Amgen Inc. v. F. Hoffmann-LaRoche LTD et al Doc. 578 Att. 20

EXHIBIT R

Lodish Decl. in Support of Opposition to Roche's Motion for Summary Judgment of Invalidity for Double Patenting Over Claim 10 of the '016 Patent

Dockets.Justia.com

L 148

Managing Editor
Stella Dutton

Assistant Editor
Monique Heald

Sub-Editor
Diane Cogan

Enzyme and Microbial Technology is an international quarterly journal published in January, April, July and October. Authors intending to submit manuscripts for consideration are urged to study the notes for authors given on the inside back cover.
Annual subscription (4 issues) including postage by surface mail £60.00 (UK and overseas). Single copies £15.00 (UK and overseas). Personal subscription £18.00 annually (UK and overseas). Airmail rates are available on request. Copies of this journal sent to subscribers in the USA and Canada are airspeeded for quicker delivery at no extra cost (annual subscription. $156.00; single copies $39.00; personal subscription $46 80 annually). US mailing agents: Expediters of the Printed Word Limited, Suite 1217, 527 Madison Avenue, New York, NY 10022, USA. Second class postage paid at New York, NY.
Subscription enquiries and orders: IPC (Sales and Distribution) Ltd, Oakfield House, Perrymount Road, Haywards Heath, Sussex RH16 3DH, England. Telephone 0444 59188. Telex and Telegrams: 25137 BISPRSFG. Back issues to the current volume are available from Wm Dawson & Sons Ltd, Cannon House, Folkestone, Kent CT19 5EE. Telephone: 0303 57421.
Microfiche. The journal is also published on 98 frame microfiche A6 size. Prices are the same as the printed edition but include binder and dispatch by airmail.
Published by IPC Science and Technology Press Ltd, P.O. Box 63, Westbury House, Bury Street, Guildford, Surrey GU2 5BH, England. Telephone 0483 31261. Telex and Telegrams: 859556 SCITEC G.
Managing Director: A. A. Smailes.
Publishing Director: C. J. Rawlins.
American Representatives:
IPC Business Press, 205 East 42nd Street, New York, NY 10017, USA.
Telephone: (212) 867 2080.
Telex: 238327 INTRA URM.
Reprints (minimum quantity 100) of papers may be ordered from the publishers. Write to the Reprints Department at Guildford.
Copyright: Readers who require copies of papers published in this journal may either purchase reprints or obtain permission to copy from the publisher at the following address, IPC Science and Technology Press Ltd, P.O. Box 63, Westbury House, Bury Street, Guildford GU2 5BH, England. For readers in the USA, permission to copy is given on the condition that the copier pay the stated per copy fee through the Copyright Clearance Center, Inc., 21 Congress Street, Salem, MA 01970 (617) 744–3350 for copying beyond that permitted by Sections 107 and 108 of the US Copyright Law. Fees appear in the code that appears at the foot of the first page of each paper.
ISSN 0141–0229
All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or be transmitted, in any form or by any means (electronic, mechanical, photocopying, recording or otherwise) without the written permission of the Publisher.
© IPC Business Press Ltd. 1981
Typeset by DJS Spools Ltd, Horsham
Printed by A. Wheaton and Co. Ltd, Exeter

enzyme and microbial technology

volume 3, number 3, july 1981, pages 177–280

EMTED2 3(3) 177–280 (1981)

Editors: Oskar R. Zaborsky and John F. Kennedy

Reviews

178 Post-translational modification of proteins
Harvey F. Lodish

189 Potential applications of bio-ligninolytic systems
T. Kent Kirk and Hou-min Chang

197 Metabolism of carbon monoxide
Robert L. Uffen

207 Biorefining of biomass to liquid fuels and organic chemicals
P. F. Levy, J. E. Sanderson, R. G. Kispert and D. L. Wise

Papers

216 Electron microscopic evidence of an immobilized living cell system
Véronique Larreta Garde, Brigitte Thomasset and Jean-Noel Barbotin

219 Catalytic properties of extracts from tyrosine phenol lyase producing cells of *Escherichia intermedia*: multienzyme complex instead of the single multisubstrate enzyme
N. G. Faleev, M. S. Sadovnikova, N. S. Martinkova and V. M. Belikov

225 Glycerol, α-glycerophosphate and other compounds as stabilizers of alcohol dehydrogenase from yeast
M. J. Brougham and D. B. Johnson

229 Effect of the nature of proteins on their coupling to different epoxide-containing supports
I. Zemanová, J. Turková, M. Čapka, L. A. Nakhapetyan, F. Švec and J. Kál

233 Glycolytic pathway as an ATP generation system and its application to the production of glutathione and NADP
Kousaku Murata, Keiko Tani, Jyoji Kato and Ichiro Chibata

243 Differential effect of cycloheximide on penicillin and cephalosporin biosynthesis by *Cephalosporium acremonium*
Y. M. Kennel, C. J. Behmer and A. L. Demain

246 Immobilization of *Bacillus subtilis* α-amylase on zirconia-coated alkylamine glass with glutaraldehyde
V. Ramesh and Chanan Singh

249 Steady state analysis of the enhancement in ethanol productivity of a continuous fermentation process employing a protein–phospholipid complex as a protecting agent
C. K. Jin, H. L. Chiang and S. S. Wang

Short paper

258 Stability of immobilized alcohol dehydrogenase prepared using the thermostable enzyme from yeast mitochondria
D. B. Johnson and M. J. Brougham

260 Literature survey

267 Patent reports

271 Industrial views

274 Conference reports

278 Books reviews

279 Publisher's announcement

280 Calendar

reviews

Post-translational modification of proteins

HARVEY F. LODISH

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA and BioInformation Associates, 65 E India Row, Boston, MA 02110, USA

Summary. *Many proteins, especially those produced by eukaryotic cells, undergo extensive, essentially irreversible, modifications after their synthesis. This review focuses on three classes of such reactions proteolytic cleavages, formation of S–S cystine bonds, and formation of asparagine-linked carbohydrate chains. Emphasis is placed on the mechanism of these reactions, and on the importance of these modifications for the proper structure, function and stability of the affected proteins. Using recombinant DNA techniques, it is now possible to synthesize the polypeptide portion of many proteins, such as mammalian peptide hormones and enzymes, in bacterial and yeast cells These host cells, however, may be unable to carry out essential post-translational modifications. Ways in which the properly modified form of these 'engineered' proteins can be produced are considered.*

Introduction

Through the use of modern techniques of molecular genetics and recombinant DNA, it is now possible to synthesize the polypeptide portion of many animal cell proteins, including enzymes, peptide hormones, and viral proteins, in bacterial or fungal (yeast) cells. Some of these, such as human insulin and human interferon, are rapidly approaching the commercial production phase, and it can be anticipated that many more proteins produced by this procedure will be important in the agricultural, pharmaceutical and chemical industries in the coming years.

The sequence of 20 common amino acids in a polypeptide is determined by the sequence of triplet nucleotide codons in the messenger RNA. This information, in turn, is a copy of that in the DNA genome. However, most, if not all, proteins are modified extensively after their synthesis by specific cuts, proteolytic cleavages, in the peptide chain, or by chemical modification of certain amino acid residues, or by the addition of one or more non-amino acid prosthetic groups, such as sugar, lipid or phosphate residues. For several reasons, these post-translational modifications are of interest and importance to the biochemical engineer. The nature of these modifications is very different for the three types of cells used for these genetic manipulations – bacterial cells (*Escherichia coli*, in particular), yeast and mammalian cells. For instance, there are few documented cases of glycoproteins in bacteria,[1] and there is no evidence that they contain asparagine-linked oligosaccharides of the sort so common in mammalian and yeast proteins. There are, too, many significant differences between the carbohydrate substituents of yeast and mammalian glycoproteins, and these will be discussed. The specific nature of these carbohydrate substituents can be of great importance in determining the stability and function of glycoproteins.

Other covalent modifications of proteins, such as proteolytic cleavages, differ not only between bacterial and eukaryotic cells, but may also differ among the various specialized cell types of a multicellular animal; these, too, are often essential for proper function and stability of a protein. Further, recent work has demonstrated that many modifications of viral proteins, such as specific proteolytic cleavages, are due to enzymes which are encoded by the viral genome, and which are different, therefore, from cellular enzymes. It is possible that intervention in the replication of these viruses could be achieved by selective inhibition of these reactions.

Covalent modification of proteins, the principal subject of this review, are of two types, reversible and irreversible. The catalytic activity of many cellular enzymes is often mediated by the covalent addition and removal of a variety of substituents, such as phosphate, AMP and UMP residues.[1a,2–6] An understanding of these reversible molecular alterations is clearly important in any application of these enzymes as catalysts for the production or modification of chemicals. Also, in 'moving' such enzymes from one cell type to another by recombinant DNA techniques, one must ensure that the enzymes which carry out these modifications are also present in the recipient cell in order for the 'transferred' protein to be optimally active. As extensive reviews are available concerning the nature and mechanism of these reversible regulatory modifications,[1a,2–6] they will not be considered further in this review. Rather, we shall focus on irreversible modifications of proteins, particularly those which take place on proteins which are secreted from the cell, such as hormones and viral proteins.

This paper will emphasize three of the well-understood types of irreversible modifications – proteolytic cleavages, glycosylation, and formation of cystine (S–S) disulphide bonds – which occur in animal cells. After describing these reactions and their importance for protein function, the possibility of synthesizing proteins with these modifications in bacterial or yeast cells will be considered.

