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
Enzyme and Microbial Technology

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

EMTEUZ 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-igninolytic systems
T. Kent Kirk and Hou-min Chang

197 Metabolism of carbon monoxide
Robert L. Utton

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
Veronique Larreta Gardé, Brigitte Thomasset and Jean-Noel Bartotin

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. Belkov

225 Glyceraldehyde-3-phosphate synthase and other compounds as stabilizers of alcohol dehydrogenase from yeast
M. J. Brougham and D. R. Johnson

229 Effect of the nature of proteins on their coupling to different epoxide-containing supports

233 Glycolytic pathway as an ATP generation system and its application to the production of glutathione and NADP
Kunesu Murai, Keiko Tani, Iyoji Kato and Ichiro Chibata

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

246 Immobilization of Bacillus subtilis α-amylase on zincia-coated alkylamine glass with gluteraldehyde
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
G. K. Jin, H. L. Chiang and S. S. Weng

Short paper

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

Literature survey

260 Patent reports

267 Industrial views

274 Conference reports

278 Books reviews

279 Publisher’s announcement

Calendar

Managing Editor
Stella Dutton

Assistant Editor
Monique Heald

Sub-Editor
Diane Coyan
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 enzymatic reactions, proteolytic cleavages, the peptidase 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 biochemist and the synthetic chemist. 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, 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 post-translational 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.

Cova lent 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. An understanding of these covalent 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, 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, hydroxylation 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 sulphonation, hydroxylation of lysine or proline residues (in collagen), vitamin K-dependent carboxylation, methylation, acetylation, and ADP-ribosylation.

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 lysosomal enzymes — have been purified to date. Importantly, these activities appear to be localized to specific subcellular regions or membrane-limited
derived from the proteolytic cleavage of one of the viral protein precursors, probably the one indicated 'replicase (P75).'

The precursor of the RNA tumor virus (retrovirus) capsid proteins is processed in much the same way as the poliovirus capsid precursor, and it also appears that a virus-specific protease might be involved. Although these virus-specifc proteases have not been purified to homogeneity, further studies of their substrate specificity and mechanism might provide the basis for design of specific anti-viral compounds.

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). 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).

Extensive proteolytic processing generally occurs in each of the above organelles.

The biogenesis of many surface membrane glycoproteins, such as the VSV glycoprotein, is the histocompatibility proteins HLA-A and HLA-B, follows an analogous pattern. The formation of these proteins is initiated with the synthesis of a pre-protein, which is subsequently processed by proteolytic cleavage (Figure 3b). In addition, these proteins remain anchored in the phospholipid membranes of the various organelles, rather than being released into the extracellular lumen (Figure 2a).

In the case of most, but not all, secretory and membrane glycoproteins in which at least the NH2-terminal segment, if not the whole protein, crosses the ER membrane, 15–26 NH2-terminal amino acid residues are removed proteolytically while the polypeptide is still growing on the ribosome. 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 proteins. The result proteins contained the extra 'pre' residues. These 'pre' sequences are often termed '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 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, 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, 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.
acids, which are always at the NH$_2$-terminus, 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 'propeptide 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 disulfide bonds, and remain attached when the C peptide is removed.\textsuperscript{31}

Examination of the amino acid sequences surrounding the cleavage sites in preproteins 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 proglucagon, a single endopeptidase with trypsin-like specificity is sufficient to generate the mature protein. In others, such as proinsulin

\textit{E. coli} can correctly process pre-proinsulin (and will secrete proinsulin),\textsuperscript{41} but other proteins such as pre-fibroblast (human) interferon are apparently not processed in this way.\textsuperscript{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 'pro' sequence is the only known proteolytic cleavage: the 'pro' form is converted directly into the mature protein.\textsuperscript{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).\textsuperscript{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.\textsuperscript{46,49,50} In contrast to the 'pro' amino
Post-translational modification of proteins: Harvey F. Lodish

Pre-procortismin
Met. Lys, 1rp, Val, Thr, Phe, Leu, Leu, Leu, Phe, Ile, Ser, Gly, Ser, Ala, Phe, Ser, Arg

Pre-βG light chain
Met. Asp, Met, Arg, Ala, Pro, Ala, Gin, Ile, Phe, Gly, Phe, Leu, Leu, Leu, Phe, Pro, Gly, Thr, Arg, Cys, Asp

Pre-lysozyme
Met. Arg, Ser, Leu, Leu, Ile, Phe, Gly, Leu, Val, