartite structure in Volvox rousseletii, not detected in other species.

Concanavalin A binding to Volvox carteri alters the membrane surface properties such that induction of sexual development and embryo inversion are inhibited. The exclusive Concanavalin A (and to a lesser extent lentil lectin) binding to the cytoplasmic membrane of the somatic cell, the cellular envelope, the flagellar membrane, specific structures within the extracellular matrix and to a much lesser extent the entire extracellular matrix, indicates a high degree of extracellular control and may indicate a common glycosylation pattern for the surface glycoconjugates (18).

The extracellular matrix is rich in hydroxyproline, containing approximately 5% by dry weight (25). Hydroxyproline, synthesized as a result of post-translational modification of proline occurs as a major component of certain extracellular glycoproteins. It is present in collagen, a constituent of the extracellular matrix in animal systems and in extensin, a plant cell wall glycoprotein. Both these glycoproteins contribute to the rigidity and resilience of the organism. Both contain hydroxyproline as a major component, which is responsible, in part, for the secondary structure of a polyproline II-type helix (23, 38). The hydroxyproline rich glycoprotein in the cell wall of Chlamydomonas, a member of the Volvocaceae, partially exists in the left handed polyproline II-type helix (94 nm pitch, 3 residues per turn) (12).

Collagen is a family of glycoproteins that exists in structurally different forms in different tissues of the same species (5). Collagen is generally composed of three polypeptide chains which are characterized for greater than 95% of their length by glycine in every third position. It is only at the amino and carboxyl termini that there is an absence of the repetitive pattern. While procollagen, the biosynthetic precursor of collagen, is still attached to the polysomes, the amino terminus enters the rough endoplasmic reticulum where specific peptidyl prolines are hydroxylated, probably stabilizing the helix. The chains align at their amino termini and super coiling begins. The triple helix formation inhibits prolylhydroxylation. At some uncertain point the polypeptide is released from the polysome. After the structure is stabilized it is secreted and the amino and carboxyl termini are cleaved by proteolytic enzymes. The collagen is incorporated into fiber by additional cross-linkages (5, 36). Collagen of lower animals is characterized by a high degree of glycosylation, with the exception of Ascaris. Higher animal collagen is often less than 2% carbohydrate.

The presence of hydroxyproline in the cell wall fraction of higher plants (35) and algae, with the exception of Characeae, Rhodophyceae and Phaeophyceae (37, 42, 10) indicates its presence is not trivial and it has been suggested by Lamport that an extensin-like protein is present in cell walls throughout the plant kingdom.

Extensin fragments have been isolated from the primary cell wall of tomato by Lamport in the form of five low molecular weight peptides which account for about one-third of the wall bound hydroxyproline. The common feature of the extensin peptides is the repeating sequence of Ser-Hyp-Hyp-Hyp-Hyp.

Other extracellular glycoproteins which contain significant amounts of hydroxyproline, in the plant kingdom, are "classical" lectins and arabinogalactan proteins (AGPS) also called β lectins.

Comparison of the two glycoproteins, collagen and extensin raises the question of their evolutionary origin(s); are these proteins homologous or analogous? Did extensin and collagen arise from a common origin via divergent evolution or did they arise from wholly different evolutionary origins, yet come to serve similar structural functions? (Table 2, Lamport 1977).

As is true for many lower forms of life, Volvox is not neatly categorized. Botanists claim Volvox as a plant, zoologists classify it as a protist. Volvox is referred to as a multicellular colony and multicellular organism. It has been regarded as an evolutionary dead end and even suggested as the origin of the blastula.

Volvox is an obligate photoautotroph with a highly glycosylated hydroxyproline rich extracellular matrix. This complex of properties places a Volvox-type organism at the

Table 2. Do Collagen and Extensin Share a Common Origin, i.e., are they in the Same Protein "Super Family"?

1.	**Position:**	**Extracellular matrix**
2.	**Function:**	**Structural**
3.	**Chemically:**	**Hydroxyproline-rich**
4.	**Structure:**	**Glycoproteins** **Helix type: Polyproline II (9.4 A pitch, 3 residues/turn)**
5.	**Biosynthesis:**	***No* hydroxyproline codon (implies "late" evolution)** **Post-translational modification** **Involves very similar proline hydroxylases utilizing molecular O_2 and cofactors Fe^{++}, ascorbate and α-ketoglutarate**
6.	Occurrence:	Extensin-like proteins occur in most photosynthetic protists Collagen does *not* occur in protozoans, i.e., HYP occurrence correlates with photosynthetic ability
7.	Evolutionary Branch Point:	*Volvox*-like organism? (Origin of the blastula?)

evolutionary branch point between plants and animals (Figure 1) (Figure 2, Lamport 1977).

Some animal-like characteristics of Volvox are:

(1) Cell division via a cleavage furrow with no cell plate formation,

(2) Formation of a hollow sphere of cells reminiscent of a blastula,

(3) Embryonic morphogenesis involving inversion via cell shape changes recalling gastrulation and neurulation,

(4) An apparent requirement for vitamin B_{12}.

Because of these unique properties, Volvox has long been recognized as a prime eukaryotic organism for the study of:

(1) Embryonic morphogenesis and the associated cellular processes regulating morphogenesis,

(2) Sexual and asexual determination,

(3) Cellular differentiation, development and senescence,

(4) Extracellular structural matrices.

To date, hundreds of mutants have been isolated (39) of Volvox carteri f. nagariensis thus facilitating these studies.

This thesis comprises a study of the extracellular matrix of Volvox including its chemical composition, form and function; the objective was to attempt to determine

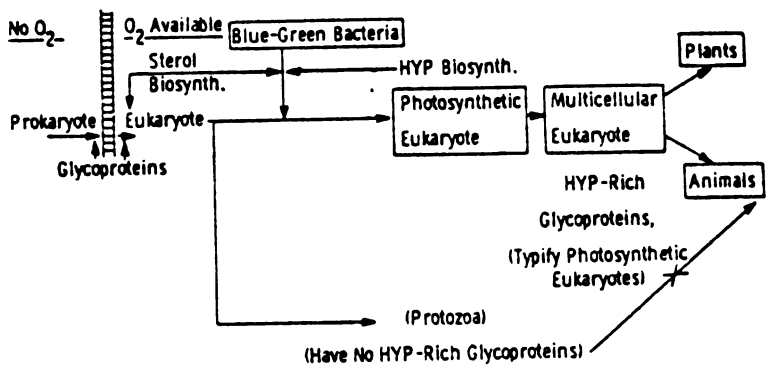


Figure 2. A Hypothetical Scheme for the Molecular Evolution of Eukaryotes.

the origin of the extracellular glycoproteins, extensin and collagen, through chemical analysis of the extracellular matrix of Volvox.

MATERIALS AND METHODS

Materials

Volvox carteri f. nagariensis, female, vegetative state, was a gift from David L. Kirk, University of Washington, St. Louis, Mo. All chemicals were analytical or sequential grade or the best commercially available. Sephadex G-25, G-50 and G-100 (fine), SP-Sephadex C-25, Sepharose CL-4B and Blue Dextran were purchased from Biorad. Gas liquid chromatography column support, Gas Chrom Q and A and coating materials SP-2100 and OV-275 were purchased from Supelco Co, Bellefonte, Pa. Trypsin was purchased from Worthington Biochemical Corporation, N.J. Protease from Streptomyces griseus type IV was purchased from Sigma Chemical Co. Fluorescamine was purchased from Pierce Chemical Co. Polyhydroxyproline was purchased from New England Nuclear Corp., Mass.

Volvox: Culture and Preparation

Volvox carteri was grown on a sixteen hour light, eight hour dark schedule. Release of daughter spheroids occurred every forty-eight to seventy-two hours. Single spheroids were grown in screw cap culture tubes and transferred to 125 ml then 500 ml Erlenmeyer flasks then 2800 ml Fernbach flasks successively, every ten to fourteen days. The growth medium presented in Table 3 is a modification of

Table 3. *Volvox* Medium

Nutrient	Molarity
$Ca(NO_3)_2$	0.50 mM
$MgSO_4$	0.16 mM
Na_2glycerophosphate	0.23 mM
KCl	0.67 mM
Tricine	5.0 mM
Vitamin B_{12}	9.30 pM
Biotin	1.25 nM
Thiamine	3.06 μM
Na_2CO_3	0.19 mM
$FeCl_3$	2.15 μM
$MnCl_2$	1.24 μM
$ZnCl_2$	0.22 μM
$CoCl_2$	0.05 μM
Na_2MoO_4	0.97 μM

Final pH adjusted to 8.0

Provasoli and Pintner's Volvox medium. Tricine, N-tris (hydroxymethyl) methylglycine, pK 8.15, was chosen as the buffer because, unlike tris, it forms soluble complexes with heavy metals and is not hydrolyzed to glycine (as is glycylglycine) thus avoiding potential glycine contamination in analyses. The spheroids were harvested every ten to fourteen days by filtration through a 10 μ Nitex screen fitted into a Büchner funnel. The resulting dense (approximately 10^4 spheroids/ml) suspension was transferred to a beaker, diluted 1:1 with a pH 6.8 SDS-mercaptoethanol, Tris-HCl buffer and heated in a boiling water bath for five minutes. The suspension was filtered through a 10 μ Nitex screen and the residue thoroughly rinsed with Volvox medium. The residue retained by the screen consisted of intact colorless Volvox spheroids, termed ghosts (Kirk, private communication 1979). Both the ghosts and ghost filtrate were dialysed extensively against distilled water. The ghosts were prepared with two different strength buffers. (1) With 1% SDS 2.5% mercaptoethanol pH 6.8 buffer the spheroids remained structurally intact during the dialysis and were dounce homogenized with a tight fitting Köntes dounce homogenizer (maximum clearance 0.0012 inches) then centrifuged at 10,000 g for 45 minutes. The more dense, loose "pellet" which contained 60-70% of the hydroxyproline was freeze dried and used for the carbohydrate analyses and early protein analyses. (2) Extraction with 2% SDS 5% mercaptoethanol pH 6.8 buffer resulted in

ghosts which lost all structural integrity upon extensive dialysis against distilled water. This structureless suspension was centrifuged at 20,000 g for 45 minutes and the supernatant was freeze dried and used for subsequent protein analyses. Figure 3 is the flow diagram for ghost preparation.

Hydroxyproline rich material was recovered from the medium as follows: twenty liters of growth medium was rotary evaporated to 200 mls and made up to one liter with 95% ethanol and was left overnight at 4° C. The precipitate was washed three times with 95% ethanol then lyophilized and chemically analyzed.

General Methods

Enzyme Digest

Trypsin - 100 mg of deglycosylated sample was taken up in 5 ml of freshly prepared 2% ammonium bicarbonate. The solution was made up to 10mmolar with 100 µl of $CaCl_2$ solution. The pH of the solution was 8.0 ± 0.2. One mg of trypsin was added and the flask was stoppered and incubated at 30.5° C overnight on a gyrotary shaker. The solution was centrifuged at 15,600 g for 15 minutes to remove any insoluble material.

Pronase - Peptides containing approximately 1 mg of hydroxyproline were taken up in 900 µl of freshly prepared buffer (50 mmolar NH_4HCO_3 containing 2 mmolar $CaCl_2$). Fifty µg of pronase was added and the stoppered flask was

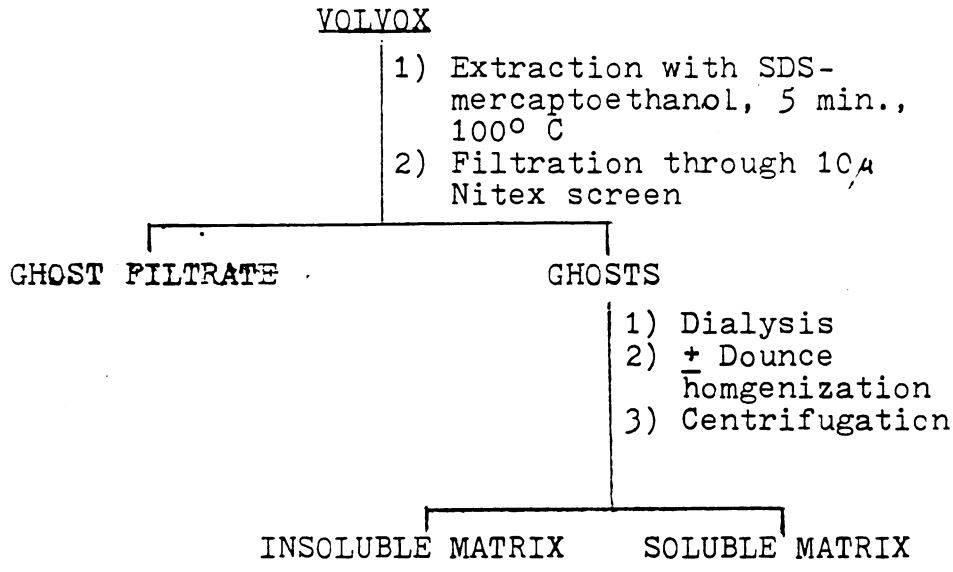


Figure 3. Ghost Preparation

incubated at 30.5° C overnight on a gyrotory shaker. The solution was centrifuged at 15,600 g for 15 minutes to remove any insoluble material.

