hormonal proteins and peptides

EDITED BY **CHOH HAO LI**

ACADEMIC PRESS INC.
A Subsidiary of Harcourt Brace Jovanovich, Publishers

volume 1

HORMONAL PROTEINS AND PEPTIDES

Edited by CHOH HAO LI

The Hormone Research Laboratory
University of California
San Francisco, California

VOLUME I

ACADEMIC PRESS New York and London 1973
A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 72-88367

PRINTED IN THE UNITED STATES OF AMERICA

List of Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

OM P. BAHL (171), Department of Biochemistry, State University of New York at Buffalo, Buffalo, New York

ROBERT B. CARLSEN (17), Department of Biological Chemistry, University of California, School of Medicine, Los Angeles, California

H. EDELHOCH (201), Clinical Endocrinology Branch, National Institute of Arthritis and Metabolic Diseases, Bethesda, Maryland

CHOH HAO LI (101), The Hormone Research Laboratory, University of California, San Francisco, California

TA-HSIU LIAO (17), * Department of Biological Chemistry, University of California, School of Medicine, Los Angeles, California

HAROLD PAPKOFF (59), The Hormone Research Laboratory, University of California, San Francisco, California

JOHN G. PIERCE (17), Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, California

M. R. SAIRAM (101), The Hormone Research Laboratory, University of California, Los Angeles, California

G. SALVATORE (201), Istituto di Patologica Generale, University of Naples, Naples, Italy

RICHARD J. WINZLER (1), Department of Chemistry, Florida State University, Tallahassee, Florida

* Present address: The Rockefeller University, New York, New York.

Preface

As chemical compounds, mammalian hormones fall into three classes: steroids, phenol derivatives, and proteins or polypeptides. Among these hormonal compounds, protein and polypeptide hormones are most complex and hence it is difficult to isolate them in pure form, to elucidate their chemical structure, and to synthesize them in the laboratory. With the remarkable development of various techniques for protein chemistry in the last twenty years, our knowledge of the chemical nature of hormonal proteins and peptides has increased enormously.

Since 1953, the structures of the following protein and peptide hormones have been elucidated: oxytocin, vasopressin, gastrin, secretin, glucagon, calcitonin, cholecystokinin-pancreozymin, insulin, parathormone, proinsulin, human chorionic somatomammotropin, and human chorionic gonadotropin. The ten adenohypophyseal hormones, namely, adrenocorticotropin, α-melanotropin, β-melanotropin, β-lipotropin, γ-lipotropin, growth hormone, prolactin, interstitial cell-stimulating hormone, follicle-stimulating hormone, and thyrotropin, have been completely purified; the amino acid sequences of nine of them, with the exception of follicle-stimulating hormone, are known. In addition, two glycoproteins possessing hormonal activities have been highly purified: thyroglobulin and pregnant mare serum gonadotropin.

The purposes of this treatise are to review critically and extensively present knowledge on the chemistry and biology of these hormones. Included in each volume is one chapter on a general subject, which is considered to be of special interest to investigators in the field.

Volume I is devoted chiefly to the chemistry of several hormonal glycoproteins. Volume II includes one chapter on the solid-phase method of peptide synthesis and two chapters on peptide hormones. It is hoped that these and future volumes of this work will provide an important forum between protein chemistry and experimental endocrinology.

CHOH HAO LI

Contents

Contents of Volume II

1

The Chemistry of Glycoproteins

RICHARD J. WINZLER

It has long been known that certain of the protein hormones contain substantial amounts of carbohydrate. Some data on carbohydrates in glycoprotein hormones are summarized in Table I. This chapter considers some aspects of the general chemistry and metabolism of glycoproteins. Extensive reviews of this subject have recently appeared (Ginsburg and Neufeld, 1969; Spiro, 1970). Attention will be focused on problems and questions pertaining to the carbohydrate components, not because these are more important than the peptide components, but because the problems are perhaps less familar. Some of these problems are as follows:

1. Number, size, and structure of oligosaccharide chains in glycoproteins.
2. Linkage of oligosaccharide chains to peptides.
3. Microheterogeneity of oligosaccharide chains.
4. Biosynthesis of oligosaccharide chains.
5. Relation of carbohydrate to biological function.