There are many other essentially irreversible modifications that will not be covered. These include sulphation, hydroxylation of lysine or proline residues (in collagen), vitamin K-dependent carboxylation, methylation, acetylation, and ADP-ribosylation.[7–10]

Proteolytic cleavages of animal cell proteins and viral proteins

Animals and their cells contain a large number of proteolytic activities, although relatively few intracellular proteases – mostly the lysozomal enzymes – have been purified to date.[11–14] Importantly, these activities appear to be localized to specific subcellular regions or membrane-limited

0141–0229/81/030178–11 $02.00
© 1981 IPC Business Press

organelles, or are extracellular, e.g. in the serum. Many proteins, such as viral structural proteins and secreted cellular proteins, are acted upon intracellularly and sequentially by a series of such enzymes; the protein can perhaps be visualized as moving on a subcellular conveyor belt, subjected to specific proteolytic (and other) modifications as it moves along the assembly line preparatory to export.[15–17] It is this multitude of successive reactions, unique to higher cells, that will be difficult to reconstruct in bacterial and possibly in yeast cells by using genetic engineering techniques.

The internal structural proteins of many picornaviruses and RNA tumour viruses are modified by a series of proteolytic activities.[18–21] Poliovirus, a typical member of the picornavirus group, is a well-studied example. The poliovirus genomic RNA is the only known polio messenger RNA. It contains ~5000 nucleotides, and is translated by ribosomes in the cytoplasm of an animal cell into a 'polyprotein' which contains >2000 amino acids (*Figure 1*). This protein contains the amino acid sequences of all of the poliovirus-specified proteins. Even while this protein is still being fabricated on the ribosomes, two specific proteolytic cleavages take place. The enzyme(s) which catalyses this reaction is believed to be a normal constituent of the cell cytoplasm.[18] The three resultant polypeptides are, after a lag period lasting several minutes, processed proteolytically into even smaller proteins. The capsid precursor (p95, MW 95 000) for instance, is cleaved into three smaller proteins – VP_0, VP_1 and VP_3, which bind with viral RNA to form a virus-like particle. In order for this particle to be infectious, however, an additional cleavage of VP_0 into two smaller proteins, VP_2 and VP_4, is essential; this reaction takes place after VP_0 is assembled into the 'provirus' particle. All of these later proteolytic cleavages are, as indicated in the Figure, probably catalysed by a virus-encoded enzyme.This active enzyme is itself

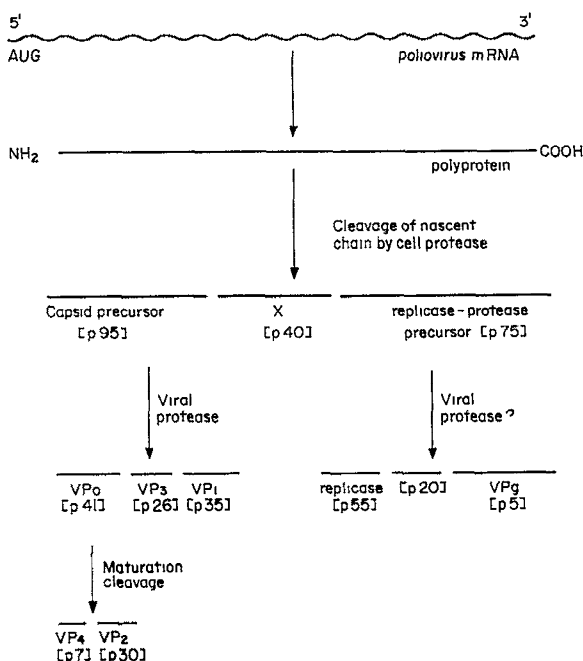

Figure 1 Proteolytic cleavage of poliovirus proteins in infected cells. The numbers given refer to estimated molecular weights (in SDS gels). The VP designation refers to a structural polypeptide, present in virus capsids. Redrawn after refs 18 and 21

derived from the proteolytic cleavage of one of the viral protein precursors, probably the one indicated 'replicase (P75)'.[18] The precursor of the RNA tumour virus (retrovirus) capsid proteins is processed in much the same way as the polio capsid precursor, and it also appears that a virus-specified protease might be involved.[22–24] Although these virus-specified proteases have not been purified to homogeneity, further studies of their substrate specificity and mechanism might provide the basis of design of specific anti-viral compounds.[18–25]

Synthesis of most proteins in bacterial and in higher cells is initiated with a methionine residue. This amino acid is usually removed proteolytically from the nascent chain. However, these proteases have never been purified, so one is uncertain of the substrate specificity of the enzyme or whether the specificities of bacterial, yeast and mammalian enzymes are the same.

Many mammalian proteins of medical or commercial interest are secreted by the cell. These include all peptide hormones, such as insulin and interferon, growth factors, serum proteins, and many enzymes, such as trypsin and amylase. These are made on the ribosome-studded membranous organelle termed the rough endoplasmic reticulum (ER), and are transported during or immediately after their synthesis into the ER lumen (*Figure 2b*).[15,16,26–32] The proteins move next to the Golgi complex of intracellular membranes. Depending on the protein, it is then secreted immediately, or is stored in secretory granules awaiting a hormonal stimulus for secretion (*Figure 3a*).[15,16,33] Extensive proteolytic processing generally occurs in each of the above organelles.

The biogenesis of many surface membrane glycoproteins, such as the Vesicular Stomatitis Virus G protein and the histocompatability proteins HLA-A and HLA-B, follows an analogous RER → Golgi → cell surface pathway (*Figure 3b*). However, these proteins remain anchored in the phospholipid membranes of the various organelles, rather than soluble in the organelle lumen (*Figure 2a*).[17,24,31]

In the case of most, but not all, secretory and membrane glycoproteins in which at least the NH_2-terminal segment, if not the whole protein, crosses the ER membrane, 15–26 NH_2-terminal amino acid residues are removed proteolytically while the polypeptide is still growing on the ribosome.[16,28–31,35–37] Thus, this so-called 'pre' piece is not normally observed on intracellular forms of the completed protein. Its existence was noted when mRNAs encoding these proteins were translated in cell-free systems in the absence of ER membranes, which are the source of these protease(s); the resultant proteins contained the extra 'pre' residues.[26,29,38] These 'pre' sequences are often called 'signal' sequences, and are believed to be involved in directing the ribosome which is translating the mRNA to the ER membrane. They have been found on a variety of both bacterial and mammalian proteins. *Figure 4* shows that there is little amino acid sequence homology among the various 'pre' sequences, although most do contain a contiguous stretch of 8–10 strongly hydrophobic amino acids, including leucine, isoleucine, and valine residues. There is also little homology with respect to the site of cleavage by this 'signal peptidase'. Generally, but not invariably, it occurs at the carboxyl side of a small, neutral amino acid side chain. Few studies have been reported on the purification and properties of these peptidases,[39,40] and it is not even clear how many different such proteases there are in a cell. Nor, more importantly, is it clear whether bacterial cells can properly cleave off the 'pre' peptides on animal proteins. There is one report that

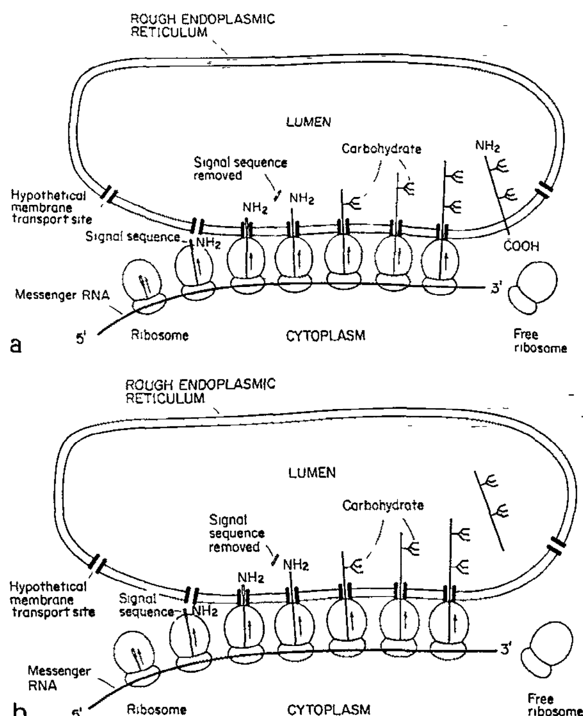

Figure 2 Co-translational insertion of a transmembrane glycoprotein (VSV G protein) into the ER membrane (a) and (b) co-translational sequestration of a secretary protein into the lumen of the ER. Among the first 50 amino acid residues of G or of a secretory protein is a 'signal sequence' that identifies the protein as one destined to be inserted into the membrane of the rough ER. Because some 40 amino acids remain buried in the ribosome, the signal sequence does not emerge until the polypeptide is ~70 amino acid units long. At that time the signal sequence is recognized by some molecule, presumably a protein, in the membrane of the ER. This hypothetical protein is thought to facilitate the passage of the polypeptide through the lipid bilayer. Once in the lumen of the reticulum at least part — 16 to 25 amino acids, depending on the protein — of the signal sequence generally is removed. The protein continues to elongate, and as it grows it is extruded through the membrane and folds up in the lumen. As it enters preformed carbohydrate side chains (*Figure 10a*) are transferred to G or to most secretory proteins from a lipid carrier. Proteins secreted by the cell pass all the way through the membrane but for reasons that are not totally clear, G becomes stuck at about the time that *translation is completed, with some 30 amino acids at the COOH* terminus remaining in the cytoplasm. Thus the completed G glycoprotein has its amino terminus, most of its bulk and all of its carbohydrates in the lumen of the ER, and a short stub that includes the carboxyl terminus on the cytoplasmic side.[17,28,32,34]

E. coli can correctly process pre-proinsulin to proinsulin (and will secrete proinsulin),[41] but other proteins such as pre-fibroblast (human) interferon are apparently not processed in this way.[42]

In the case of some secreted proteins, such as growth hormone, placental lactogen, lysozyme, and ovomucoid, and certain viral membrane proteins, such as the Vesicular Stomatitis Virus glycoprotein, removal of the 'pre' sequence is the only known proteolytic cleavage; the 'pre' form is converted directly into the mature protein.[17,43] In most cases, however, there is an additional, relatively long-lived intracellular form, termed the pro-protein or pro-hormone (*Figures 5* and *6*).[31,44–47] Proteolytic conversion of the pro-protein to the mature molecule occurs at a late stage in intracellular maturation, probably within the Golgi membrane complex.[46,49,50] In contrast to the 'pre' amino acids, which are always at the NH_2terminus, the additional amino acids in the 'pro'-form can be either at the NH_2- (proPTH, proalbumin) or COOH- (proglucagon) terminal region of the proprotein (*Figure 5*). In the case of proinsulin, the 'extra' amino acids, termed the C peptide, are located internally in the polypeptide. The NH_2-terminal B chain and the COOH-terminal A chain of mature insulin are linked, in proinsulin, by disulphide bonds, and remain attached when the C peptide is removed.[51]

Examination of the amino acid sequences surrounding the cleavage sites in preprotein to protein conversions suggests a common mechanism for conversion of many proproteins. The cleavage site is invariably comprised of two or three adjacent basic amino acids – lysine or arginine. In the case of proalbumin, proPTH and progastrin, a single endopeptidase with trypsin-like specificity is sufficient to generate the mature protein. In others, such as proinsulin

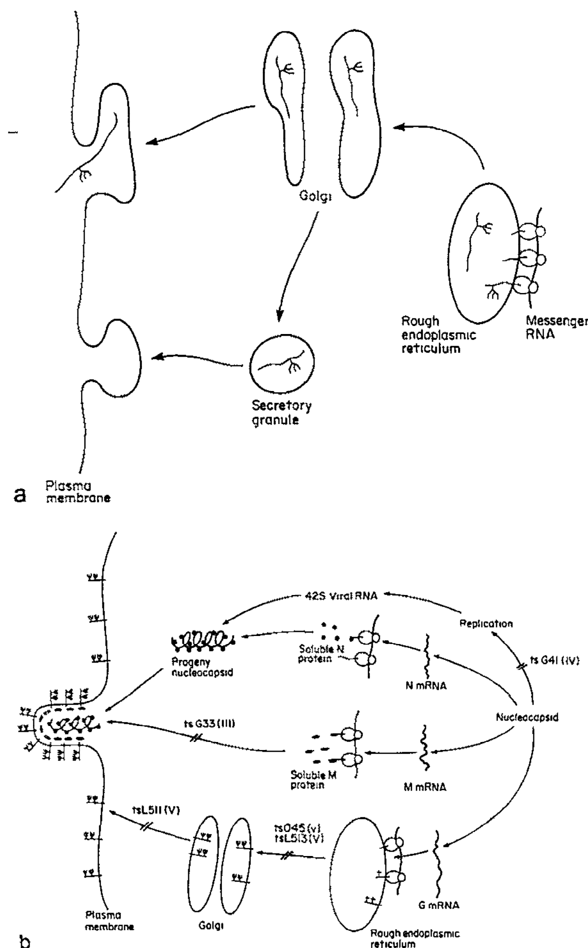

Figure 3 Maturation of secretory and membrane glycoprotein. (a) Secretory proteins. They migrate first to the Golgi complex of membranes, where there occurs certain proteolytic cleavages and also modifications to Asn-linked oligosaccharides. In some cells, such as the liver, the proteins are secreted constitutively; in others, such as the exocrine pancreas, they are stored in secretory granules awaiting a hormonal signal for secretion.[16] (b) Schematic diagram illustrating the pathways of maturation of the major structural proteins of VSV, especially the G glycoprotein. Shown are the blocks in virion assembly due to temperature-sensitive lesions in viral structural proteins; ts L511, ts L513, and ts 045 are mutants in the G polypeptide.[115] +, Denotes a high mannose oligosaccharide; Ψ, denotes a complex oligosaccharide in G. Redrawn from ref 17

Case 1:05-cv-12237-WGY Document 578-21 Filed 06/29/2007 Page 6 of 13

	−30 … −10 … ⁝ +1
Pre-proalbumin	Met. Lys. Trp. Val. Thr. Phe. Leu. Leu. Leu. Leu. Phe. Ile. Ser. Gly. Ser. Ala. Phe. Ser. A⁝rg. . . .
Pre-IgG light chain	Met. Asp. Met. Arg. Ala. Pro. Ala. Gln. Ile. Phe. Gly. Phe. Leu. Leu. Leu. Leu. Phe. Pro. Gly. Thr. Arg. Cys. A⁝sp. . . .
Pre-lysozyme	Met. Arg. Ser. Leu Leu. Ile. Leu. Val. Leu. Cys. Phe. Leu. Pro. Leu. Ala. Ala. Leu. Gly. L⁝ys. . . .
Pre-prolactin	Met. Asn. Ser. Gln. Val. Ser. Ala. Arg. Lys. Ala. Gly. Thr. Leu. Leu. Leu. Leu. Met. Met. Ser. Asn. Leu. Le⁝u. Phe. Cys. Gln. Asn. Val. Gln. Thr. Leu. . . .
Pre-penicillinase (*E. coli*)	Met. Ser. Ile. Gln. His. Phe. Arg. Val. Ala. Leu. Ile. Pro. Phe. Phe. Ala. Ala. Phe. Cys. Leu. Pro. Val. Phe. Ala. H⁝is. . . .
Pre-vesicular stomatitis virus glycoprotein	Met. Lys. Cys. Leu. Leu. Tyr. Leu. Ala. Phe. Leu. Phe. Ile. His. Val. Asn. Cys. L⁝ys. . . .
Pre-lipoprotein (*E. coli*)	Met. Lys. Ala. Thr. Lys. Leu. Val. Leu. Gly. Ala. Val. Ile. Leu. Gly. Ser. Thr. Leu. Leu. Ala. Gly. C⁝ys. . . . ↑ CLEAVAGE POINT

Figure 4 Amino acid sequences of pre-pieces of several secretory and membrane proteins. The vertical dotted line denotes the site of cleavage of these residues: this generally occurs during elongation of the growing polypeptide chain. Redrawn after ref 28

and proglucagon, both the action of this endopeptidase and an exopeptidase with the specificity of carboxypeptidase B are essential to generate the mature protein (see *Figure 5* and refs cited therein). Highly efficient conversion of proPTH to PTH in >85% yield could be accomplished by dilute solutions of pancreatic trypsin.[46] Likewise, preinsulin can be converted efficiently to insulin *in vitro* by the combined action of trypsin and carboxypeptidase B.[55] The nature of the enzymes which catalyse these cleavages within animal cells is unknown, but they appear to be localized to the Golgi.

A particularly interesting example concerns the biogenesis of the peptide hormones ACTH, β-lipotropin, β endorphin,

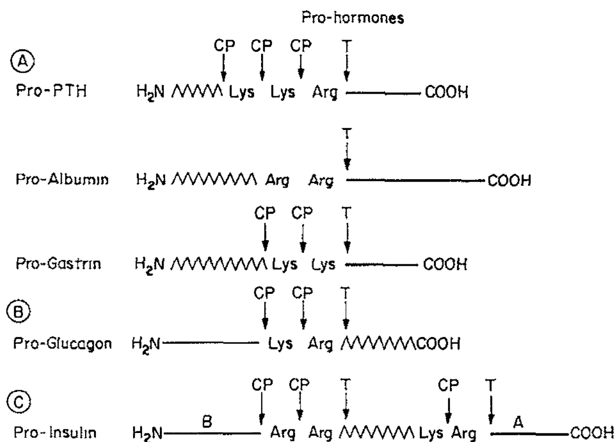

Figure 5 Schemes for combined actions of tryptic (T) and carboxypeptidase B-like (CP) activities in the conversion of proproteins (prohormones) to the authentic product: A, Mechanism whereby tryptic cleavage alone results in formation of final product; B, combined action of the two enzymes is necessary for production of glucagon; C, both cleavage mechanisms illustrated by (A) and (B) are involved in the formation of insulin and C peptide from proinsulin. Data sources are as follows: ProPTH, ref 52; proalbumin, ref 53; progastrin, ref 54; proglucagon, ref 31; proinsulin, ref 50. *Solid lines denote the mature form of the protein; jagged lines* denote the 'pro' sequences

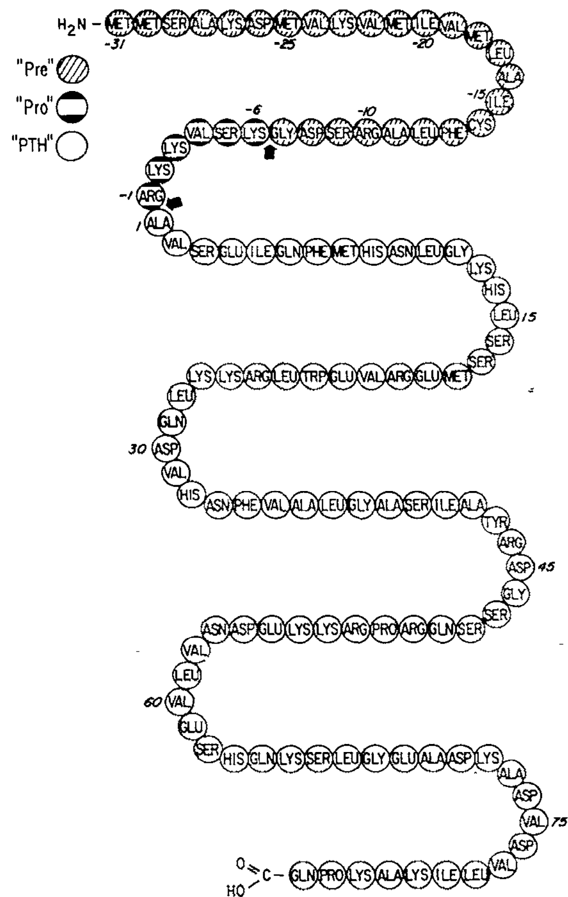

Figure 6 Amino acid sequence of pre-proparathyroid hormone (pre-proPTH). Residues in hatched circles constitute the NH_2-terminal 'pre' or signal sequence. Shaded residues are specific to the prohormone. Residues in open circles denote PTH, the principal secreted form of the hormone. The arrows denote the peptide band cleaved in removal of, first, the pre, and later, the pro extensions. Redrawn from ref 47

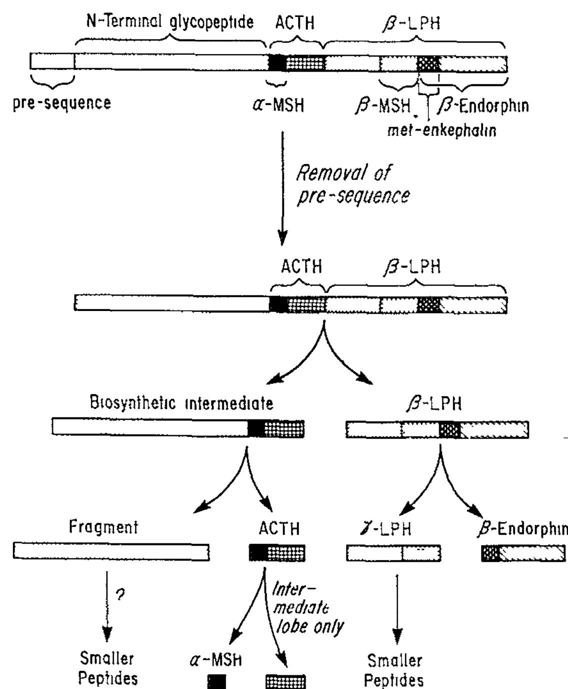

Figure 7 Structure of the common precursor molecule of ACTH and β-endorphin, and pathways of synthesis of the smaller peptides in various pituitary cells. Note that the cleavage of ACTH to α MSH (and also other cleavages) takes place only in intermediate lobe cells. Data taken from refs 57 and 58

and α MSH by the rat pituitary. As is indicated in *Figure 7*, all are synthesized as part of a long precursor molecule which contains ~300 amino acids.[56–58] The initial proteolytic cleavage of the common precursor for ACTH and β LPH occurs ~20 min after synthesis. This is followed rapidly by a second proteolytic cleavage which separates the glycosylated NH_2-terminal segment from the mature ACTH. On a slower time scale, β LPH is cleaved to produce β endorphin and γ-lipotropin (γ LPH). In the intermediate lobe of the pituitary, but not the anterior lobe, ACTH is further cleaved to the hormone α MSH. It is believed that the enzyme which catalyses this specific cleavage is localized in the intermediate lobe cells.

Analogous proteolytic cleavages occur during maturation of many viral surface glycoproteins, and are important for infectivity of the virus. Many animal viruses contain a lipid envelope, and bud off from the surface of infected cells. Examples include influenza virus, Newcastle disease virus (a pest of bird flocks), measles virus, Vesicular Stomatitis Virus, and others. In all cases, these viruses encode one or more glycoproteins which accumulate on the surface membranes of the cell and on the surface of the budded viral particles (see *Figure 3b*). These glycoproteins are essential for adsorption of viruses to the target cells, and determine the host range of the virus. Also, they are the targets of immunological response to the infecting virus. *Figure 8* shows the maturation process of the influenza HA glycoprotein, one of many of this class that undergoes a late proteolytic cleavage. The two polypeptides, HA_1 and HA_2, which comprise the haemaglutinin glycoprotein are formed as a precursor HA_0 glycoprotein. The intracellular proteolytic cleavage of HA_0 to HA_1 and HA_2 is essential for infectivity of the resultant virus;[59–62] this processing can be reproduced *in vitro* by treatment of non-infectious virions, containing HA_0, with trypsin.[59–61] Similar cleavages occur during the maturation of the F (fusion) glycoprotein of the unrelated paramyxovirus Newcastle Disease Virus.[63]

It is important to note, finally, that precursors of some secreted proteins are cleaved after secretion. Well-studied examples include the zymogen precursors of digestive enzymes, fibrinolysins, complement, and blood clotting factors. Here, cleavage is essential for the induction of enzyme activity. The reactions can, in some cases, occur in cell-free reactions using purified proteases.[11,14] Likewise, proteolytic removal of sequences from both the NH_2- and COOH-terminal of the three polypeptide chains of the procollagen molecule occur at, or just after, secretion. These cleavages facilitate assembly of the proper multichain quaternary structure.[7,64]

It is most unlikely that bacterial cells (and probably yeast cells) contain enzymes which can correctly process the prepro or pro-forms of mammalian proteins. How, then, will it be possible to convert the proprotein or pre-proprotein, synthesized in bacteria as a result of recombinant DNA methodology, to the mature protein?

A first possibility, as noted above, is to conduct the cleavage reaction *in vitro*, using one or more purified enzymes. This should work most efficiently for proteins, such as proPTH, in which only a single proteolytic cleavage, by an enzyme of known specificity is required (*Figure 6*). Human proinsulin can be made in *E. coli* cells, but it is not known whether it can subsequently be processed to insulin *in vitro*.[41]

Second, one can manipulate the gene for the protein in question so that an AUG (methionine) codon on the mRNA which specifies initiation of bacterial protein synthesis now immediately precedes the mRNA codons for the *mature* (rather than the prepro or pro) form of the polypeptide. The DNA splicing methodology involved here is rather complex, but is well within the range of current recombinant DNA technology. The cloned gene (or cDNA) for the protein is cut with a specific restriction enzyme at a region just inside the coding sequence for the NH_2-terminus of the mature form of the protein. It is then necessary to synthesize chemically a DNA segment which contains, in addition to the signals for initiation of bacterial RNA synthesis and binding of bacterial ribosomes to the mRNA, the AUG 'start' codon and the balance of the NH_2-terminal sequence of the protein in question. This DNA is enzymatically spliced to the fragment of cloned DNA which contains the balance of the protein, and the reconstructed chimeric

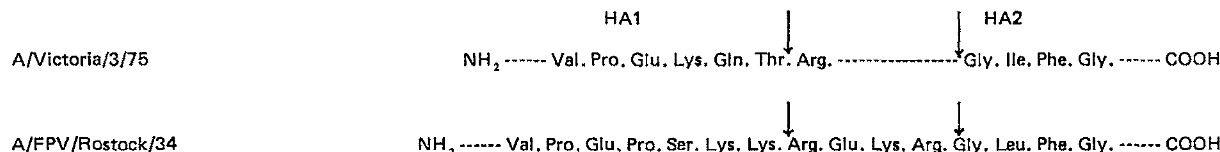

Figure 8 Proteolytic processing of the influenza HA_0 surface glycoprotein. The partial sequence of HA_0 from two influenza strains is presented. In both cases, two proteolytic cuts are required to generate the mature HA_1 and HA_2 glycoproteins, but the number of amino acids removed is different. These cleavages are essential for infectivity of the virus[62]

Case 1:05-cv-12237-WGY Document 578-21 Filed 06/29/2007 Page 8 of 13

DNA is recloned in a suitable bacterial cell. The protein produced should be the mature form of the mammalian polypeptide, although there could be an extra NH_2-terminal methionine residue if the initiator methionine is not removed. This approach has been employed for the production of human growth hormone.[65] A related method was used for human fibroblast interferon. Here, methionine is the natural NH_2-terminal residue of the mature polypeptide, and a DNA 'gene' was constructed in which the AUG codon for this methionine residue, rather than the normally employed AUG at the beginning of the 'pre' sequence, was used to initiate interferon synthesis.[42]

In a different type of cloning strategy, the coding sequence (cDNA) of the mammalian protein is inserted into an internal position in a bacterial gene so that the sequences are in phase with the bacterial coding sequences. This results in the synthesis of a hybrid bacterial–mammalian protein, from which the mammalian protein has to be released by chemical or enzymatic means. An example is that of human somatostatin, which was synthesized as a precursor containing, at its NH_2-terminus, a fragment of *E. coli* β-galactosidase. Fortunately, there are no methionine residues in somatostatin, and the construction of the cloned gene was such that a single methionine residue separated the β-galactosidase and somatostatin segments of the chimeric precursor protein. It was possible chemically to cleave off the somatostatin, at the methionine residue, using CNBr, but this cannot be said to be a general procedure.[66] Related procedures have been used to produce human insulin[67] and murine β-endorphin[68] from such hybrid polyproteins.

Finally, attention should be given to the use of mammalian cells as systems for production of some peptide hormones or other secretory proteins, since most of the relevant proteolytic (and glycosyl transferase) enzymes are present in most lines of cultured cells. The economical use of such systems awaits the development of new cloning methodologies which will permit the production of large amounts of a particular protein whose gene has been cloned and introduced into the cell. Examples of this approach are given in refs 69–72.

Formation of S–S disulphide bonds

Disulphide bonding between two cysteine resides is one of the most important stabilizing forces in the secondary structure of a polypeptide. In proteins which contain more than two cysteine residues, formation of the proper arrangement of S–S bonds is essential for normal structure and enzymatic or homonal action. In higher cells, S–S bonds are generally confined to secreted proteins and certain membrane proteins. Possibily, because of a greater reducing potential, cytoplasmic proteins in higher cells generally do not utilize the more oxidized ($2SH \rightarrow S{-}S + 2H$) S–S bond as a stabilizing force. This point is important for the genetic engineer who wishes to synthesize such a secreted protein in bacterial cells, as, in most cases studied thus far, the mammalian proteins remain in the bacterial cytoplasm. It is not known whether the right S–S bonds could form under these conditions.

In animal cells it is believed that the formation of Cys–Cys bonds occurs while the polypeptide is still growing on the ribosome (*Figure 9*).[73] In the case of the secreted immunoglobin light chain polypeptide, which is portrayed in *Figure 9*, the S–S bonds are formed on the lumenal surface of the ER, and bond formation occurs sequentially, in that the first cysteine pairs with the second, and the third with the fourth. This sequential ordering does not occur in all S–S bonded proteins.

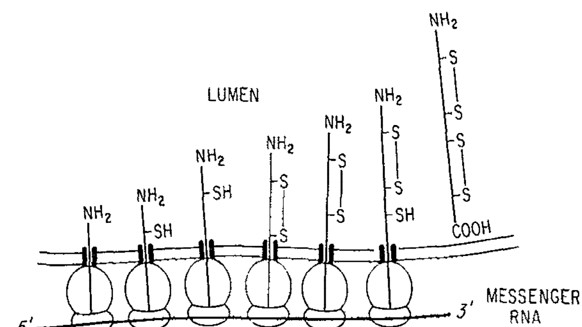

Figure 9 Formation of disulphide bonds on nascent secreted immunoglobin light chain polypeptides. Note in particular that the cystine bonds form on the lumen side of the ER and that, in this molecule, they form sequentially while the polypeptide is growing. Redrawn from ref 73

Will bacterial cells form the proper S–S bonds of mammalian proteins, such as insulin or proinsulin? We do not know, as there are little relevant published data. There are numerous examples of S–S bonds of mature proteins undergoing reduction (forming S–H cysteine) and then reoxidization, with the protein achieving its original, mature configuration and activity.[74,75] It is not known whether this can be achieved on a commercial scale. Questions of the purity of the reformed oxidized protein are very real if the material, e.g. as insulin, is to be used in humans or other animals. Abnormally folded proteins may well elicit an immune response which could restrict the long-term use of the material.

Protein glycosylation

Many secreted proteins of industrial or pharmaceutical interest are glycoproteins. This class includes serum proteins (certain complement proteins, transferrin); hormones (fibroblast and immune interferon, possibly leucocyte interferon), and virus surface glycoproteins (influenza, measles and mumps viruses). Bacterial cells, as noted above, contain little if any protein-linked saccharide residues. In yeast cells and in mammals, sugar residues are commonly linked to three different amino acid residues; they are classified as *O*-linked (serine, threonine) and *N*-linked (asparagine). The structures of *N*- and *O*-linked oligosaccharides are very different, and different sugar residues are usually found in each type. Oligosaccharides which are *O*-linked are generally shorter, often one to three sugar residues, and more variable than the *N*-linked structures. Unfortunately, little is known of the function of these *O*-linked sugars, or of the details of their biogenesis (see refs 76 and 77).

Much recent interest has focused on the 'complex', galactose and sialic acid-containing asparagine-linked oligosaccharides. The structures of most 'complex' *N*-linked oligosaccharides of membrane and secreted glycoproteins are quite similar;[76,78] an example, the structure of the two identical *N*-linked oligosaccharides on the VSV G protein, is shown in *Figure 10b*. Differences among different proteins reflect primarily the number of branches (two or three), the number of sialic acid residues (zero to three), and the chemical nature of their

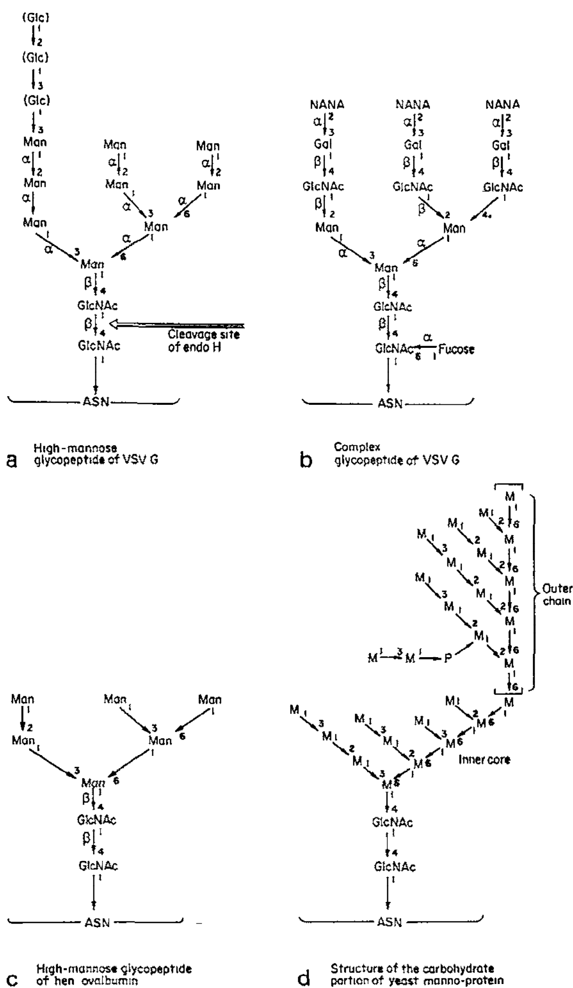

Figure 10 Structures of asparagine-linked oligosaccharides. (a) Structure of the two high mannose carbohydrate chains found on the microsomal form of VSV G protein (drawn from refs 76 and 15); including the site of cleavage by endoglycosidase H. (b) Structure of the asparagine-linked oligosaccharide found on the VSV glycoprotein.[87] (c) Structure of one of the related high-mannose glycopeptides of hen ovalbumin.[80] (d) Structure of the carbohydrate portion of yeast mannoprotein[99]

linkage to galactose.[76,78] The oligosaccharides on the major membrane protein on human erythrocytes contain variable numbers of the repeating disaccharide (Gal-GlcNAc) penultimate to the sialic acid residues, but are otherwise similar to that on VSV G.[79] The structure of the *N*-linked 'high mannose' oligosaccharide on certain glycoproteins is different from that of complex oligosaccharides (*Figures 10b* and *10c*), in that they lack fucose, sialic acid, galactose and the three peripheral GlcNAc residues. Rather, they contain six to nine Man residues, as well as two GlcNAc residues adjacent to the Asn residue.[76,78] As an example, *Figure 10c* depicts one of several related 'high mannose' oligosaccharides on hen ovalbumin.[80]

Despite this difference in structure, it is believed that both the 'complex' and 'high mannose' oligosaccharides are derived from a common precursor. The biosynthesis of these asparagine-linked oligosaccharides is a multistep process which involves both the addition and removal of specific saccharide residues (*Figure 11*).[81,82] A branched oligosaccharide precursor containing three glucose, nine mannose and two *N*-acetylglucosamine residues, is preformed on a polyisoprenoid lipid carrier molecule (dolichol) and is localized in the rough ER. The biogenesis of this intermediate utilizes a complex of membrane-bound enzymes, and involves a step-by-step sequential addition of GlcNAc, Man, and Glc residues to dolichol phosphate.[81,82]

This oligosaccharide chain is transferred, *en bloc*, to the nascent polypeptide. Many studies of the biogenesis of glycoproteins use viral membrane proteins as model systems. In particular, it was possible to show by synchronized *in vitro* translation studies that one of the two chains of the VSV G protein is added when the nascent chain is about one-third completed, the other when it is about 70% complete.[84] The *N*-linked oligosaccharides are invariably found in the tripeptide sequences Asn-X-Ser/Thr. Recent work suggests that this tripeptide is the minimal substrate for the oligosaccharide–protein transferase.[82,85] The oligosaccharides invariably face the lumenal side of the ER, and it is believed that the enzyme which transfers the saccharide to the polypeptide is localized on the lumenal surface of the ER (*Figure 2*).[86]

Immediately after transfer to the polypeptide, one or two of the glucose residues are removed. Further processing of the oligosaccharide to the 'complex' form begins only 10–20 min after synthesis of the protein, presumably at the time it is transferred to the Golgi complex. In a stepwise, concerted set of reactions, the remaining glucose residue and six of the nine mannose residues are removed from the oligosaccharide, and the 'peripheral' sugar residues *N*-acetylglucosamine (three residues per chain), galactose

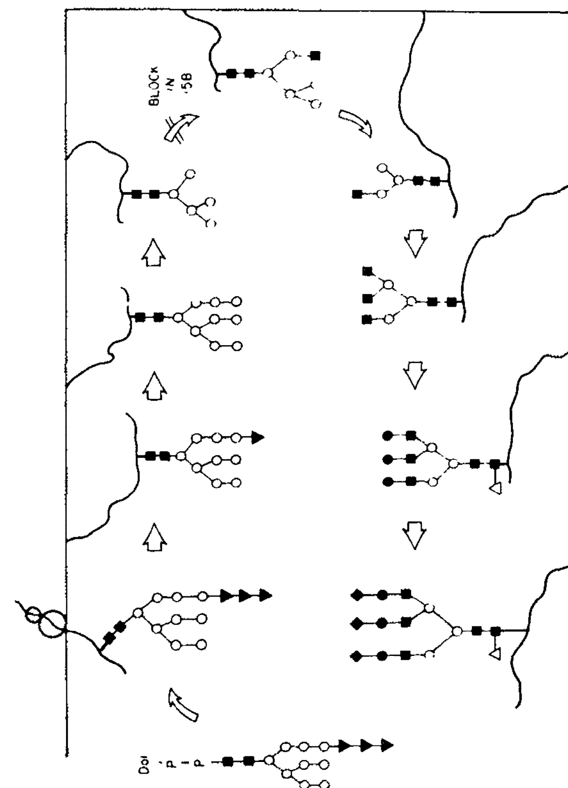

Figure 11 Proposed sequence for the synthesis of complex-type oligosaccharides (refs 81, 83). Dol = dolichol, the lipid carrier of the oligosaccharide. Symbols ■, *N*-acetylglucosamine, ○, mannose; ▲, glucose; ●, galactose; ◆, sialic acid; △, fucose

(three residues), sialic acid (one to three residues per chain) and fucose (one residue) are added (*Figures 10* and *11*).[81,83,87–92] Oligosaccharide processing is completed ~10 min before the protein reaches the cell surface[93] or is secreted.[94]

Note that the 'high mannose' oligosaccharides which are found on the mature form of certain glycoproteins resemble an intermediate ($Man_9GlcNAc_2$) in processing of the 'complex' oligosaccharides (*Figures 10a, 10c* and *11*). It is believed that both classes of *N*-linked oligosaccharides do derive from this same intermediate, the difference being the action of an α-mannosidase on the ($Man_{6-9}GlcNAc_2$) oligosaccharides.[87] Since both 'complex' and 'high mannose' oligosaccharides are occasionally found on the same protein molecule, the differences in processing may reflect the accessibility of a particular sugar side chain to a Golgi mannosidase.

Yeast cells are increasingly used as cloning hosts for the synthesis of mammalian (and other) proteins. They do contain *N*-linked oligosaccharides, but it is important to emphasize that the structure of these is very different than those on mammalian glycoproteins (*Figure 10d*). Linked to the common 'core' disaccharide GlcNAc-GlcNAc-Asn are 20–40 Man residues, in a highly specific configuration.[95,96] Interestingly, the oligosaccharide which is transferred initially to the polypeptide Asn residue is identical to that found in mammalian cells – $Glc_3Man_9GlcNAc_2$.[97] Clearly, the difference in structure between the multiple types of mature oligosaccharides – the yeast mannans and the mammalian 'complex' and 'high mannose' oligosaccharides – resides in the later processing steps. The yeast *Saccharomyces cerevesie* contains several specific GDP-mannose $\alpha1\to6$, $\alpha1\to3$, and $\alpha1\to2$ glycosyl transferases which are involved in mannan synthesis.[98–100] Viable mutants exist in several of these enzymes, a result indicating that many, but probably not all, of the peripheral mannose residues are not essential for cell viability.[99–100]

What is the function of these *N*-linked oligosaccharides? In many, but not all cases, they appear to be required for secretion of secretory proteins, or for maturation to the cell surface of membrane glycoproteins. Many of these studies utilize tunicamycin, an antibiotic which blocks the first stage in formation of the oligosaccharide–lipid donor; in its presence, the polypeptide is synthesized but it contains no *N*-linked sugar chains. Secretion of some but not all proteins will take place even in the presence of tunicamycin.[82] For instance, the rate and extent of secretion of glycosylated and unglycosylated fibronectin by fibroblasts[101] or transferrin by a cultured line of rat hepatoma cells is the same.[94] Secretion of some classes of immunoglobulins, but not others, will occur if addition of the asparagine-linked oligosaccharide is blocked,[102] as will secretion of immune interferon.[102a] Unglycosylated HLA-A (histocompatability antigen) protein matures to the cell surface at the same rate, and to the same extent, as does the normally glycosylated form.[103] A similar result has been obtained with glycophorin.[104] However, maturation of most viral glycoproteins requires the carbohydrate moiety. Clearly, the requirement for glycosylation is an idiosyncratic property of the individual protein species. It has been suggested that addition of the high mannose oligosaccharide chain modifies the conformation of the polypeptide with respect to the membrane phospholipid;[105] in some cases, this alteration may be essential for subsequent movement.

Because, in general, non-glycosylated proteins are not secreted, it has not been possible to study the metabolic fate of, for instance, a non-glycosylated derivative of a serum protein or hormone in an experimental animal, in comparison to its normally modified form. Non-glycosylated human leucocyte interferon, produced in *E. coli* using recombinant DNA methodology, is functional as an antiviral agent both in cell culture and in monkeys, but it is not known whether the potency or the stability of the molecule is normal.[106] Human leucocyte interferon probably contains attached carbohydrates, but the exact nature of this substituent is not clear.[106a]

Oligosaccharides may act to stabilize protein structure.[107] For example, fibronectin is a cell surface glycoprotein which is involved in binding cells to their substrate. Non-glycosylated fibronectin is as active as the normal form in agglutinating sheep erythrocytes, and in mediating the attachment of cultured cells to collagen. Unglycosylated fibronectin is twice as sensitive to protease digestion as is the glycosylated form, suggesting that the oligosaccharide chains stabilize the molecule against proteolytic digestion.[108] Vesicular Stomatitis Virus particles which contain unglycosylated surface viral glycoprotein are as infectious as normal virions; other properties of these particles, such as antigenicity or stability, were not tested.[109]

Some proteins in nature are only partially glycosylated. About 75% of bovine pancreatic ribonuclease contains no attached carbohydrate (RNase A) while ~25% contains one Asn-linked oligosaccharide (RNase B). Both enzymes appear equally active, although their differential stability in the intestinal lumen is not known.

While the total absence of *N*-linked sugars on a secreted protein may (or may not) affect its physiological properties, recent work makes it clear that incompletely glycosylated serum proteins are physiologically unstable. Ashwell and coworkers showed that serum glycoproteins from which the terminal sialic acid residues have been removed are rapidly (half-life <20 min) removed from the circulation. Enzymic readdition of the sialic acid residues restores normal stability (half-life of several days) (*Figure 12*). Rapid clearance of asialoglycoproteins is due to a surface receptor protein on hepatocytes which binds to, and directs internalization and destruction of, proteins which contain exposed galactose residues.[110,111] Other receptors, localized in other cells, particularly phagocytes and macrophages, specifically bind glycoproteins with terminal GlcNAc or mannose residues.[111–114] These incompletely glycosylated proteins are also unstable in the circulating blood.

How, then, is one to synthesize large amounts of glycoproteins by recombinant DNA techniques? Clearly, the ideal situation would be if the non-glycosylated protein, synthesized in bacteria, is completely normal with respect to structure, activity, stability and immunological reactivity. There are insufficient data to predict whether this hope will be realized for any particular glycoprotein. Yeast would appear to be a very poor host cell for the production of mammalian glycoproteins, if these are to be introduced into animals or humans. The presence of a mannan side chain (*Figure 10*) would certainly induce an immunological reaction by the animal against the protein. Also, the exposed mannose residues might cause a rapid clearance of the protein from the circulation by binding to one of the lectin-like receptors present on hepatocytes or phagocytes.[112–114] In principle, it should be possible enzymatically to convert the yeast mannan oligosaccharide to one

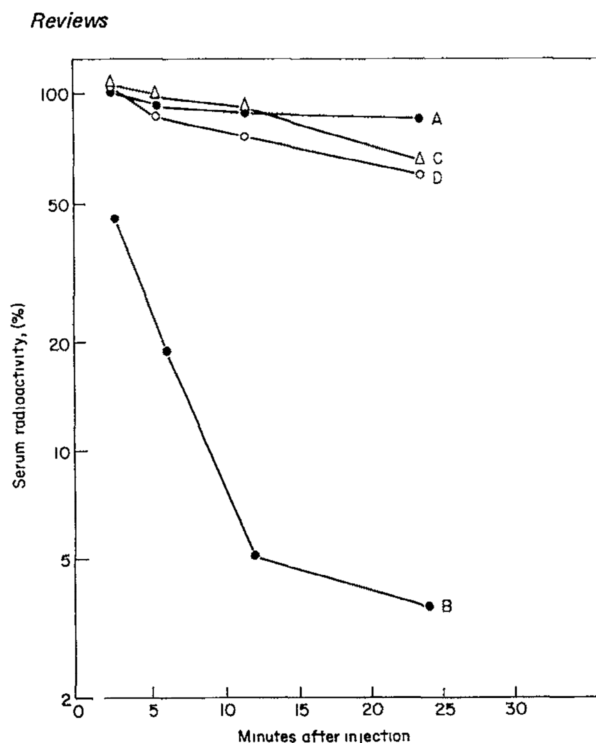

Figure 12 Disappearance from serum of [^{64}Cu] labelled native and modified rabbit ceruloplasmin injected into rabbits. Each point is the average value obtained from two animals. A, ceruloplasmin; B, asialoceruloplasmin (i.e exposed galactose); C, oxidized ceruloplasmin; D, asialoagalactoceruloplasmin, i.e. both galactose and sialic acid removed, exposing GlcNAc residues. Redrawn from ref 110

resembling the 'complex' oligosaccharide found on mammalian glycoproteins. Most of the requisite mannosidases (for removal of the 'excess' Man residues) and GlcNAc-, Gal, and sialyl-transferases have been purified, and are active in soluble form; activity is often low when an intact glycoprotein is tested as a substrate.[77] Yeast mutants exist in which some, but not all, of the outer mannose residues are missing.[99,100] All of the 'inner-core-mannose' residues apper to be present normally in these strains. Thus, to use even these mutant forms as a substrate for synthesis of complex oligosaccharides, considerable and controlled α-mannosidase degradation would be an essential prerequisite. The problem, however, with the enzymic approaches is that the glycosyl transferases must modify 100% of the saccharide chains; as noted above, chains with exposed GlcNAc or Gal residues may be unstable physiologically.

If the 'complex' *N*-linked oligosaccharide is essential for function or stability of a protein, there may be no real alternative at present to the use of cultured mammalian cells.

Conclusions

Recombinant DNA technology now permits the synthesis of virtually any polypeptide in yeast or bacterial cells, in relatively high amounts. Whether these proteins will be useful for industrial or pharmaceutical applications without further modifications is not yet clear. Most proteins, secreted proteins in particular, are extensively modified after their synthesis by proteolytic cleavages, S–S bond formation, and glycosylation. The roles of each of these modifications in the structure, function or stability of any particular protein must be determined directly as it is not yet possible to make any generalizations or predictions concerning the physiological importance of these post-translational alterations of any specific glycoprotein or secreted protein. While it may be possible to reconstruct, on bacterially engineered proteins, the proper proteolytic cleavages and S–S bonds, it is unlikely that this can be achieved for the complex set of glycosylation reactions.

Much basic research in this area needs to be done in order to assess the ultimate applicability of these engineered proteins. In particular, many of the requisite proteolytic enzymes, glycosidases and glycosyl transferases have not been purified, and we cannot estimate when it will be feasible to modify proteins made in yeast or bacterial cells with isolated enzymes *in vitro*. Ultimately, this may be the best approach to the synthesis of rare peptide hormones or growth factors which need to be made in relatively small amounts.

Acknowledgements

Research in the authors' laboratory is supported by grants AI-08814, GM27989, and GM29897 from the NIH, and NSF grant 79-00839PCM. The author thanks Miriam Boucher for her patience and expertise in preparing the manuscript, and Drs Asher Zilberstein, Anthony Sinsky and Rex Chisholm for their helpful comments.

References

1 Snapa, R. and Sutherland, B. *Biochemistry* 1980, 19, 4201–4208
1a Krebs, E. G. and Beavo, J. A. *Annu Rev Biochem* 1979, 48, 923–959
2 Chock, P. B., Rhee, S. G. and Stadtman, E. R. *Annu Rev. Biochem.* 1980, 49, 813–843
3 Nimmo, H. G. and Cohen, P *Adv. Cyclic Nucleotide Res* 1977, 8, 145–266
4 Hunt, T. *Miami Winter Symp.* 1979, 16, 321–346
5 Cohen, P. *Curr Top Cell. Regul.* 1978, 14, 117–196
6 Ochoa, S. and de Haro, C. *Annu. Rev Biochem* 1979, 48, 549–580
7 Fessler, J. H. and Fessler, L. I. *Annu Rev Biochem* 1978, 47, 129–162
8 Adams, E. and Frank, L. *Annu. Rev. Biochem* 1980, 49, 1005–1062
9 Stenflo, J. and Sutti, J. W. *Annu Rev. Biochem.* 1977, 46, 157–172
10 Hayashi, O. and Veda, K. *Annu Rev. Biochem.* 1977, 46, 95–116
11 Porter, R. R. and Reid, K. B. M. *Adv Protein Chem* 1979, 33, 1–71
12 Barrett, A. J. *Proteinases of Mammalian Cells and Tissues* North Holland, Amsterdam, 1976
13 Holtzman, E. *Lysozomes. a Survey* Springer-Verlag, 1975
14 Davie, E. W., Fujikawa, K., Kurachi, K. and Kisiel, W. *Adv Enzymol.* 1979, 48, 277–318
15 De Robertis, E. D. P. and De Robertis, E. M. F *Cell and Molecular Biology* Saunders, Philadelphia, 7th ed., 1980, pp. 206–248
16 Palade, G. *Science* 1975, 189, 347–358
17 Lodish, H. F., Braell, W. A., Schwartz, A. L., Strous, G. J. A. M. and Zilberstein, A. in *Int. Rev. Cytol.* 1981, in press
18 Korant, B. D., Chow, N. L., Lively, M. O. and Powers, J. C. in *Precursor Processing in the Biosynthesis of Proteins* (Zimmerman, M., Mumford, R. A. and Steiner, D. F , eds); *Ann. NY Acad. Sci.* 1980, 343, 304–318
19 Pelham, H. *Eur. J. Biochem.* 1978, 85, 457–462
20 Jacobson, M., Asso, J. and Baltimore, D. *J. Mol. Biol.* 1970, 49, 657–669

21 Korant, B. in *Molecular Basis of Biological Degradative Processes* (Berlin, R., ed.) Academic Press, New York, 1978, pp. 171–224
22 Von der Helm, K. *Proc. Natl. Acad. Sci. USA* 1977, **74**, 911–915
23 Dittman, K. J. and Moelling, K. *J. Virol.* 1978, **28**, 106–118
24 Yoshinaka, Y., Ishigame, K., Ohno, T., Kakeyama, S., Shibata, K. and Luftig, R. B. *Virology* 1980, **100**, 130–140
25 Lawrence, C. and Thach, R. *J. Virol* 1975, **15**, 918–928
26 Blobel, G. and Dobberstein, B. *J. Cell Biol.* 1975, **67**, 835–851
27 Blobel, G. and Dobberstein, B. *J. Cell Biol.* 1975, **67**, 852–862
28 Leader, D. P. *Trends in Biochemical Sciences* 1979, **4**, 205–208
29 Davis, B. D. and Tai, P-C. *Nature (London)* 1980, **283**, 433–438
30 Schechter, I., Wolf, O., Kantor, F., Schechter, B. and Burstein, Y. in *Precursor Processing in the Biosynthesis of Proteins* (Zimmerman, M., Mumford, R. A. and Steiner, D. F., eds), *Ann NY Acad. Sci.* 1980, **343**, 218–231
31 Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. and Tager, H. in *Precursor Processing in the Biosynthesis of Proteins* (Zimmerman, M., Mumford, R. A. and Steiner, D. F., eds), *Ann. NY Acad. Sci.* 1980, **343**, 1–16
32 Blobel, G. *Proc. Natl. Acad. Sci. USA* 1980, **77**, 1496–1500
33 Geuze, J. J., Slot, J. W. and Tokuyasu, K. T. *J. Cell Biol.* 1979, **82**, 697–707
34 Katz, F. N., Rothman, J. E., Knipe, D. M. and Lodish, H. F. *J. Supramol. Struct* 1977, **7**, 353–370
35 Jackson, R. C. and Blobel, G. in *Precursor Processing in the Biosynthesis of Proteins* (Zimmerman, M., Mumford, R. A. and Steiner, D. F., eds) *Ann. NY Acad. Sci.* 1980, **343**, 391–404
36 DiRienzo, J. M., Nakamura, K. and Inouye, M. *Annu. Rev Biochem.* 1978, **47**, 481–532
37 Bassford, P. J., Emr, S., Silhavy, T., Beckwith, J., Bedouelle, H., Clement, J-M., Hedgpeth, J. and Hofnung, M. in *Methods and Perspectives in Cell Biology* 1981, in press
38 Milstein, C., Brownlee, G. G., Harrison, T. M. and Mathews, M. B. *Nature New Biol.* 1972, **239**, 117–120
39 Jackson, R. C. and Blobel, G. *Proc. Natl. Acad Sci. USA* 1977, **74**, 5598–5602
40 Zwizinski, C. and Wickner, W. *J. Biol. Chem.* 1980, **255**, 7973–7977
41 Talmadge, K., Kaufman, J. and Gilbert, W. *Proc. Natl. Acad. Sci. USA* 1980, **77**, 3988–3992
42 Taniguchi, T., Guarente, L., Roberts, T., Kimelman, D., Douhan, J. and Ptashne, M. *Proc. Natl. Acad. Sci. USA* 1980, **77**, 5230–5233
43 Habener, J. F. and Potts, J. T. *N. Engl. J Med.* 1978, **299**, 580–585
43a Habener, J. F. and Potts, J. T. *N. Engl. J. Med.* 1978, **299**, 635–644
44 *Precursor Processing in the Biosynthesis of Proteins* (Zimmerman, M., Mumford, R. A. and Steiner, D. F., eds); *Ann NY Acad. Sci.* 1980, **343**
45 Steiner, D. F, Kemmler, W., Tager, H. S. and Peterson, J. D. *Fed. Proc.* 1974, **33**, 2105–2115
46 Habener, J. F., Chang, H. T and Potts, J. T. *Biochemistry* 1977, **16**, 3910–3917
47 Potts, J. T., Kronenberg, H. M., Habener, J. F. and Rich, A. in *Precursor Processing in the Biosynthesis of Proteins* (Zimmerman, M., Mumford, R. A. and Steiner, D. F., eds); *Ann. NY Acad Sci.* 1980, **343**, 38–54
48 Quinn, P. S. and Judah, J. D. *Biochem. J.* 1978, **172**, 301–309
49 Eipper, B. A., Mains, R. E. and Guenzi, D. *J. Biol. Chem.* 1976, **251**, 1421–1426
50 Steiner, D. F., Kemmler, W., Clark, J. L., Oyer, P. E. and Rubenstein, A. H. The Biosynthesis of Insulin in *Handbook of Physiology – Endocrinology I* (Steiner, D. F. and Freinkel, N., eds) Williams and Wilkins, Baltimore, 1972, pp. 175–198
51 Steiner, D. F. *Diabetes* 1977, **26**, 322–340
52 Hamilton, J. W., Niall, H. D., Jacobs, J. W., Keutmann, H. J., Potts, J. T. and Cohen, D. V. *Proc. Natl. Acad. Sci. USA* 1974, **71**, 653–658
53 Russell, J. H. and Geller, D. M. *J Biol. Chem.* 1975, **250**, 3409–3413
54 Noyes, B. E., Stein, M. and Agarwal, K. L. *Proc. Natl Acad. Sci. USA* 1979, **76**, 1770–1774
55 Kemmler, W., Peterson, J. D and Steiner, D. F. *J Biol. Chem.* 1971, **246**, 6786–6791
56 Mains, R. E., Eipper, B. A and Ling, N. *Proc. Natl. Acad Sci USA* 1977, **74**, 3014–3018
57 Herbert, E., Baduf, M., Phillips, M., Rosa, P Policastro, P., Oates, E., Roberts, J., Seidah, N. and Chretien, M. in *Precursor Processing in the Biosynthesis of Proteins* (Zimmerman, M., Mumford, R. A. and Steiner, D. F., eds); *Ann NY Acad Sci.* 1980, **343**, 79–93
58 Mains, R. E. and Eipper, B in *Precursor Processing in the Biosynthesis of Proteins* (Zimmerman, M , Mumford, R A and Steiner, D. F., eds), *Ann. NY Acad. Sci.* 1980, **343**, 94–108
59 Lazarowitiz, S. G and Choppin, P. W. *Virology* 1975, **68**, 440–454
60 Klenk, H-D , Rott, R. and Orlich, M. *J. Gen Virol.* 1977, **36**, 151–161
61 Bosch, F. X., Orlich, M., Klenk, H-D. and Rott, R. *Virology* 1979, **95**, 197–207
62 Min Jou, W , Verhueyen, M., Devos, R., Saman, E., Fang, R , Hoylebroeck, D., Fiers, W., Threlfall, G., Barker, C., Carey, N. and Emtage, S. *Cell* 1980, **19**, 683–696
63 Scheid, A. and Choppin, P. W. *Virology* 1974, **57**, 475–490
64 Bornstein, P. and Sage, H. *Annu Rev. Biochem.* 1980, **49**, 957–1003
65 Goeddel, D., Heyneker, H., Hozumi, T., Arentzen, R., Itakura, K., Yansura, D., Ross, M., Miozzari, G., Crea, R. and Seeburg, P. H. *Nature (London)* 1979, **281**, 544–549
66 Itakura, K., Hirose, T., Crea, R., Riggs, A., Heyneker, H., Bolivar, F. and Boyer, H. *Science* 1977, **198**, 1056–1063
67 Goeddel, D. *et al Proc. Natl. Acad. Sci USA* 1979, **76**, 106–110
68 Schine, J., Fettes, I., Lan, N., Roberts, J. and Baxter, J. *Nature (London)* 1980, **285**, 456–461
69 Mulligan, R. C. and Berg, P. *Science* 1980, **209**, 1422–1428
70 Hamer, D. H. and Leder, P. *Cell* 1979, **18**, 1229–1238
71 Wigler, M., Pellicer, A., Silverstein, S. and Axel, R. *Cell* 1978, **14**, 725–731
72 Pellicer, A., Wagner, E. F., Kareh, A., Dewey, M. J., Reuser, A., Silverstein, S., Axel, R. and Mintz, B. *Proc. Natl. Acad. Sci. USA* 1980, **77**, 2098–2102
73 Bergman, L. W. and Kuehl, W. M. *J. Biol. Chem.* 1979, **254**, 8869–8876
74 Baldwin, R. L. *Annu. Rev. Biochem.* 1975, **44**, 453–475
75 Anfinsen, C. B. and Scheraga, H. A. *Adv Protein Chem.* 1975, **29**, 205–299
76 Kornfeld, R. and Kornfeld, S. in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, E. J., ed.) Plenum Press, New York, 1980 pp. 1–34
77 Schachter, H. and Roseman, S. in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, E. J., ed.) Plenum Press, New York, 1980, pp. 85–160
78 Kornfeld, R. and Kornfeld, S. *Annu. Rev. Biochem.* 1976, **45**, 217–237
79 Tsuji, T., Irimura, T. and Osawa, T. *Biochem. J* 1980, **187**, 677–686
80 Tai, T., Yamashita, K., Ito, S. and Kobata, A. *J. Biol Chem.* 1977, **252**, 6687–6693
81 Gibson, R., Kornfeld, S. and Schlesinger, S. *Trends Biochem. Sci.* 1980, **5**, 290–293
82 Struck, D. K. and Lennarz, W. J. in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, E. J., ed.) Plenum Press, New York, 1980, pp. 35–83
83 Kornfeld, S., Li, E. and Tabas, I. *J. Biol. Chem.* 1978, **253**, 7771–7778
84 Rothman, J. E. and Lodish, H. F. *Nature (London)* 1977, **269**, 775–780
85 Hart, G. W., Brew, K., Grant, G. A., Bradshaw, R. A. and Lennarz, W. J. *J. Biol. Chem.* 1979, **254**, 9747–9753
86 Hanover, J. A. and Lennarz, W. J. *J. Biol. Chem.* 1980, **255**, 3600–3604
87 Robbins, P. W., Hubbard, S. C., Turco, S. J. and Wirth, D. F. *Cell* 1977, **12**, 893–900
88 Hunt, L. A., Etchison, J. R. and Summers, D. F. *Proc. Natl. Acad. Sci. USA* 1978, **75**, 754–758
89 Reading, C. L., Penhoet, E. E. and Ballou, C. E. *J. Biol. Chem.* 1978, **253**, 5600–5612
90 Turco, S. J. and Robbins, P. W. *J. Biol. Chem.* 1979, **254**, 4560–4567

91 Tabas, I. and Kornfeld, S. *J. Biol. Chem.* 1979, **254**, 11655–11663
92 Bretz, R., Bretz, H. and Palade, G. E. *J. Cell Biol.* 1980, **84**, 87–101
93 Knipe, D., Lodish, H. F. and Baltimore, D. *J. Virol.* 1977, **21**, 1121–1127
94 Strous, G. J. A. M. and Lodish, H. F *Cell* 1980, **22**, 709–717
95 Ballou, C. E. *Adv. Enzymol* 1974, **40**, 239–270
96 Nakajima, T. and Ballou, C. E. *J. Biol Chem* 1974, **249**, 7685–7694
97 Trimble, R., Maley, F. and Tarentino, A L. *J Biol Chem.* 1980, **255**, 10232–10238
98 Nakajima, T. and Ballou, C. E. *Proc Natl. Acad. Sci. USA* 1975, **72**, 3912–3916
99 Lehle, L., Cohen, R. E. and Ballou, C. E. *J. Biol. Chem.* 1979, **254**, 12209–12218
100 Ballou, L., Cohen, R. E. and Ballou, C. E. *J. Biol. Chem.* 1980, **255**, 5986–5991
101 Olden, K., Pratt, R. M. and Yamada, K. *Cell* 1978, **13**, 461–473
102 Hickman, S. and Kornfeld, S. *J. Immunol.* 1978, **121**, 990–996
102a Mizrahi, A., O'Malley, J., Carter, W., Takatsuki, A., Tamura, G. and Solkowski, E. *J. Biol. Chem* 1978, **253**, 7612–7615
103 Owen, M. J., Kissonerghis, A-M. and Lodish, H. F. *J. Biol. Chem.* 1980, **255**, 9678–9684
104 Gahmberg, C. G., Jokinen, J., Darhi, D. D. and Anderson, L. C. *J. Biol. Chem.* 1980, **255**, 2169–2175
105 Gibson, R., Schlesinger, S. and Kornfeld, S *J. Biol. Chem* 1979, **254**, 3600–3607
106 Goeddel, D. V. *et al. Nature (London)* 1980, **287**, 411–416
106a Otto, M., Sedmak, J. and Grossberg, S. *J. Virol.* 1980, **35**, 390–399
107 Pazur, J., Knull, H and Simpson, D. *Biochem. Biophys Res. Commun.* 1970, **40**, 110–116
108 Olden, K., Pratt, R. M. and Yamada, K. M. *Proc. Natl Acad. Sci. USA* 1979, **76**, 3343–3347
109 Gibson, R., Leavitt, R., Kornfeld, S. and Schlesinger, S. *Cell* 1978, **13**, 671–679
110 Ashwell, G. and Morell, A. G. *Adv. Enzymol. Relat Areas Mol. Biol* 1974, **41**, 99–128
111 Neufeld, E. F. and Ashwell, G. in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, E. J., ed.) Plenum Press, New York, 1980, pp. 241–266
112 Kawasaki, T., Etoh, R. and Yamashina, I. *Biochem. Biophys. Res. Commun.* 1978, **81**, 1014–1018
113 Achord, D. T., Brot, F E. and Sly, W. S. *Biochem. Biophys. Res. Commun.* 1977, **77**, 410–418
114 Schlesinger, P. H., Doebker, T. W., Mandell, B., White, R., de Schriyver, C., Rodman, J., Miller, M. and Stahl, P. *Biochem J.* 1978, **176**, 103–109
115 Zilberstein, A., Snider, M., Porter, M. and Lodish, H. F. *Cell* 1980, **21**, 417–427