Acid Hydrolysis

The sample was placed in a Köntes microflex tube with 5.5 N HCl. The vial and sample was flushed with N_2, tightly sealed with a tefbond septum and cap, and heated to 110° C for the specified period of time. The hydrolyzate was cooled to room temperature and blown dry under a stream of N_2 at 40° C.

Performic Acid Oxidation

The sample, containing less then 1 mg of protein was added to a 1 ml microflex tube with 20 µl performic acid (0.5 ml 30% w/w hydrogen peroxide plus 10 ml formic acid) which had been allowed to stand for one hour at room temperature. The reaction mixture was incubated 20 minutes at room temperature. Four hundred µl of water was added to quench the reaction which was then blown dry under N_2.

Deglycosylation

HF/pyridine - The matrix was routinely deglycosylated in 70% HF/pyridine according to the procedure of U.V. Mani et al. (29) developed in this laboratory. Material to be deglycosylated was thoroughly desiccated, then stirred continuously in HF/pyridine with anhydrous methanol (as a scavenger for glycosylfluorides, preventing polymerization

and alkylation of aromatic residues) in a ratio of 10:1 for 90 minutes at room temperature. The reaction was quenched with water and dialysed against distilled water for 2 days with 8 water changes, then centrifuged at 10,000 g for 45 minutes.

Anhydrous HF - When it was necessary to recover the sugars, deglycosylation was performed with anhydrous HF according to the procedure of Mort and Lamport (31). The reaction was allowed to go one hour at 0° C to recover neutral and acidic sugars or three hours at room temperature to obtain amino sugars. After the HF was completely evaporated the residue was taken up in 0.01 N NH_4OH and centrifuged at 15,600 g for 15 minutes to remove any insoluble material.

Hydroxyproline Estimation

The method of Lamport and Miller (21), involved oxidative decarboxylation and dehydroxylation of hydroxyproline to give a pink colored pyrrole whose absorbance was monitored at 560 nm. Sample preparation entailed either an 18 hour acid hydrolysis or, for column eluents, a one hour alkaline hydrolysis (5N NaOH) at 121° C. The alkaline hydrolyzate was neutralized with 5 N HCl. Hydroxyproline standards were run with each set of samples.

Amino Group Detection

The amino groups of peptides from column chromatography were reacted with fluorescamine in a basic buffer.

Fluorescence was detected by a Gilson fluorometer, with a maximum sensitivity of 1 nanomole norleucine equal to full scale deflection (8).

Analytical Methods

Column Chromatography

Gel permeation - Unless otherwise stated, gel permeation chromatography was performed using columns of 100 ml volume (123 x 1.2 cm) eluted with 0.1 N NH_4OH. Two-ml fractions were collected. Exclusion limits were determined with Blue Dextran and picric acid M.W. 188.

Ion exchange - SP-Sephadex C-25 cation exchange chromatography was performed as follows. The sample was dissolved in 0.01 N HCl and centrifuged to remove any insoluble material. The sample was applied to a column of 30 ml volume (1.2 x 26 cm) and washed with 1.5 bed volumes of 0.01 N HCl then eluted with a five bed volume ionic gradient of 0.01 N HCl to 1 N NaCl in 0.01 N HCl. Three-ml fractions were collected. Aminex A5 cation exchange chromatography was performed according to the procedure of Lamport et al.(22). The sample was adjusted to pH 2.5 and applied to an Aminex A5 column eluted with a pyridine acetate pH gradient of 2.7 to 5.0 at 60° C, column pressure less than 600 PSI.

Sugar Analysis via Gas Liquid Chromatography

Alditol acetate - The method described by Albersheim (2) involved a sixty minute trifluoroacetic acid hydrolysis at

121° C, reduction of the monosaccharides to their alcohols and then acetylation with glacial acetic acid. A 6' x 1/8" column packed with 0.2% poly(ethylene glycol adipate) 0.2% poly(ethylene glycol succinate) and 0.4% XF 1150 on Gas Chrom Q was employed with the temperature programmed from 130° C to 195° C at 2° per minute.

Trimethylsilylated methyl glycosides - The method of Bhatti et al. (3) involved a ninety minute methanolysis in 1.5 N methanolic hydrochloric acid at 95° C. The anhydrous methylglycosides were trimethylsilylated and chromatographed on a 12' x 1/8" column packed with 3% SP-2100 on Gas Chrom Q, the initial temperature was held at 120°C for four minutes, then programmed 1° per minute to a final temperature of 185° C.

Amino Sugar Analysis

Gas liquid chromatography - Amino sugars were analyzed as N-acetylated alditol acetates (same protocol as for alditol acetates of neutral sugars), chromatographed on a 6' x 1/8" column packed with 1% OV-275 on unsilanized Gas Chrom A with the temperature programmed from 130° to 230° C at 2° per minute. Amino sugars were also analyzed after N-protonation in pyridine/methanol then acetylation with acetic anhydride prior to trimethylsilylation of the methylglycosides, as described by Kozulic et al. (16), and chromatographed on the SP-2100 column.

High pressure liquid chromatography, automated amino acid analyzer (Dionex) - The sample was hydrolyzed for three hours at 100° C in 5.5 N HCl. The supernatant was blown dry under N_2 and was taken up in 40 µl water with 20 nanomoles norleucine as an internal standard and applied to a DC 5A column under approximately 1600 PSI via a 20 µl injection loop. The eluent from the column was automatically reacted with ninhydrin and monitored at 540 and 420 nm. The peak areas were computed as nanomoles of amino acid by an SP-4100 computing integrator.

Amino Acid Analysis

Gas liquid chromatography as heptafluorobutyric isobutyl esters - The sample was subjected to the 18 hour acid hydrolysis, blown dry under N_2 at 40° C and derivitized by the method of Mackenzie and Tenaschuk (26, 27), which involved esterification of the carboxy groups with 3 N HCl in isobutanol, thenacylation of the amino groups with heptafluorobutyric anhydride. The sample was chromatographed on the SP-2100 column with a temperature program of 95° C to 240° C at 8° per minute, with an initial delay of 4 minutes.

High pressure liquid chromatography, automated amino acid analyzer (Dionex) - The sample was subjected to the 18 hour acid hydrolysis, blown dry under N_2 at 40° C and taken up in 40 µl of water containing 20 nanomoles of norleucine as an internal standard. The sample was injected

and analyzed as described under amino sugar analysis via HPLC automated amino acid analyser (Dionex).

High pressure liquid chromatography, automated amino acid analyzer (Technicon) - The sample was acid hydrolyzed by the standard procedure for 18 hours. The hydrolyzate was blown dry under N_2 and was taken up in 200 µl 0.001 N HCl. Aliquots were applied to a 73 x 0.6 cm column filled with Chromobead C with 100 nanomoles norleucine as an internal standard. The column was operated under 200-600 PSI at 60° C, and was eluted with a pH-ionic strength gradient as described by Lamport (20). The eluent was fed into an automated ninhydrin analyzer and monitored at 560 and 420 nm. The peak areas were computed by a Spectra Physic System IV Autolab Integrator.

Cyanoethylation

Twenty-five micrograms of peptide were incubated with 100 µl 0.1 M trimethylamine and 25 µl acrylonitrile overnight at 37° C. The reaction mixture was then blown dry under N_2 and hydrolyzed in 200 µl 5.5 N HCl for 18 hours at 110° C. Amino acid analysis was performed by HPLC automated amino acid analyzer (9).

Protein Sequencing

Peptide sequencing via automatic Edman degradation with HPLC analysis of the phenylthiohydantoin amino acid derivatives were graciously performed by Dr. Michael Jermyn and Sharon Mohrlok. The PTH amino acid derivatives were

analyzed by reverse phase chromatography on a 4.6 cm x 25 cm C_{18} column at 60° C, under approximately 500 PSI, eluted with acetonitrile buffer. The amino acid derivatives were monitored at 254 nm, and were automatically recorded and computed as peak area by an SP-4100 computing integrator.

All gas liquid chromatography was performed with a Perkin-Elmer 900 or 910 gas chromatograph fitted with dual columns, with the output connected to a Spectra Physics System IV Autolab integrator or Spectra Physics SP-4100 computing integrator.

RESULTS

Chemical Analysis of Volvox Culture Medium

The isolation, purification and characterization of a protein is made many-fold more straight forward if the protein can be easily solubilized. Cell walls and connective tissues, both highly cross-linked, are notorious for their insolubility. One might expect less cross-linked, more soluble glycoproteins to be present in the growth medium of cell or tissue cultures as precursors, degradation products or as functionally or structurally closely related glycoproteins. Such is the case for the arabinogalactan proteins (AGPs) secreted into the medium by cultured sycamore cells. AGPs (or β-lectins) are hydroxyproline rich highly glycosylated soluble glycoproteins present throughout the plant kingdom (13). The proposed model for the AGPs is a glycoprotein with a hydroxyproline rich region and a hydroxyproline poor region. There are significant differences between extensin and AGPs in their amino acid content, glycosylation and solubility properties, yet there is a distinct possibility that a sequence homology exists between extensin and the hypothetical hydroxyproline rich regions of the AGPs (1).

A major goal in this study was to obtain soluble, low molecular weight matrix peptides which could be sequenced,

and compared to collagen and extensin. Because hydroxyproline is the "marker" amino acid in these three glycoproteins, it was the amino acid assayed for through all purification and fractionation steps. An attempt was made to obtain soluble, low molecular weight hydroxyproline rich material either in the form of extracellular material secreted into the medium, as are AGPs, or as autolyzed degradation products of senescent spheroids. Protein precipitated from the medium with ethanol was analyzed for sugar and protein (Table 4) then deglycosylated with HF/pyridine. After removal of the sugars, the ethanol precipitated protein remained in solution in the dialysis bag. The sample was lyophilized and its amino acid content (Table 5) and solubility properties (Table 6) were determined. The ethanol precipitated material contained significant amounts of hydroxyproline (approximately 15% of all the amino acids present were hydroxyproline). The material was about 15% protein and 15% carbohydrate by dry weight; that is, 1 mole of sugar per mole of protein. This was the minimal amount of carbohydrate present because the sugars were obtained by trifluoroacetic acid or methanolic hydrochloric acid hydrolysis, not anhydrous HF deglycosylation. After HF/pyridine deglycosylation the material was especially soluble in formic acid. It was also significantly soluble in chaotropic salt solutions. It was not soluble in nonpolar solvents. Fifty milligrams of the lyophilized, deglycosylated sample was taken up in 3 ml of 22% formic

Table 4. Ethanol Precipitated Medium: Amino Acid and Sugar Analyses

Amino Acid	Mole %
Asp	8.4
Thr	5.9
Ser	6.6
Glu	8.1
Pro	5.1
Gly	8.6
Ala	9.2
Val	6.8
Cys	0.7
Met	0.9
Ile	3.6
Leu	6.7
Tyr	2.2
Phe	3.1
Lys	8.1
His	2.9
Arg	4.6
Hyp	11.2

15% protein by dry weight

Sugar	Mole %
Arabinose	10.1
Rhamnose	3.4
Fucose	1.3
Xulose	7.4
Mannose	56.5
Galactose	12.2
Glucose	9.0

15% carbohydrate by dry weight

Table 5. Ethanol Precipitated, Deglycosylated Medium: Amino Acid Analyses Before and After Sephadex G-100 Fractionation

Fraction: Amino Acid	Freeze Dried Medium Mole %	Soluble in 22% Formic Acid	G-100 A Mole %	G-100 B Mole %	G-100 C Mole %
Asp	10.5	9.4	9.2	9.8	10.2
Thr	6.5	5.6	6.4	6.1	6.4
Ser	10.4	9.4	8.3	10.7	6.7
Glu	7.8	7.1	6.5	7.2	10.6
Pro	n.d.	5.8	5.8	4.9	5.5
Gly	7.9	9.4	8.1	7.9	9.2
Ala	10.7	9.6	9.0	10.9	10.1
Val	5.9	5.9	6.1	6.9	7.1
Cys	n.d.	0.6	n.d.	0.7	1.3
Met	n.d.	n.d.	n.d.	n.d.	1.1
Ile	3.2	2.7	3.1	3.2	3.3
Leu	7.2	7.0	7.0	7.2	6.5
Tyr	2.3	2.1	2.4	2.0	1.6
Phe	3.4	3.2	3.4	3.4	2.9
Lys	2.9	2.8	2.7	4.2	3.6
His	0.8	0.7	0.7	0.9	1.0
Arg	4.3	4.3	4.5	5.1	4.5
Hyp	16.2	14.4	16.6	10.1	8.4
% Protein:	58%	30%	67%	42%	100%

n.d. not determined

Table 6. Solubility Properties of the Ethanol Precipitated, Deglycosylated Medium

Solvent	Molarity	+% Soluble
50% Acetic Acid	8.3	97.7
88% Formic Acid	19.1	100.0
44% Formic Acid	9.6	100.0
22% Formic Acid	4.8	100.0
11% Formic Acid	2.4	70.0
Guanidine HCl	20.0	91.6
KSCN	10.0	88.8
KSCN	5.0	92.3
KSCN	2.5	85.6
KSCN	1.3	75.4
KSCN	0.6	75.9
KSCN	0.3	67.7
KI	6.0	87.7
Urea	8.0	98.6
LiCl	10.0	57.5
10% TCA	0.6	52.0
Phenol:Acetic acid: water 2:1:1	v:v:v	91.1
DMSO		86.0
Glycerol		11.0
n-Butanol		6.0

+% solubility measured as $\frac{\mu g \text{ Hyp in supernatant}}{\mu g \text{ Hyp in supernatant plus pellet}}$

after centrifugation at 15,600 g for 15 minutes.

acid, centrifuged at 15,600 g for fifteen minutes and was applied to a 100 ml column, 123 x 1.2 cm, of Sephadex G-100 calibrated with Blue Dextran and picric acid, eluted with 22% formic acid. Two ml fractions were collected and aliquots were analyzed for hydroxyproline. The majority of the hydroxyproline voided the column with an apparent high molecular weight. The elution profile (Figure 4) shows an assortment of hydroxyproline containing peptides (with a range of MW) not clearly separated under these conditions. The variance of hydroxyproline content across the elution profile (Table 5) is consistent with the model proposed for the arabinogalactan proteins, that is, a highly glycosylated portein containing a hydroxyproline rich core and a hydroxyproline poor tail. Figure 5 is a flow sheet for the partial purification of hydroxyproline rich proteins from the medium.

Hydroxyproline Distribution in Volvox; Chemical Analysis of the Ghost Filtrate

When investigating a fraction obtained from the medium one is initially compelled to speculate from whence it originated. Analysis of the organism's extracellular glycoprotein matrix was initiated by extraction of SDS-mercaptoethanol soluble material, utilizing two different strength buffers. Hydroxyproline distribution and preliminary chemical analyses of the fractions obtained are presented in Table 7. Throughout the subsequent analyses of the fractions obtained from Volvox, the major portion of the

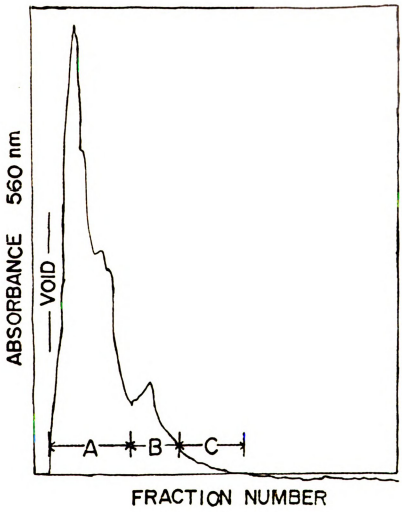


Figure 4. Deglycosylated Medium: Sephadex G-100 Elution Profile.

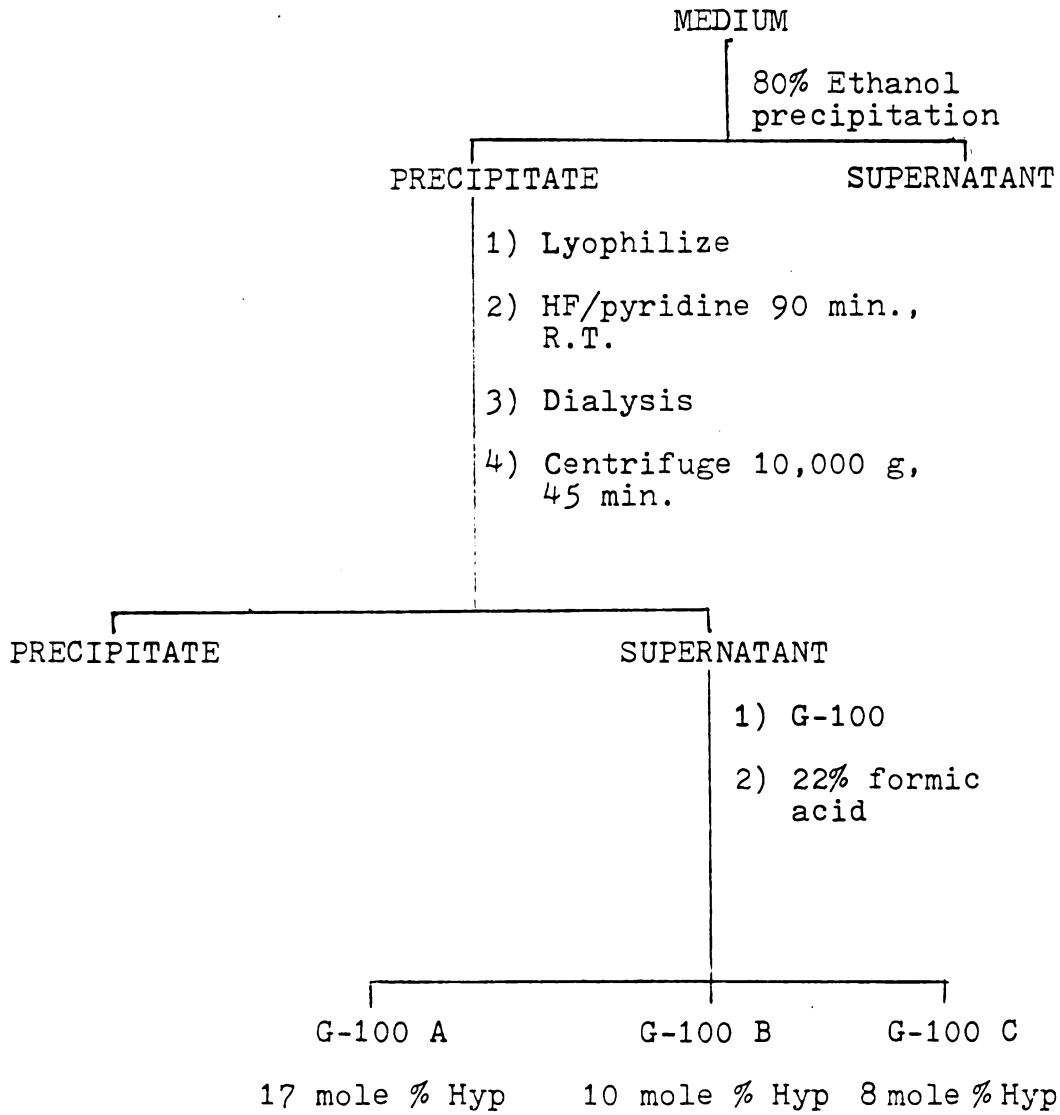


Figure 5. Flow Sheet for Partial Purification of Hydroxyproline Rich Proteins from the Medium.

Table 7. Volvox: Hydroxyproline Distribution and Preliminary Chemical Analyses

1% SDS 2.5% Mercaptoethanol Tris-HCl pH 6.8 Buffer

Fraction	% Total Hyp	Preliminary Chemical Analyses
Ghost Matrices	50-60%	
insoluble matrix	(35-40%)	13% protein by dry weight, 40 mole % Hyp ≙ 33% carbohydrate by dry weight
soluble matrix	(15-20%)	7% protein by dry weight, 25 mole % Hyp. ≙ 20% carbohydrate by dry weight
Ghost Filtrate	40-50%	60% protein by dry weight, 3 mole % Hyp

2% SDS 5% Mercaptoethanol Tris-HCl pH 6.8 Buffer

Fraction	% Total Hyp
Ghost Matrices	50-60%
insoluble matrix	(15-20%)
soluble matrix	(35-40%)
Ghost Filtrate	40-50%

hydroxyproline consistently appeared in the Sephadex G-100 void fraction. This may indicate an enormous amount and widespread location of the hydroxyproline rich glycoprotein throughout the structural matrices of Volvox, or it may represent incomplete separation of the fractions in the initial preparation and separation steps.

Membrane proteins can be solublized and extracted with detergents such as sodium dodecylsulfate. It was likely, therefore, that hydroxyproline rich components of the structural matrices of Volvox would also be solublized by detergents (approximately 50% of the hydroxyproline in Volvox is found in the ghost filtrate. The 2% SDS 5% mercaptoethanol ghost filtrate was examined for soluble, low molecular weight hydroxyproline rich peptides. Ethanol precipitated protein from the ghost filtrate was taken up in 10% trichloroacetic acid and stirred continuously overnight at room temperature in an attempt to solublize it. Highly glycosylated glycoproteins tend to be soluble in 10% TCA; most proteins are not. The 10% TCA supernatant contained approximately 10% protein by dry weight and 25 mole per cent hydroxyproline but accounted for only 3% of the total hydroxyproline in the ghost filtrate (see Figure 6 and Table 8). Although this fraction may have been reasonably pure, there was insufficient material for further analyses. It was reasonable to believe that this hydroxyproline rich fraction originated from the matrix and was a contaminant from the ghosts. The 10% TCA pellet, containing

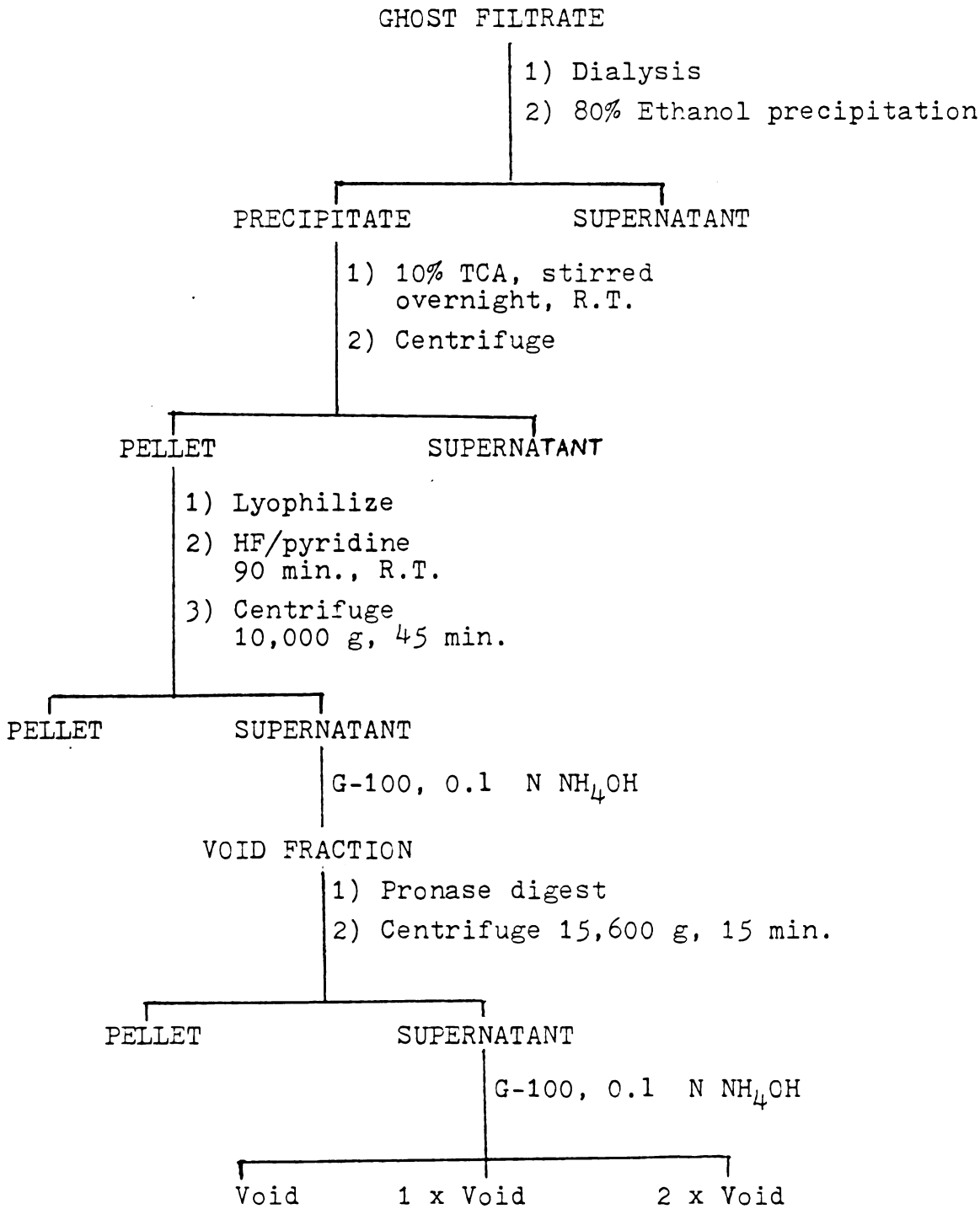


Figure 6. Flow Sheet for Partial Purification of the Ghost Filtrate

Table 8. TCA Soluble, Ethanol Precipitated Ghost Filtrate: Amino Acid Analyses

Amino Acid	Mole %
Asp	6.4
Thr	3.1
Ser	11.1
Glu	8.8
Gly	12.9
Ala	5.6
Val	1.8
Ile	1.1
Leu	2.7
Phe	3.5
Lys	3.8
Arg	9.1
Pro	4.7
Hyp	25.2

the major portion of the total hydroxyproline from the ghost filtrate, was deglycosylated in HF/pyridine and centrifuged. The supernatant was applied to the Sephadex G-100 column. The hydroxyproline rich fraction voided with an apparent molecular weight greater than 100,000. This fraction was digested with pronase and the supernatant reapplied to the G-100 column. Three hydroxyproline containing fractions were obtained, approximately 50% of the hydroxyproline voided the column, 25% was held up at 2 x void and 25% came out in the salt region. No amino acid analyses of these fractions were performed. The sum of the amounts of hydroxyproline in the three fractions was equal to about 10% of the total starting hydroxyproline with the major loss occurring as a pronase resistant pellet.

Hydroxyproline Enrichment of Deglycosylated Matrix

The HF/pyridine deglycosylated matrix was insoluble in formic acid, DMSO and chaotropic salts (similar insolubility properties are possessed by the inner cell wall of Chlamydamonas) (6). The deglycosylated matrix was digested with trypsin, which cleaves specifically at lysyl and arginyl residues, in an attempt to solubilize it. Although trypsin is specific for lysine and arginine, there may not be 100% cleavage at these residues if the lysine and arginine are adjacent to each other or next to proline, hydroxyproline or negatively charged amino acids, or if there is physical inaccessibility due to protective carbohydrates, aggregation

or insolubility. Approximately one-third of the hydroxyproline was solubilized by trypsin digestion. The digest was centrifuged and the supernatant was applied to the Sephadex G-100 column, 2-ml fractions were collected from which aliquots were taken for hydroxyproline estimation. The hydroxyproline rich fraction voided the column. Analysis of the column eluent for primary amine groups via reaction with fluorescamine (Figure 7) suggests two possible primary structures for the protein backbone. (1) The protein contains a hydroxyproline rich region which voided the G-100 column and a hydroxyproline free region which was held up in the salt region of the G-100 column after trypsin digestion. Or (2) the protein of interest, containing all the hydroxyproline, was resistant to trypsin attack and voided the G-100 column, while contaminating proteins were digested and held up in the salt region as tryptides. The minute peak, eluting just after the void peak on the G-100 column, (1.8 x void, Figure 7) containing 3% of the total hydroxyproline, suggests that the native protein may contain a hydroxyproline rich region (40-50 mole per cent hydroxyproline) and a hydroxyproline free region with a transitional region containing an intermediate amount of hydroxyproline (20 mole per cent) (Table 9). This trend of variance in the hydroxyproline content was seen in the other enzyme digested Volvox fractions, the autolytically digested medium and pronase digested ghost filtrate. The relatively nonspecific enzymatic cleavages gave rise to an assortment of polypeptides with a range of molecular weights and hydroxyproline content as expected. What is remarkable is the

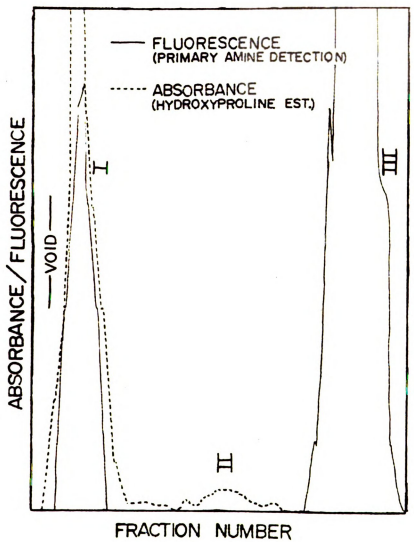


Figure 7. Trypsinized, Deglycosylated Matrix: Sephadex G-100 Elution Profile.

Table 9. Trypsinized, Deglycosylated Matrix, Sephadex G-100 Fractionation: Amino Acid Analyses

Method:	Technicon	GLC	GLC	GLC
Fraction	I	I	II	II
Amino Acid	Mole %	Mole %	Mole %	Mole %
Asp	6.2	5.5	10.8	10.5
Thr	5.1	4.5	5.5	6.8
Ser	7.2	6.4	14.2	9.1
Glu	4.6	3.4	6.2	4.2
Gly	7.4	6.0	11.0	10.7
Ala	6.7	4.7	9.5	13.7
Val	4.7	4.4	5.0	7.9
Ile	2.8	1.9	3.0	5.2
Leu	5.4	4.2	6.5	11.1
Phe	1.6	2.2	2.7	4.1
Lys	2.7	1.5	2.4	6.8
Arg	3.3	2.0	1.9	9.6
Tyr	0.7	1.0	n.d.	n.d.
Pro	2.1	3.6	n.d.	n.d.
Hyp	40.9	48.7	21.2	n.d.

n.d. not determined

large mole percentage of hydroxyproline in the major, high molecular weight fractions, i.e. the Sephodex G-100 void fractions.

The amino acid analysis of the trypsinized hydroxyproline rich G-100 void is given in Table 9. This fraction was digested with pronase and reapplied to the G-100 column. It voided the column, but the mole per cent of hydroxyproline had been increased (Table 10) to 80 to 90 mole per cent. If the pronase resistant peptide is 80 mole per cent hydroxyproline, it lends itself very nicely to the extensin model of Ser-Hyp_4 where Ser might be any other amino acid. If the peptide is 90 mole per cent hydroxyproline it would more likely be polyhydroxyproline with rare or no other amino acids. When analyzing amino acids at the 2 or 1 or lower mole percentage level the distinction between amino acid peak area and baseline noise becomes crucial and a major source of error. Further purification (to reduce baseline noise) and peptide sequencing may decide between an extensin periodicity (Ser-Hyp_4) and polyhydroxyproline with an extended or no periodicity.

One can repeat the argument presented in the case of trypsin digestion, that is, pronase, which cleaves at virtually all peptide bonds, except proline and hydroxyproline, either cleaved more hydroxyproline free regions from the hydroxyproline rich protein or digested contaminating proteins. With either possibility the mole per cent of hydroxyproline in the void fraction would be increased.

Table 10. Pronase Peptide from Sephadex C-100 Void of Trypsinized Deglycosylated Matrix: Amino Acid Analyses

Method	Techicon	Dionex
Amino Acid	Mole %	Mole %
Asp	0.8	2.0
Thr	0.6	1.2
Ser	2.8	3.6
Glu	0.8	1.0
Gly	1.0	1.7
Ala	0.9	1.7
Val	0.6	1.9
Ile	0.3	0.7
Leu	1.2	0.5
Tyr	0.2	1.1
Phe	0.3	0.7
Lys	0.6	0.6
Arg	0.7	0.7
Hyp	88.2	83.1
Pro	0.3	n.d.
Cys	0.5	n.d.

n.d. not determined

Because the hydroxyproline rich protein is in the void fraction, it is not necessarily a pure protein; nor can anything be learned about the primary structure of the protein by the effect of enzymes on its size (MW), i.e. were hydrozyproline free peptides cleaved from: (1) the protein of interest; decreasing its MW or (2) a contaminating protein; not affecting the MW of the hydroxyproline rich protein?

The fact that the pronase resistant deglycosylated fraction from the matrix voided a Sephadex G-100 column with an apparent MW greater than 100,000 raised the question of solubility. Was this peptide soluble in 0.1 N NH_4OH (used in column chromatography), was it being excluded from the G-100 gel pores by shape alone, was it aggregating, was it cross-linked to itself? Any combination of these factors may contribute to the picture of an enormous peptide containing a remarkable amount of hydroxyproline.

The pronase digested deglycosylated matrix was routinely centrifuged at 15,600 g in 2 ml centrifuge tubes for 15 minutes. The clear, pale yellow supernatant absorbed or scattered light at 325 nm. Neither proteins nor yellow colored solutions absorb at 325 nm. If the protein is in a particulate, insoluble form, light scattering will occur. A rigid rod shaped molecule (i.e. with a polyproline II helix) permeates dextran gels with an apparent molecular weight greater than its real molecular weight. Proteins

in solution will aggregate if it is thermodynamically favorable, by hydrogen bonding, hydrophobic interactions, etc. A small percentage of sugar and amino sugars are present (0.2 moles of sugar per mole of amino acid in the deglycosylated matrix) and may provide the means for cross-likage, as would cysteine, glutamic and aspartic acid, and arginine or lysine which are present in minute amounts (0 to 2 mole per cent).

Tri-Hydroxyproline

In order to obtain soluble, low molecular weight peptides, partial acid hydrolyses were performed. One hour hydrolysis of the deglycosylated matrix in 5.5N HCl at 110° C, gave rise to two approximately equal peaks on Sephadex G-50. (Figure 8). The 2.4 x void peak was lyophilized and taken up in 3 ml 0.01 N HCl and applied to a SP-Sephadex C-25 cation exchange column eluted with an ionic gradient. One major peak eluted at 0.26 M NaCl (Figure 9). GLC amino acid analysis showed the peak to be 88% hydroxyproline, 12% glutamic acid. Sequence data from this peak suggests the bulk of the hydroxyproline is tri-hydroxyproline with lesser amounts of tetra-, penta- and hexa-hydroxyproline with no other amino acids present (Figure 10). This and several subsequent fractions were analyzed before and after cyanoethylation in order to confirm the residue length data obtained by sequencing via automatic Edman degradation.

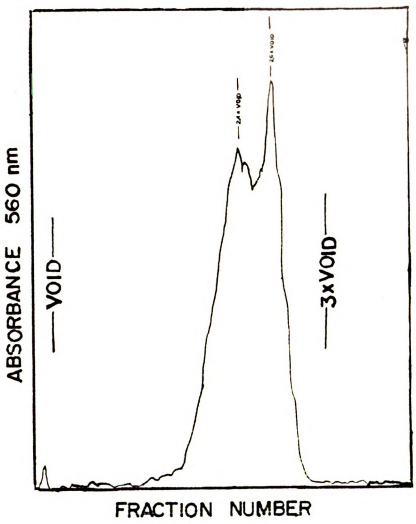


Figure 8. SG-50 Elution Profile of the Deglycosylated Matrix after a One Hour Partial Acid Hydrolysis.

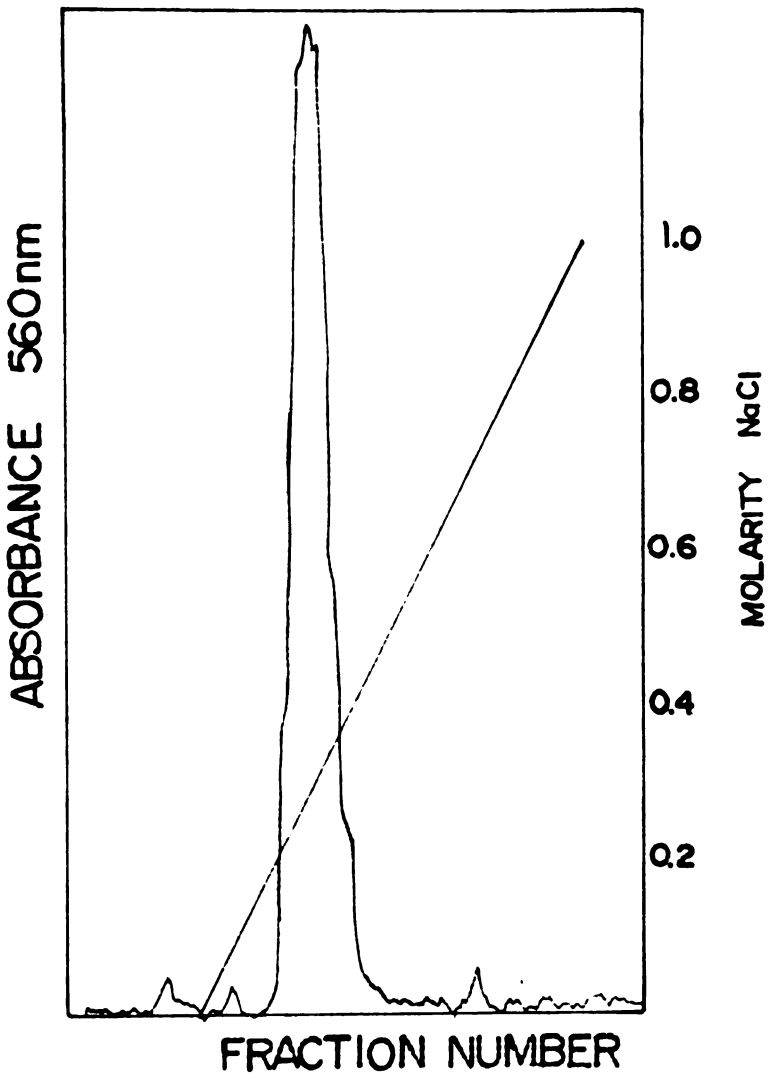


Figure 9. SP - Sephadex C-25 Elution Profile of the SG-50 2.4 x Void Peak.

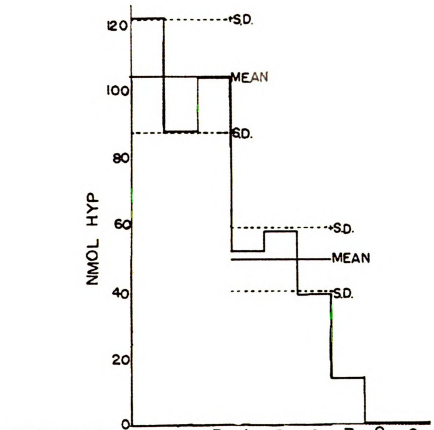


Figure 10. Sephadex G-50 2.4 x Void Peak: Sequence Profile (500 nmole polyhydroxyproline).

Cyanoethylation of amino groups by reaction with acrylonitrile at basic pH provides a means of subtractive amino terminus determination and thereby an estimate of the number of amino acid residues present in the polypeptide. Analyses were performed: (1) without hydrolysis to assay for free hydroxyproline and other free amino acids, (2) after an 18 hour hydrolysis and (3) after cyanoethylation followed by an 18 hour hydrolysis. A control experiment was performed with polyhydroxyproline, MW 11,000. Amino acid analysis of the polyhydroxyproline, MW 11,000, after cyanoethylation and acid hydrolysis gave a double peak at the elution time of hydroxyproline indicating racemization had occurred. If both isomers, or peaks, were quantitated the result was as expected, i.e. the loss of about one amino acid.

Cyanoethylation of the G-50 2.4 x void peak (obtained from one hour hydrolysis of deglycosylated matrix and further purified on SP-Sephadex C-25) indicated tri-hydroxyproline, which suggests that the occurrence of tetra-, penta- and hexa-hydroxyproline during sequencing was due to incomplete coupling or cleavage of the phenythiocarbamyl peptide to the phenylthiocarbamyl amino acid.

The G-50 2.6 x void was lyophilized and taken up in one ml of pH 2.5 pyridine-acetate buffer and applied to the Aminex A5 column. There were five major fluorescamine positive peaks, one of which was rich in hydroxyproline (A5-1, the void fraction). The other peak containing

hydroxyproline, A5-2 did not react with fluorescamine. The Aminex A5 elution profile is given in Figure 11, the amino acid analyses of peaks A5-1 through A5-4 are presented in Table 11. The amino terminus of A5-1 was cleaved by automatic Edman degradation, resulting in the loss of 8% of the hydroxyproline and all the glutamic acid suggesting the peptide to be polyhydroxyproline with a maximum of 11 to 12 residues of hydroxyproline with free amino acid contamination. Cyanoethylation of this fraction resulted in the loss of 58% of the hydroxyproline (Table 12) which suggests the peptide was dihydroxyproline contrary to the data obtained from the NH_2 terminal cleavage by automatic Edman degradation.

The A5-2 peak was determined to be free hydroxyproline by Edman degradation, i.e. virtually all the hydroxyproline was recovered in the first cleavage cycle. The amino acid analysis showed it be be 97% hydroxyproline and 3% glutamic acid (Table 11).

The sum of the hydroxyproline in the SP-Sephadex C-25 fraction (tri-Hyp), the Aminex A5-1 and A5-2 fractions account for 75% of the hydroxyproline present before the one hour hydrolysis, 40% of the total starting hydroxyproline in the ghosts and 20% of the total hydroxyproline present in the intact organism (Figure 12).

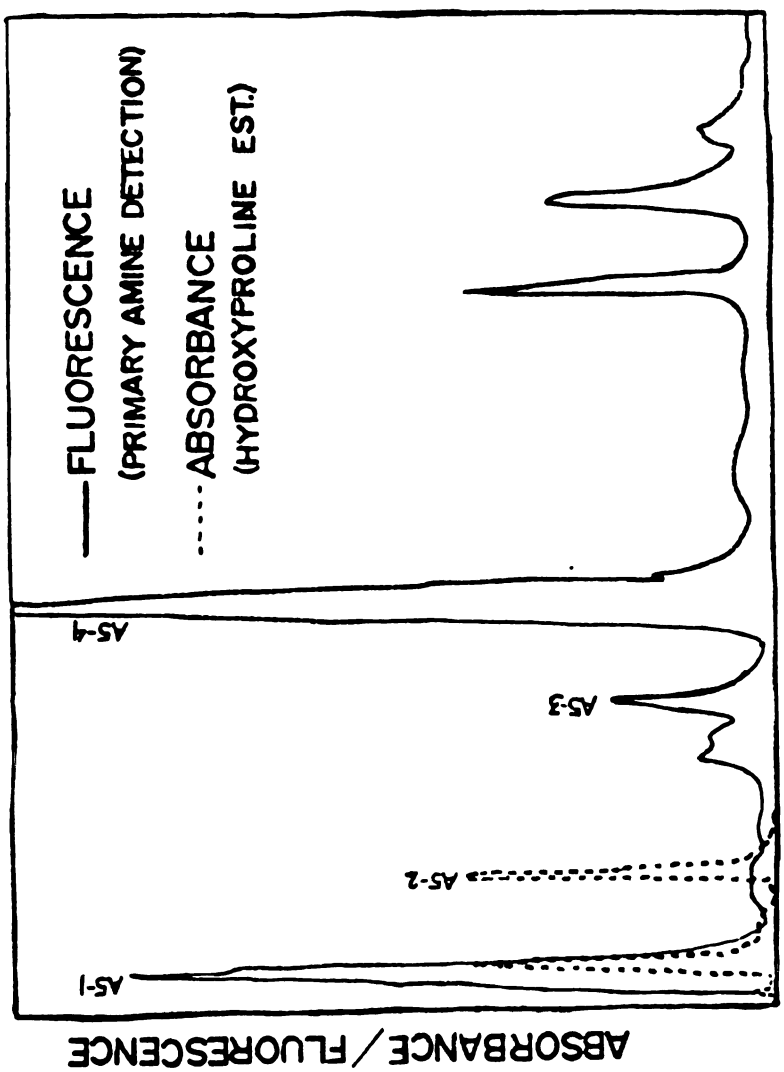


Figure 11. Aminex A5 Elution Profile of the Sephadex G-50 2.6 x Void Peak.

Table 11. Aminex A5 Fractionation: Amino Acid Analyses of Fractions A5-1 Through A5-4

Amino Acid	Mole Per Cent by GLC A5-1	A5-2	A5-3	A5-4
Asp	n.d.	n.d.	8.4	2.0
Thr	n.d.	n.d.	5.0	2.4
Ser	n.d.	n.d.	24.0	2.4
Glu	20.0	3.0	15.7	11.8
Gly	n.d.	n.d.	n.d.	25.5
Ala	n.d.	n.d.	n.d.	22.9
Val	n.d.	n.d.	n.d.	2.4
Cys	n.d.	n.d.	n.d.	n.d.
Met	n.d.	n.d.	n.d.	n.d.
Ile	n.d.	n.d.	n.d.	n.d.
Leu	n.d.	n.d.	n.d.	10.6
Tyr	n.d.	n.d.	n.d.	1.4
Phe	n.d.	n.d.	n.d.	3.9
Lys	n.d.	n.d.	n.d.	2.4
Arg	n.d.	n.d.	n.d.	0.8
Pro	n.d.	n.d.	6.2	2.3
Hyp	80.0	97.0	41.4	6.7
Suggested sequence:	di-Hyp [minimal peptide length] *	free Hyp	di-peptides and/or free amino acids	

n.d. not determined

* see Table 12 and text

Table 12. Cyanoethylation of Aminex A5-1 Fraction: Amino Acid Analyses

Conditions:	(-) Hydrolysis	(+) 18 Hour Hydrolysis		Cyanoethylation (+) 18 H Hydrolysis	
Amino Acids	nmole	nmole	mole %	nmole	mole %
Asp	n.d.	4.6	5.2	1.2	3.4
Thr	n.d.	2.8	3.2	0.5	1.4
Ser	n.d.	3.3	3.7	0.7	2.0
Glu	n.d.	2.7	3.0	2.3	6.5
Gly	n.d.	1.5	1.7	1.1	3.1
Ala	n.d.	1.0	1.1	0.8	2.2
Val	n.d.	0.4	0.5	0.2	0.5
Leu	n.d.	0.7	0.8	0.7	2.0
Phe	n.d.	1.5	1.7	0.6	1.7
Arg	n.d.	3.6	4.1	n.d.	n.d.
Hyp	n.d.	66.6	75.1	27.7	77.4

n.d. not determined

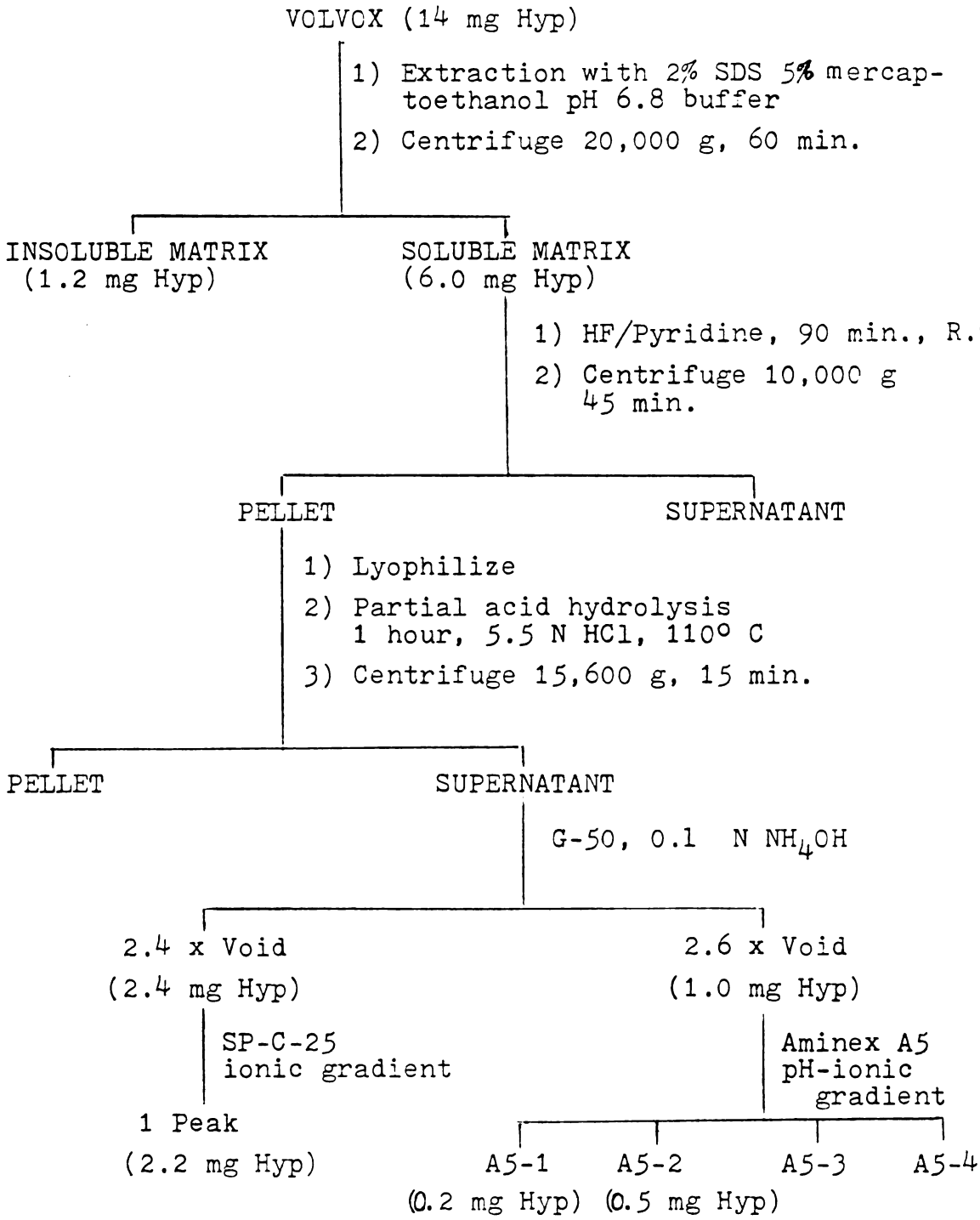


Figure 12. Hydroxyproline Flow Sheet.

Penta-Hydroxyproline

The "soluble" pronase resistant hydroxyproline rich core from deglycosylated matrix was partially hydrolyzed for one half hour in 5.5 N HCl at 110° C and sized on Sephadex G-25. The sample eluted at 1.6 x void, apparent MW 3500, as a broad peak, whereupon it was freeze dried, taken up in one ml pH 2.5 pyridine-acetate buffer and applied to an Aminex A5 column. One major hydroxyproline rich peptide was obtained in the void and two minor hydroxyproline rich peaks were slightly held up (Figure 13). Amino acid analyses by GLC are in Table 13. Sequence data suggest the void peak contains polyhydroxyproline peptides of which at least 70% are penta-hydroxyproline. The remaining peptides contain 6 to 12 hydroxyproline residues. Serine is present at the 11th and 12th residues in amounts equal to hydroxyproline (Figure 14). The peptides Hyp_{6-12} may be real or may be the result of incomplete coupling and cleavage or may be a combination of the two. Cyanoethylation of this fraction removes about 13% or 1/8 of the hydroxyproline which is consistent with the sequence data that suggests the fraction is composed of penta-hydroxyproline with some (approximately 30%) of the hydroxyproline accounted for as longer polyhydroxyproline peptides. The presence of serine in the 11th and 12th residues supports the model for the extracellular glycoprotein of Volvox as consisting of polyhydroxyproline or a polyhydroxyproline domain with periodic other amino acids. A flow diagram of the one-half

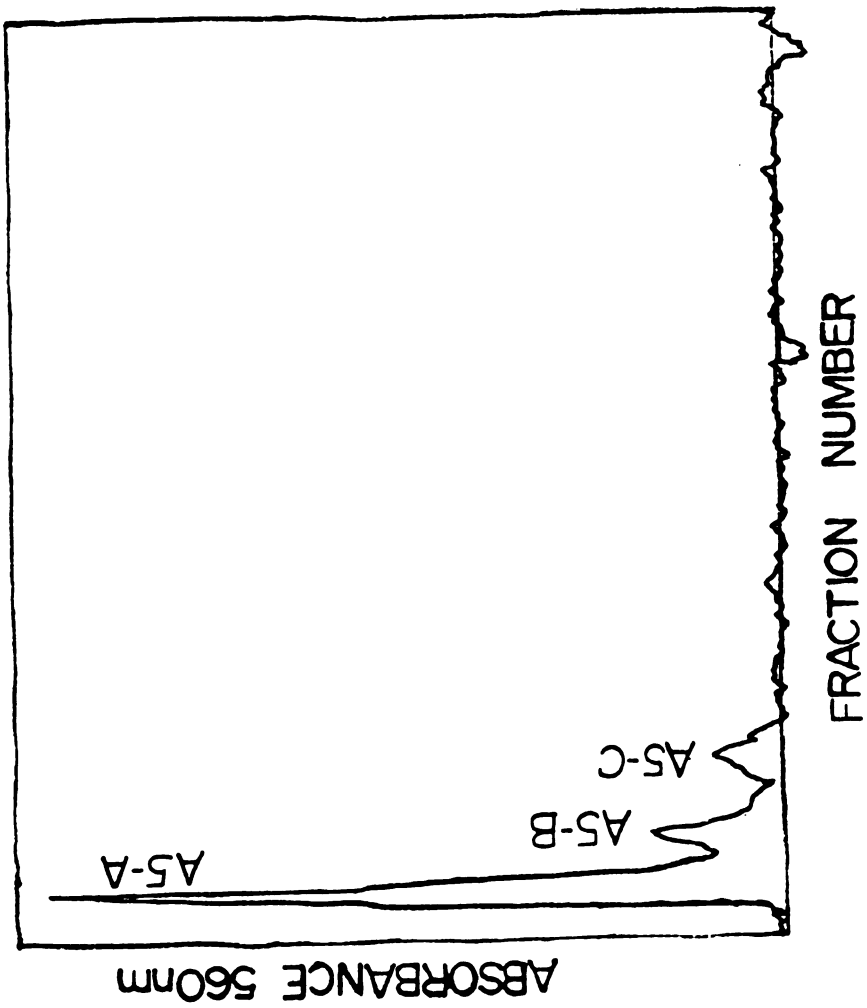


Figure 13. Aminex A5 Elution Profile of the Sephadex G-25 1.6 x Void Peak.

Table 13. Aminex A5, Fractions A5-A, A5-B and A5-C: Amino Acid Analyses

Amino Acid	Mole Per Cent by GLC A5-A	A5-B	A5-C
Ala	1.8	n.d.	n.d.
Gly	7.3	0.8	n.d.
Val	2.2	1.0	0.4
Thr	2.6	1.5	1.9
Ser	3.9	6.8	3.2
Glu	2.3	1.4	0.9
Ile	1.2	0.7	n.d.
Leu	2.0	1.1	n.d.
Pro	2.6	n.d.	n.d.
Hyp	64.4	81.6	88.0
Met	0.5	n.d.	n.d.
Phe	2.9	2.3	0.9
Asp	5.0	0.5	2.0
Lys/Tyr	5.3	0.5	0.4
Arg	1.9	n.d.	0.5

n.d. Not determined

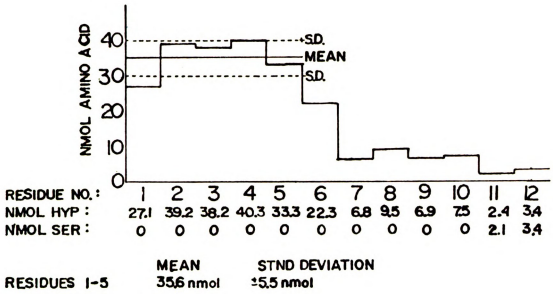


Figure 14. Aminex A5-A Sequence Profile (250 nmole polyhydroxyproline).

hour partial acid hydrolysis of the deglycosylated matrix is given in Figure 15.

High Molecular Weight Polyhydroxyproline

The pronase resistant, hydroxyproline rich core from the deglycosylated matrix was further "purified" on SP-Sephadex C-25 and sequenced in this crude form (Figure 16). The amino acid analysis, before and after passage through an SP-C-25 column is given in Table 14. The protein was not held up on the SP-C-25 column, being either too large or uncharged at pH 2, but there was hydroxyproline enrichment. Sequence data from this crude sample, (SP-C-25 "purified", pronase resistant, hydroxyproline rich peptides from the deglycosylated matrix) gave a maximum peptide length of 1000 residues of hydroxyproline, corresponding to a maximum molecular weight of 131,000. There was a background of other amino acids present in amounts too low to quantitate, this is consistent with the amino acid analysis of the sample (10 mole per cent other amino acids). So the question remains are the other amino acids constituents of the peptide or contaminants? The pentapeptide, obtained after one-half hour hydrolysis, indicates this protein does not exist in the form of extensin (Ser-Hyp_4), but the recurrent presence of amino acids in amounts about 1/10 that of hydroxyproline, suggests that an extended purehydroxyproline periodicity may be present.

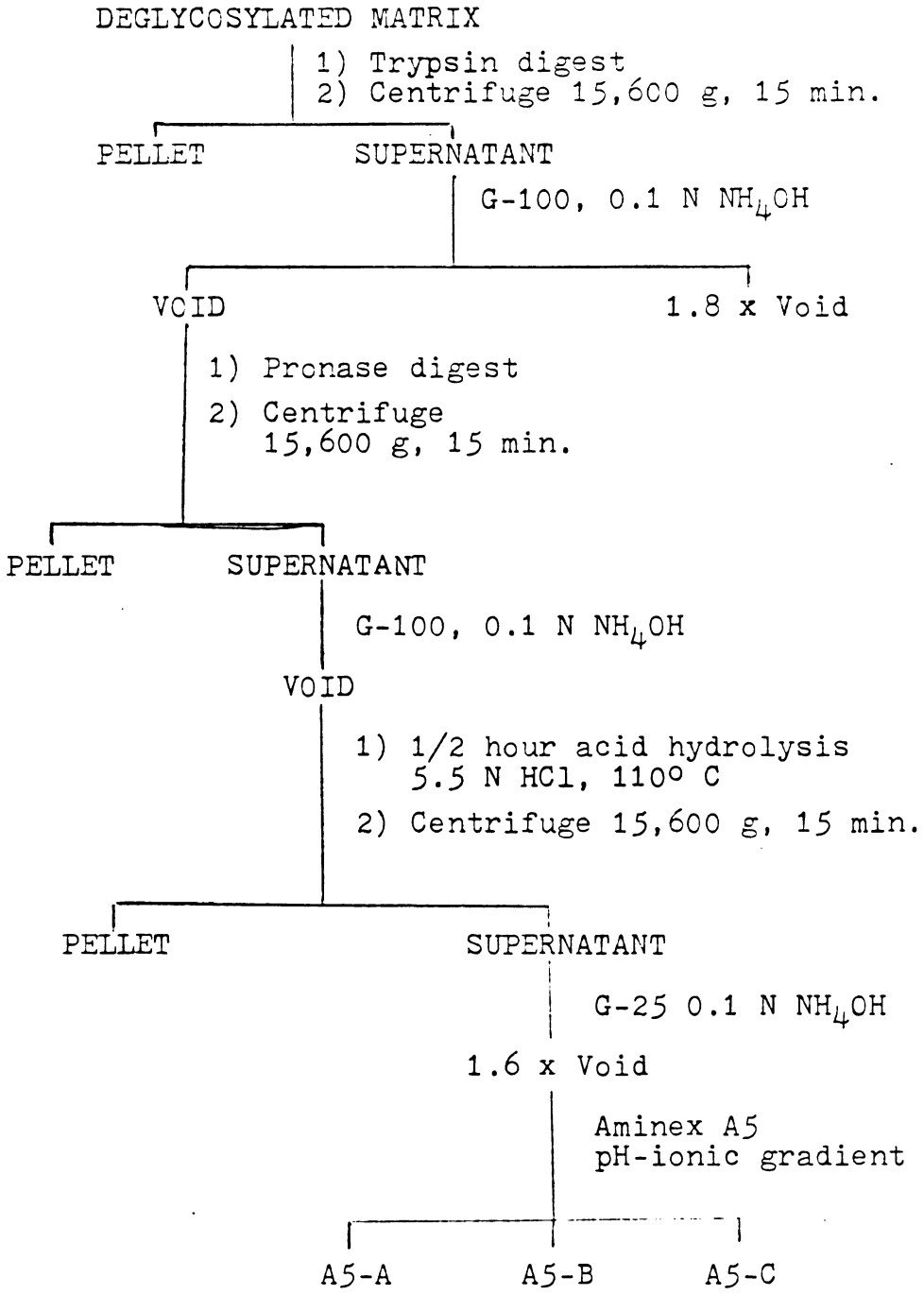


Figure 15. Flow diagram for 1/2 Hour Partial Acid Hydrolysis of the Deglycosylated Matrix.

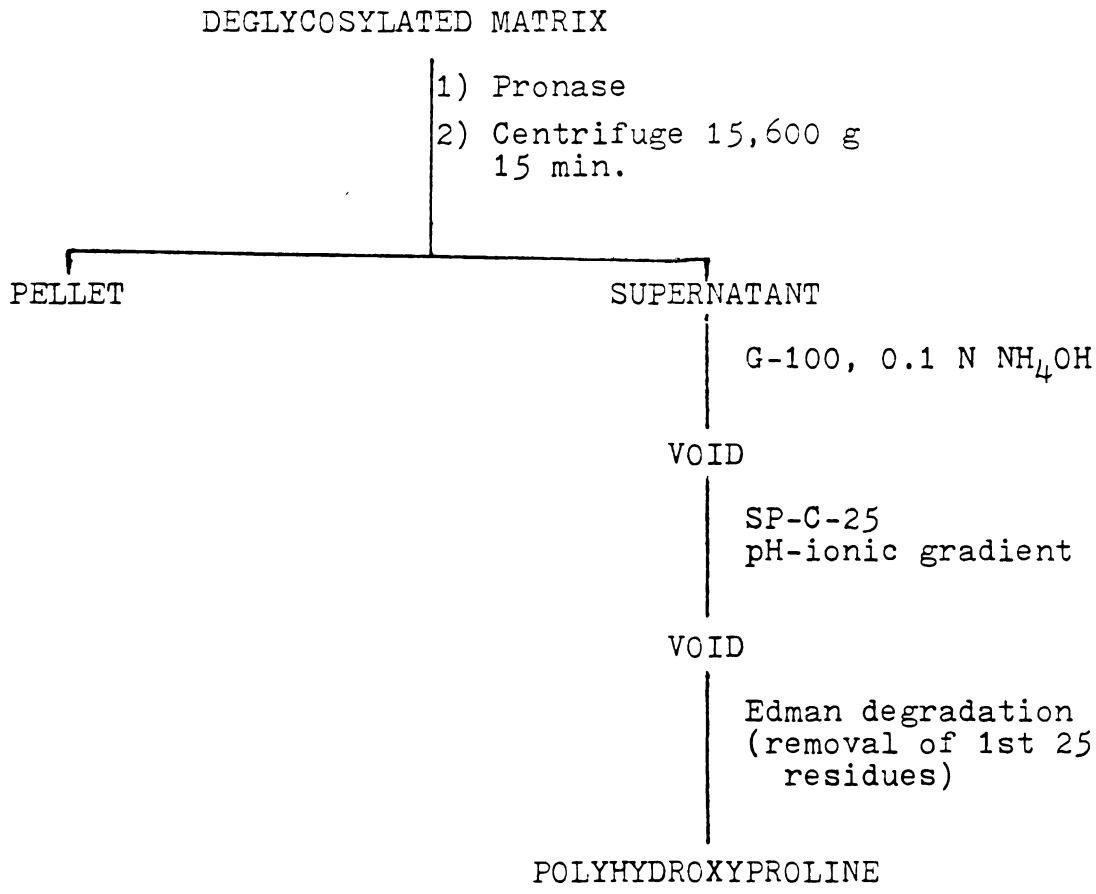


Figure 16. Flow Sheet for "Purification" of Pronase Resistant Polyhydroxyproline from the Deglycosylated Matrix

Table 14. Polyhydroxyproline: Amino Acid Analysis Before and After SP-Sephadex C-25 Purification

Amino Acid	Before	After
Asp	2.0	1.3
Thr	1.2	0.6
Ser	3.6	3.0
Glu	1.0	0.5
Gly	1.7	0.8
Ala	1.7	0.9
Val	1.9	0.4
Ile	0.7	0.2
Leu	0.5	0.7
Tyr	1.1	0.5
Phe	0.7	0.5
Lys	0.6	0.3
Arg	0.7	n.d.
Hyp	83.0	90.5

n.d. Not determined

Sugar Analyses of the Matrix

Often it is necessary to remove the protective carbohydrates from glycoproteins in order to expose the protein backbone to digestive enzymes. HF, a well known deprotecting agent in protein synthesis, is extraordinarily effective in depolymerizing and cleaving carbohydrate from protein. This lab has developed an HF solvolysis technique to deglycosylate glycoproteins and to remove cellulose from plant cell walls (31). Seventy per cent HF in pyridine was also effective in removing carbohydrate from glycoprotein (29). It is a safer reagent to work with than is HF and does not require extensive reaction apparatus. HF/pyridine does not rapidly depolymerize cellulose to a significant extent (Table 15). Sugar analyses of the 1% SDS 2.5% mercaptoethanol extracted structural matrices from Volvox carteri were performed by GLC of their alditol acetates and TMS derivitives. The GLC analyses (Table 16a & b) show the major sugars to be glucose, galactose, mannose and arabinose. Fifty percent of the analyzed carbohydrate is an acid resistant glucan. Four to seven μg of glucose was hydrolyzed per 100 μg of matrix by conventional means (90 minute methanolysis in 1.5 N methanolic HCl or 60 minute trifluoroacetic acid hydrolysis). A three-fold increase in analyzed glucose occurred when solvation of the carbohydrate in anhydrous HF preceded the acid hydrolysis and derivation. Although anhydrous HF depolymerizes cellulose, HF/pyridine does so only very slowly. The fact that very little

Table 15. HF/Pyridine and Cellulose

Form of Cellulose	Initial Weight	Final Weight	% Remaining After HF/Pyridine 90 min. at Room Temp.
Cotton	244.5 mg	233.5 mg	95.5%
Whatman Filter Paper	328.1 mg	308.4 mg	94.0%

Table 16a. Carbohydrate Analyses of Glycosylated Matrix

Treatment II - Hydrolysis conditions:			60 min. Trifluoroacetic acid				90 min. MeOH-HCl	
GLC column:			[a]PEGS		OV-275		SP-2100	
Sample	Treatment I	Sugar	µg	Mole %	µg	Mole %	µg	Mole %
100 µg matrix (100 nmole amino acid)		Arabinose	3.2	19.9			2.4	14.8
		Mannose	3.0	18.6			0.9	5.5
		Galactose	4.0	24.8			4.3	26.5
		Glucose	5.9	36.6			8.6	53.2
		[b]% Carbohydrate:	16.1				16.2	
100 µg matrix (100 nmol amino acid)	HF 0° C 1 hour	Arabinose	4.3	13.5	6.1	17.1	6.0	14.4
		Mannose	1.8	5.7	0.8	2.1	1.1	2.6
		Galactose	6.5	20.4	5.6	15.7	6.2	14.6
		Glucose	19.2	60.4	23.0	65.0	28.4	68.2
		[b]% Carbohydrate:	31.8		35.5		41.7	
100 µg matrix (100 nmole amino acid)	HF R.T. 3 hours	Arabinose	4.9	10.2	9.8	16.3	8.8	18.3
		Mannose	9.6	20.1	4.8	8.0	4.4	9.1
		Galactose	3.8	8.0	17.3	28.7	4.5	9.3
		Glucose	29.4	61.6	28.2	46.9	30.5	63.3
2.8 mole sugar / mole amino acid		Amino sugar	-		+		+	
		[b]% Carbohydrate:	47.7		60.1		48.2	

[a] 0.2% poly(ethylene glycol adipate) 0.2% poly(ethylene glycol succinate) and 0.4% XF 1150

[b] Per cent carbohydrate by dry weight

Table 16b. Carbohydrate Analyses of HF/Pyridine Deglycosylated Matrix

Treatment II - Hydrolysis conditions:			60 min. Trifluoroacetic acid				90 min. MeOH-HCl	
GLC column:			[a]PEGS		OV-275		SP-2100	
Sample	Treatment I	Sugar	µg	Mole %	µg	Mole %	µg	Mole %
100 µg HF/pyridine deglycosylated matrix (350 nmole amino acid)	HF 0° C 1 hour	Arabinose	0.7	4.2	0.4	3.4	0.7	5.4
		Mannose	4.6	27.4	5.7	47.9	6.9	53.1
		Galactose	1.7	10.1	4.2	35.3	1.7	13.0
		Glucose	9.8	58.3	1.6	13.4	3.7	28.5
		[b]% Carbohydrate:	16.8		11.9		13.0	
100 µg pronase resistant deglycosylated matrix 0.2 mole sugar mole amino acid	HF R.T. 3 hours	Arabinose			0.8	6.7		
		Mannose			4.2	35.0		
		Galactose			4.2	35.0		
		Glucose			2.8	23.3		
		Amino sugar			+			
		[b]% Carbohydrate			12.0			

[a] 0.2% poly(ethylene glycol adipate) 0.2% poly(ethylene glycol succinate) and 0.4% XF 1150

[b] Per cent carbohydrate by dry weight

additional glucose was released when the matrix, deglycosylated in HF/pyridine, was treated with anhydrous HF indicates the resistant glucan is not cellulose.

Amino sugars appear to be present at low levels, but identification and quantification by GLC and HPLC automated amino acid analyzer after acid hydrolysis was not definitive.

Treatment with HF/pyridine removes 75-80% of the initial carbohydrate, as measured by HF solvolysis (Table 16 a & b). The remaining sugar may be a result of the inaccessibility of the HF/pyridine to the insoluble portions of matrix that persisted even with continuous stirring. The matrix glycoprotein contains approximately 2.8 moles of sugar per mole amino acid. After HF/pyridine treatment 0.2 moles of sugar remain per mole of amino acid.

Molecular Weight Determination and Sequencing

Size determination by gel permeation chromatography on dextran gels gave relative or maximum values rather than absolute values. A standard of free hydroxyproline was held up at 1.9 x void on Sephadex G-25 and 2.4 x void on Sephadex G-50. One would expect free hydroxyproline to totally permeate the column based on its molecular weight alone. If the SDS-mercaptoethanol extracted, deglycosylated protein from the structural matrices of Volvox had the tertiary structure of a rigid rod, as does polyhydroxyproline, its molecular weight on Sephadex G-100 would be anomalously high because a rod has a much larger effective

hydrodynamic volume than a globular molecule of similar molecular weight. A standard of polyhydroxyproline, MW 18,000, eluted the G-100 column at 2.0 x void (an apparent MW twice that of a globular protein of the same MW.) The pronase resistant hydroxyproline rich core, from deglycosylated extracellular matrix, which voided the G-100 column must be greater than 18,000 MW, cross-linked to itself, aggregating in solution or insoluble. UV-visible spectra of the G-100 void fraction of SDS-mercaptoethanol extracted deglycosylated matrix gave a typical light scatter curve, indicating insolubility, aggregation or enormous molecules.

Approximately one out of every one hundred residues of the pronase resistant hydroxyproline rich core from the deglycosylated matrix was cysteine. If the rod shaped polypeptide was cross-linked at every hundredth residue or so by cystine, the tertiary structure would become a rigid fence, totally excluded from the gel. The performic acid oxidation of the pronase resistant hydroxyproline rich deglycosylated matrix gave 100% conversion of cysteine to cysteic acid. When the performic oxidized fraction was rechromatographed on Sepharose CL-4B column in 0.1 N NH_4OH the major hydroxyproline rich fraction eluted as a very broad peak at 1.9 x void, an apparent MW of 160,000 with a range of 50,000 to 500,000.

SDS binds to protein in a ratio of 1.4 g/g of protein (11). SDS dissociates oligomeric proteins and totally

swamps out the protein's own charge so there is a constant charge per unit surface area of protein. SDS binding also alters the protein's shape, causing a rigid rod like conformation with a width that is constant and a length that is a function of molecular weight. Because SDS is not completely removed by the dialysis procedure used during preparation of the extracellular matrix, there remained a significant possibility that the bound SDS after dialysis altered the protein's conformation so that it was a rigid rod by experimental and not natural design. A control experiment ruled out the apparent high molecular weight as a function of SDS. For this experiment one harvest of Volvox was extracted in fresh Volvox medium. The spheroids were broken open by douncing once with the tight fitting dounce, then centrifuged at 20,000 g for 45 minutes. The clear, viscous supernatant was freeze dried, deglycosylated in HF/pyridine and digested with pronase. The pronase digest was centrifuged, to remove insoluble material, and applied to the G-100 column. The water-extracted, deglycosylated pronase resistant fraction voided the G-100 column. It was applied to and voided the Sepharose CL-4B column. The SDS extracted, pronase resistant deglycosylated matrix, which was oxidized with performic acid, was held up at 1.9 x void on the Cl-4B column. This suggests that disulfide bridges, not reduced during the initial extraction (with mercaptoethanol), may be the cross-linkages that resulted

in the water-extracted fraction voiding the CL-4B column (apparent MW 5 x 10^6).

The pronase resistant, Hyp rich core from deglycosylated matrix was applied to Whatman No. 4 filter paper in a 1 cm streak (50 μg) and was electrophoresed in pH 1.9 formic-acetic acid buffer for three hours at 5 K volts. The electrophoresis paper was dried and dipped in a ninhydrin/acetone/pyridine/acetic acid mixture 0.6/200/2/2 (w/v/v/v), dried, and developed at 60° C for 20 minutes. One spot appeared, at the origin, stained yellow. This indicates a very high molecular weight peptide. The characteristic yellow color when ninhydrin reacts with a secondary amine indicates a proline or hydroxyproline amino terminus with rare or no arginine or lysine residues.

Sequencing of the first 25 amino acid residues of the SDS-mercaptoethanol extracted, deglycosylated pronase resistant fraction gave 8.0$\pm$2.4 nanomoles of hydroxyproline per cleavage cycle. There was a background (not easily distinguished from baseline) of various amino acids in amounts too small to quantitate. Eight thousand and five hundred nanomoles of Hyp were placed into the sequencer, 176 nanomoles were recovered as PTH derivatives during the 25 cleavage cycles. The peptide minus the first 25 residues, was recovered from the sequencer, acid hydrolyzed and its amino acid composition determined. The result was 6200 nanomoles of hydroxyproline recovered with a very low

percentage of other amino acids (less than 10%) (Table 13). The estimated peptide length was not less than 750 and not more than 1000 hydroxyproline residues.

DISCUSSION

The major glycoprotein in the extracellular matrix of Volvox is composed of polyhydroxyproline. This glycoprotein is at least 50% carbohydrate, the major sugars are glucose, arabinose, galactose and mannose. Fifty per cent of the analyzed carbohydrate from this extracellular glycoprotein (by anhydrous HF deglycosylation) is a trifluoroacetic acid resistant glucan but is not cellulose. (The matrix appears to be rich in carbohydrate, noncovalently bound to this glycoprotein.) Glycopeptide linkages known to occur in plants and protists include: β-D-GlcNAc-Asn, β-D-Gal-Ser, Xyl-Thr, Gal-Hyp, β-L-Ara-Hyp and Glc-Hyp (24). The ratio of carbohydrate to amino acid in the SDS-mercaptoethanol extracted matrix is 2.8 : 1.0. Assuming a combination of all these N- and O-glycosidic linkages to be present in the extracellular matrix of Volvox, the ratio of carbohydrate to asparagine, serine, threonine and hydroxyproline is 5.5 : 1.0. The polysaccharide sidechain may fold back on the protein backbone and hydrogen bond with it (23,24), adding tensile strength and protecting it from proteolytic attack.

The presence of cysteine, lysine, arginine, glutamic acid and aspartic acid in the extracellular matrix provides

sites for peptide cross-linkage in the intact organism, although the thick viscous nature of the extracellular material is more consistent with a highly disordered (i.e. uncross-linked) matrix of long filamentous, hydrated glycoprotein. The insolubility of the matrix during isolation and purification may be more a function of denaturation and aggregation than of cross-linkage.

The extracellular matrix glycoprotein may exist as a one, two or three domain protein; a polyhydroxyproline region plus or minus a hydroxyproline free region and a transitional hydroxyproline rich region. The polyhydroxyproline domain may consist of pure hydroxyproline (750 to 1000 residues of hydroxyproline) and/or polyhydroxyproline with periodic other amino acids, most frequently serine. The only periodicity detected is a $Hyp_{10}Ser$ peptide after partial acid hydrolysis. This periodicity could not be confirmed because of the impure nature of the unhydrolyzed polypeptide. The hydroxyproline rich transitional region is hypothetical but may exist as a region that is 20 mole per cent hydroxyproline, 14 mole per cent serine, 11 mole per cent glycine and 11 mole per cent aspartic acid (Table 9). The hydroxyproline free domain, also hypothetical, is rich in alanine, glycine, aspartic acid and serine (Table 9). The hydroxyproline rich transitional region and hydroxyproline free region, which were obtained by trypsin digestion of the deglycosylated extracellular matrix, will remain

hypothetical until it is proven that these peptides originated from the protein containing the polyhydroxyproline domain. Knowledge of the entire structural picture of the protein backbone awaits a method for purifying the insoluble, deglycosylated matrix glycoprotein. The nature of the protein's insolubility is uncertain.

The polyhydroxyproline protein backbone of the extracellular matrix glycoprotein has the same tertiary structure as extensin fragments and the polypeptide chains of collagen. Poly-L-hydroxyproline exists in a poly-L-proline II-type helix as do the individual chains of collagen, but it does not form the super helix seen in collagen (38). The most primitive organism from which collagen has been purified is the sea urchin; the collagen isolated is a single, not triple chain protein (32). Another parallel between the extracellular glycoprotein in Volvox and collagen of lower forms of animal life, sponges and sea urchins, is the high degree of glycosylation.

Extending the hypothesis put forth by Lamport, the archetypal protein may have had a glycine rich collagen-like region and a polyhydroxyproline region. The collagen-like region was selected for in the animal kingdom. Perhaps the polyhydroxyproline region was selected for in the plant kingdom. A one base change in every fifth codon of the genome would allow conversion of polyhydroxyproline into the repeating pentapeptide periodicity of extensin (Ser-Hyp_4).

The major extracellular glycoprotein from Volvox is not the archetypal glycoprotein from which collagen and extensin arose. Perhaps a minor component of the extracellular matrix will more closely approximate the archetypal protein. (Glycine rich fractions have been isolated from various Volvox preparations, Table 8, 9, 11).

Do the structural glycoproteins containing hydroxyproline as a major constituent belong to one "Super Family" of proteins, that is, did they arise from one archetypal protein, if so, what are their relationships? Continuing to follow the phylogenetic trail may provide the answer.

REFERENCES

1. Akiyama, Y. M. S. Thesis. Michigan State University, E. Lansing, MI (1979).

2. Albersheim, P., D. J. Nevin, P. D. English and A. Karr. Carbohyd. Res. 5 340 (1967).

3. Bhatti, T., R. E. Chambers and J. R. Clamp. Biochim, Biophys, Acta 222 339 (1970).

4. Bisaulptra, T. and J. R. Stein. Can. J. Bot. 44 (12) 1697 (1966).

5. Bornstein, P. Ann Rev. Bioch. 43 567 (1974).

6. Catt, J. W., G. J. Hills and K. Roberts. Planta 138 91 (1978).

7. Darden, W. H. Jr. J Protozool 13 (2) 239 (1966).

8. Davidson E., S. Fowler and B. Poole. Anal Bioch 69 49 (1975).

9. Fletcher, J. C. Biochem. J. 98 34c (1966).

10. Gotelli, I. B. and R. Cleland. Amer. J. Bot. 55 (8) 907 (1968).

11. Haschemeyer, R. H. and A. E. V. Haschemeyer. "Proteins: a guide to study by physical and chemical methods" New York: J. Wiley (1973).

12. Homer, R. B. and K. Roberts. Planta 146 217 (1979).

13. Jermyn, M. A. and Y. M. Yeow. Aust. J. Plant Physiol. 2 501 (1975).

14. Kirk, D. L. and M. M. Kirk. Dev. Biol. 50 413 (1976).

15. Kochert, G. J. Protozool. 15 (3) 438 (1968).

16. Kozulic, B., B. Ries and P. Mildner. Anal. Biochem. 94 36 (1979).

17. Kudo, R. R. "Protozoology" Thomas. 5th ed. 3rd. printing (1977).

18. Kurn, N. and B. Sela. FEBS LETT 104 (2) 249 (1979).

19. Lang, N. J. Amer. J. Bot. 50 (3) 280 (1963).

20. Lamport, D. T. A. Biochem 8 1155 (1969).

21. Lamport, D. T. A. and D. H. Miller. Plant Physiol. 48 454 (1977).

22. Lamport, D. T. A., L. Katona and S. Roerig. Bioch. J. 133 125 (1973).

23. Lamport D. T. A. "Recent Advances in Phytochemistry" (eds. F. A. Loewus and V. C. Runeckles), New York: Plenum Press (1977), Vol. II, p. 79.

24. Lamport, D. T. A. "The Biochemistry of Plants" (ed. Stumpf), Academic Press, Inc. (1980), Vol. 3, p. 501.

25. Lamport D. T. A. and G. Kochert. Unpulbished results. (1974).

26. MacKenzie, S. L. and Tenaschuk, D. J. J. Chromatography 171 195 (1979).

27. MacKenzie, S. L. and Tenaschuk, D. J. J. Chromatography 173 53 (1979).

28. Margulis, L. "Handbook of Genetics" (ed. R. C. King) Plenum Press (1979) p. 1.

29. Mani, U. V., Y. Akiyama, S. Mohrlok and D. T. A. Lamport. Federation Proceedings Abstract 418 (1979).

30. McCracken, M. D. and W. J. Barcellona, J. Histochem. and Cytochem. 24 (5) 668 (1976).

31. Mort, A. J. and D. T. A. Lamport. Anal. Bioch. 82 289 (1977).

32. Nordig, A., E. Rogall and U. Hayduk. "Chemistry and Molecular Biology of the Intercellular Matrix". (ed. E. A. Balazs), Boston: Academic Press (1970), Vol. I, p. 27.

33. Pickett-Heaps, J. D. Planta (Berl) 90 174 (1970).

34. Prescott, G. W. "How to Know Fresh Water Algae". Dubuque, Iowa: Wm. C. Brown Co. (1954).

35. Preston R. D. "The Physical Biology of Plant Cell Walls". (eds. Chapman and Hall), London: Halsted Press (1974) p. 32.

36. Prockop, D. J., K. I. Kivirikko, L. Tuderman and N. A. Guzman. N. Engl. J. Med. 301 (1) 13 (1979).

37. Punnet, T. and E. C. Derrenbacker. J. Gen Microbiol. 44 105 (1966).

38. Ramachandran, G. N. "Treatise on Collagen " (ed. G. N. Ramachandran), New York: Academic Press (1967) Vol. I, p. 103.

39. Sessoms, A. H. and R. J. Huskey. Proc. Nat. Acad. Sci. USA 70 (5) 1335 (1973).

40. Sessoms, A. H. Ph. D. Dissertation. University of Virginia, Charlottesville, Va. (1974).

41. Starr, R. C. Devel. Biol. Supp. 4 59 (1970).

42. Thompson, E. W. and R. D. Preston. Nature 213 684 (1967).

43. Viamontes, G. I. and D. L. Kirk. J. Cell Biol. 75 719 (1977).

44. Viamontes, G. I., L. J. Fochman and D. L. Kirk. Cell 17 537 (1979).

45. Yates, I., M. Darley and G. Kochert. Cytobios 12 211 (1975).

46. Whittaker, R. H. Science. 63 (3863) 150 (1969).