A protein containing a single covalently linked monosaccharide can be considered to be a glycoprotein. A protein of the molecular weight of serum albumin containing a single molecule of hexose would have a carbohydrate content of 0.28%. Usually, however, glycoproteins contain more than this amount of carbohydrate. The range is very broad, however, and from less than 1% to more than 80% of the mass of glycoproteins may be carbohydrate.

Table I—Carbohydrates in Glycoprotein Hormones

Hormone	Source	Neutral sugar (%)	Acetyl hexosamine (%)	Sialic acid (%)	Fucose (%)	Reference
Chorionic gonadotropin	Human urine	14.0	11.4	10.3	0.6	Bahl, 1969
Serum gonadotropin	Pregnant mare serum	18.6	17.5	10.4	1.4	Bourrillon *et al.*, 1959
Follicle-stimulating hormone	Ovine pituitary	5.7	5.5	2.8	1.5	Papkoff *et al.*, 1967a
	Porcine pituitary	3.6	4.6	—	1.1	Papkoff, 1966
	Human pituitary	3.9	2.9	1.4	0.4	Papkoff *et al.*, 1967b
Luteinizing hormone (ICSH)	Ovine pituitary	6.5	7.8	—	1.4	Papkoff and Gan, 1970
						Kathan *et al.*, 1967
						Walborg and Ward, 1963
	Human pituitary	11.3	4.9	2.0	—	Kathan *et al.*, 1967
	Bovine pituitary	5.3	7.0	—	0.9	Papkoff and Gan, 1970
	Porcine pituitary	5.8	8.2	—	0.81	Hennen *et al.*, 1971
	Rat pituitary	4.4	11.4	—	1.0	Ward *et al.*, 1971
Thyrotropin	Bovine pituitary	14.2	9.0	—	0.5	Liao *et al.*, 1969
	Human pituitary	5.9	4.1	—	0.5	Kim *et al.*, 1967
Erythropoietin	Ovine anemic plasma	9.0	9.2	10.8	0	Goldwasser and Kung, 1971
Gonadotropin-transporting protein	Equine serum	14.0	13.3	12.0	—	Bourrillon *et al.*, 1958
Thyroglobulin	Human thyroid	4.8	4.2	1.1	0.5	McQuillan and Trikojus, 1966
	Porcine thyroid	4.0	3.4	1.2	0.5	McQuillan and Trikojus, 1966
	Ovine thyroid	4.0	3.2	1.5	0.4	McQuillan and Trikojus, 1966
	Bovine thyroid	3.7	3.2	1.4	0.4	McQuillan and Trikojus, 1966

Table II—Size, Number, and Type of Oligosaccharides in Some Glycoproteins

Protein	% CHO	Oligosaccharide (units/molecule)	Sugar residues per oligosaccharide	Number of types of oligosaccharide
Orosomucoid	41.4	5	18	1
Haptoglobin	18.6	13	14	1
Fetuin	22.9	3	>3,20<	2
Ribonuclease B	11.3	1	8	1
Transferrin	5.9	2	12	1
Thyroglobulin	10.6	24	>6,29<	2
Ovalbumin	3.2	1	8	1
Ovine submaxillary mucin	39.4	800	2	1
Erythrocyte membrane glycoprotein	64.5	20	>4,12<	2

The second column of Table II gives a range of carbohydrate values found in several well-studied glycoproteins. This carbohydrate consists primarily of two hexoses (D-galactose and D-mannose), two acetylated hexosamines (*N*-acetyl-D-glucosamine and *N*-acetyl-D-galactosamine), one methyl pentose (L-fucose), and a sialic acid (*N*-acetylneuraminic acid or *N*-glycolylneuraminic acid). In addition glucose may occur in a few glycoproteins such as collagen. The structures of these sugars are shown in Fig. 1.

The carbohydrate in glycoproteins may occur in a single relatively large oligosaccharide, or as a large number of relatively small oligosaccharides linked to the peptide chain. The number of oligosaccharide units per glycoprotein molecule ranges from 1–800 as is shown in the third column of Table II. Some of these are simple disaccharides. The ovine submaxillary mucin, for example, contains primarily a disaccharide consisting of *N*-acetylneuraminyl-*N*-acetyl galactosamine linked to the hydroxyl groups of serine or threonine in the peptide chain. Other glycoproteins, such as orosomucoid, contain large, highly branched oligosaccharides comprised of *N*-acetylneuraminic acid, fucose, galactose, mannose, and *N*-acetyl glucosamine. The range in size of the oligosaccharides in several glycoproteins is shown in the fourth column in Table II. The structures of the oligosaccharides from ovine submaxillary mucin and from orosomucoid are shown in Fig. 2.

Most glycoproteins contain carbohydrate units of only one structural pattern. However, a few proteins have oligosaccharide chains of more than one type (column five of Table II). These include thyroglobulin which contains one type of oligosaccharide comprised of sialic acid, galatose, acetyl glucosamine, and mannose, and another containing only mannose and acetyl glucosa-

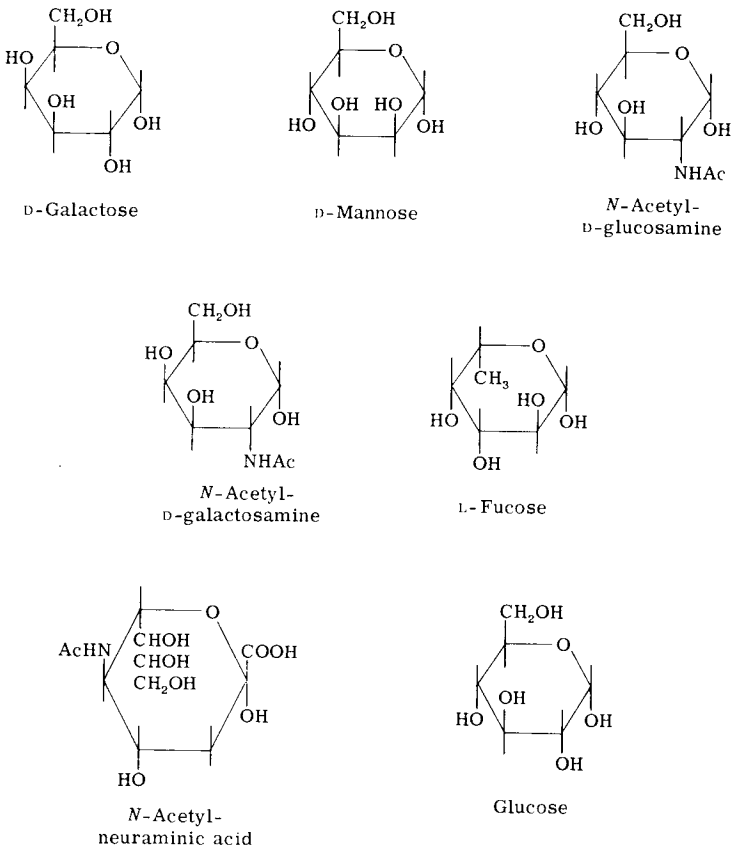

FIG. 1. Structures of monosaccharides occurring in glycoproteins.

mine. The major glycoprotein of human erythrocytes contains a tetrasaccharide consisting of two sialic acids, one galactose, and one acetyl galactosamine, and another containing sialic acid, fucose, galactose, mannose, and acetyl glucosamine. Fetuin also contains two types of carbohydrate, one of the sialic acid, galactose, acetyl glucosamine type, and a smaller unit consisting of sialic acid, galactose, and acetyl galactosamine.

Complete or partial structures have been worked out for the carbohydrate units of relatively few glycoproteins up to this time. The major chemical approaches to the elucidation of the structure of oligosaccharides in glycoproteins have included sequential removal of sugars by specific glycosidases, periodate oxidation, sequential degradation by alternate periodate oxidation, borohydride reduction, and mild acid hydrolysis (Smith degradation), and

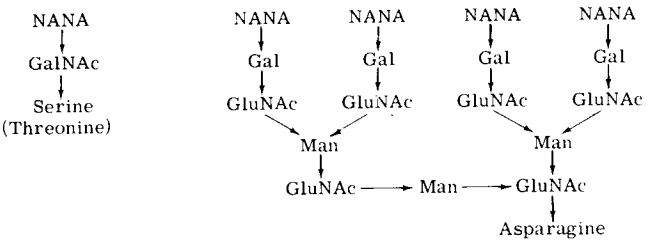

FIG. 2. Structures of disaccharide from ovine submaxillary mucin (left) and oligosaccharide from orosomucoid (right).

methylation and identification of the methyl ethers of monosaccharides following acid hydrolysis. Table III lists the glycoproteins whose oligosaccharide structures have largely been elucidated. There is little doubt that the structures of many other oligosaccharide units of glycoproteins will shortly be established.

Table III—Some Glycoproteins whose Oligosaccharide Structures Have Largely Been Elucidated

Glycoprotein	Source	Reference
α_2-Macroglobulin	Human serum	Dunn and Spiro, 1967
Blood group substances	Human ovarian cysts	Lloyd and Kabat, 1969
Barium α_2-glycoprotein	Human serum	Kamiyama and Schmid, 1961
Chorionic gonadotropin	Human urine	Bahl, 1969
Collagen, basement membrane	Various sources	Spiro, 1969
Deoxyribonuclease	Bovine pancreas	Catley *et al.*, 1969
Erythrocyte glycoprotein	Human erythrocytes	Thomas and Winzler, 1969 Kornfeld and Kornfeld, 1970
Fetuin	Bovine serum	Spiro, 1964
Fibrinogen	Bovine plasma	Bray and Laki, 1968
IgG immunoglobulins	Human serum	Clamp and Putnam, 1964
IgA immunoglobulins	Human myeloma serum	Dawson and Clamp, 1968
Orosomucoid	Human serum	Wagh *et al.*, 1969
Ovalbumin	Chicken eggs	Montgomery *et al.*, 1965 Makino and Yamashina, 1966
Ribonuclease B	Bovine pancreas	Plummer *et al.*, 1968
Submaxillary mucin	Ovine submaxillary	Graham and Gottschalk, 1960
Submaxillary mucin	Porcine submaxillary	Carlson, 1968
Thyroglobulin	Bovine thyroid	Spiro, 1965
Ceruloplasmin	Human serum	Jamieson, 1965

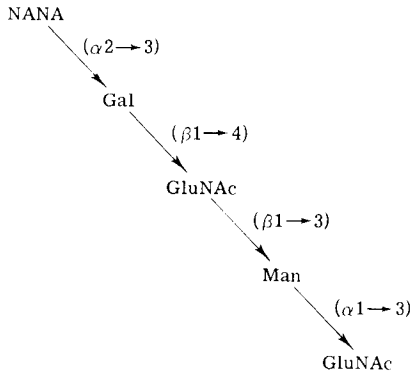

FIG. 3. A frequently occurring oligosaccharide in glycoproteins.

Detailed comparison of the structures of the relatively few oligosaccharides that have been established so far reveals that there are certain linkages which occur very frequently in mammalian glycoproteins. An especially frequently occurring structure is the heteropolysaccharide consisting of the sequence sialic acid (or fucose) →galactose→*N*-acetyl glucosamine→mannose. This structure, shown in Fig. 3, occurs in fetuin, α_2-macroglobulin, orosomucoid, thyroglobulin, barium α_2-glycoprotein, fibrinogen, chorionic gonadotropin, glomerular basement membranes, immunoglobulins and others.

The most unique structural feature of glycoproteins is the attachment of carbohydrate to the peptide chain. Three distinct types of peptide–carbohydrate bonds have been demonstrated in mammalian glycoproteins. These are shown in Fig. 4. The first of these to be clearly demon-

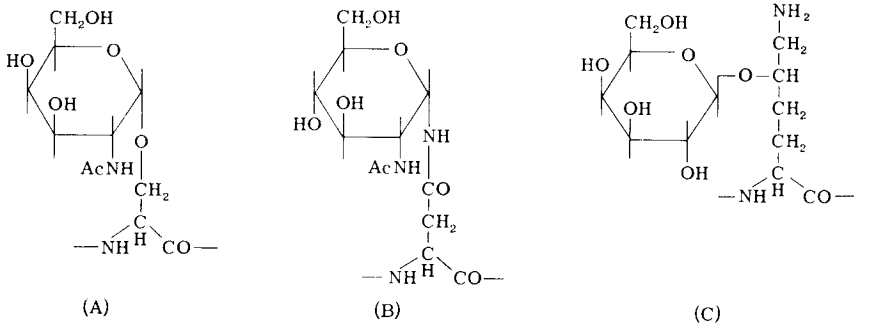

FIG. 4. Carbohydrate peptide linkages. A, *O*-glycoside; B, *N*-glycoside; and C, *O*-glycoside with hydroxylysine.

strated was that between the anomeric carbon of acetyl glucosamine and the amide group of asparagine. At the present time the only method of conclusively demonstrating this linkage is the isolation, usually in low yield, of 2-acetamido-1-(L-β-aspartamido)-1,2-dideoxy-β-D-glucose (Glucosaminyl asparagine or β-asparaginyl glucosamine) from acid hydrolyzates of the glycopeptides containing only aspartic acid. This isolation is possible since the glycosylamine bond is somewhat more stable to acid hydrolysis than are the glycosidic linkages between monosaccharides. This type of bond is found in the circulating plasma proteins synthesized in the liver as well as in the circulating immunoglobulins. It is also found in ovalbumin, ribonuclease B, takamalase, and many other proteins, including the glycoprotein hormones. A more quantitative and definitive method to identify this type of linkage in glycoprotein is badly needed.

The second type of bond between carbohydrate and protein is an α-*O*-glycosidic linkage between the anomeric carbon of a sugar (usually *N*-acetyl galactosamine in mammalian glycoproteins) and the hydroxyl group of serine and threonine. This bond is labile to alkali and is readily identified and quantitated by treatment with alkaline borohydride, and noting the loss of galactosamine and of the hydroxy amino acids and the appearance of galactosminitol. This linkage occurs in most of the epithelial mucins, including the submaxillary mucins, the blood group substances, and also in the major glycoprotein of human erythrocytes. Some glycoproteins have been shown to have both the *N*-glycoside type of linkage and the *O*-glycoside type of linkage (the major glycoprotein of human erythrocytes, fetuin, human chorionic gonadotropin and certain immunoglobulins).

The third type of linkage has been demonstrated in collagen and in basement membrane (Butler and Cunningham, 1966; Spiro, 1969). This consists of a β linkage between the reducing group of galactose and the hydroxyl group of hydroxylysine. This linkage is stable to alkali and the hydroxylysine-linked glucosyl galactose can be isolated in high yield following alkaline digestion.

Not all of the potential linkage amino acids (asparagine, serine, threonine, hydroxylysine) in a given glycoprotein are linked to oligosaccharide units. Therefore some kind of recognition site in the peptide chain must be involved in dictating the location of the oligosaccharide chains. Since the oligosaccharide units appear to be built up in stepwise fashion by addition of monosaccharides to a growing saccharide chain, the first glycosylation step must involve recognition sites on the peptide chain. The sequence of amino acids around the asparagine linked to *N*-acetyl glucosamine by an *N*-glycosidic linkage has been investigated in a number of glycoproteins, and is shown in Table IV. From this table it would appear that the sequence Asn-X-Ser (or Thr) could be the recognition sign for the formation of this type

Table IV—Amino Acid Sequence around Asparagine-linked Oligosaccharide Units of Several Glycoproteins

CHO ↓				Reference
Glu-Lys-Tyr-	Asn	-Leu-Thr-Ser	Ovalbumin	Nuenke and Cunningham, 1961
Leu-Ile-His-	Asn	-Arg-Thr-Gly	Ovotransferrin	Williams, 1968
Leu-Gly-Ser-	Asn	-Met-Thr-Ile	Avidin (Asn-17)	DeLange, 1970
Gln-Gln-Tyr-	Asn	-Ser-Thr-Tyr	IgG, H chain, human (Asn-297)	Edelman *et al.*, 1969
Gln-Gln-Phe-	Asn	-Ser-Thr-Ile	IgG, H chain, rabbit	Hill *et al.*, 1967
Ala-Ser-Gln-	Asn	-Ile-Ser-Asn	IgG, L chain, mouse	Melchers, 1969
Gln-Val-Glu-	Asn	-Lys-Thr-Ser	Fibrinogen, human γ chain (Asn-60)	Iwanaga *et al.*, 1968
Ala-Leu-Glu-	Asn	-Ala-Thr-Arg	Thyroglobulin, human	Rawitch *et al.*, 1968
Ser-	Asn	-Ala-Thr	DNase, bovine	Catley *et al.*, 1969
Lys-Ser-Arg-	Asn	-Leu-Thr-Lys	RNase B, bovine (Asn-34)	Plummer and Hirs, 1964
Ser-Ser-Ser-	Asn	-Ser-Ser-Asn	RNase, porcine (Asn-21)	Jackson and Hirs, 1970
Ser-Arg-Arg-	Asn	-Met-Thr-Gln	RNase, porcine (Asn-34)	Jackson and Hirs, 1970
Tyr-Gln-Ser-	Asn	-Ser-Thr-Met	RNase, porcine (Asn-76)	Jackson and Hirs, 1970
Cys-Ile-	Asn	-Val-Thr-Thr	Human chorionic gonadotropin (1)	Bahl, 1900
Arg-Pro-Ile-	Asn	-Ser-Thr-Leu	Human chorionic gonadotropin (2)	Bahl, 1900
Pro-Ile-	Asn	-Ile-Thr	TSH α chain (Asn-56)	Pierce, 1900
Arg-Val-Glx-	Asn	-His-Thr	TSH α chain (Asn-83)	Pierce, 1900
Ile-	Asn	-Thr-Thr	TSH β chain (Asn-22)	Pierce, 1900
Val-Pro-Lys-	Asn	-Ile-Thr-Ser	LH α chain, ovine (Asn-56)	Liu *et al.*, 1972; Papkoff *et al.*, 1971
Arg-Val-Glu-	Asn	-His-Thr-Glu	LH α chain, ovine (Asn-82)	Liu *et al.*, 1972; Papkoff *et al.*, 1971
Glu-Pro-Ile-	Asn	-Ala-Thr-Leu	LH β chain, ovine (Asn-13)	Liu *et al.*, 1970; Papkoff *et al.*, 1971
Pro-Ile-Thr-	Asn	-Ala-Thr-Leu	Orosomucoid	Schmid *et al.*, 1971
	Asn	-Lys-Ser	Orosomucoid	Schmid *et al.*, 1971
	Asn	-Lys-Thr	Orosomucoid	Schmid *et al.*, 1971
	Asn	-Thr-Thr	Orosomucoid	Schmid *et al.*, 1971
	Asn	-Gly-Thr	Orosomucoid	Schmid *et al.*, 1971

of glycopeptide bond. The sequence on the amino terminal side of the asparagine linkage does not seem to bear any specific pattern.

The question of when in the biosynthetic process this first acetyl glucosamine is attached to the peptide chain is still a matter of controversy, al-

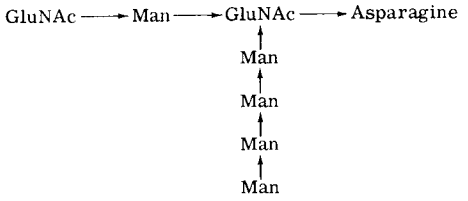

FIG. 5. Complete glycopeptide from ovalbumin.

though considerable evidence indicates that it may occur while the nascent peptide is still associated with the ribosome.

Comparable studies have not been made with the *O*-glycosidic type of carbohydrate peptide bond. However, in this case, too, specific recognition signs on the peptide are likely to be involved, since not all serines and threonines are glycosylated. It has been possible to transfer enzymically *N*-acetyl galactosamine from UDP *N*-acetyl galactosamine to ovine submaxillary mucin from which carbohydrate had been removed by treatment with neuraminidase and an α-galactosidase (McGuire and Roseman, 1967). In this work no ribosomes were involved. It will probably turn out that glycosylation occurs on the ribosome in some glycoproteins and postribosomally in others.

One of the major problems facing the investigator of glycoprotein structure is the occurrence of microheterogeneity in the oligosaccharide chains. This heterogeneity may be of several types. One type of heterogeneity is the occurrence in almost all glycoproteins so far studied of incompleted oligosaccharide chains.

Cunningham *et al.* (1965) found markedly different ratios of mannose to acetyl glucosamine in glycopeptides prepared from ovalbumin. Since this glycoprotein contains only one oligosaccharide unit per molecule, the variations must represent differences in composition of the oligosaccharides occurring in different molecules. Similar variations were observed even from ovalbumin from a single hen. It appears that this heterogeneity results from incompleted versions of the largest and "most complete" oligosaccharide shown in Fig. 5. In some molecules the complete oligosaccharide is made, whereas in others the ovalbumin is secreted before the oligosaccharide is completed.

Other examples of microheterogeneity due to incompleted oligosaccharides include those in the α_2-macroglobulin of human plasma (Dunn and Spiro, 1967), in the oligosaccharide from porcine submaxillary mucins (Carlson, 1968), in human plasma orosomucoid (Yamashina *et al.*, 1965; Wagh *et al.*, 1969), and in the alkali-stable and alkali-labile oligosaccha-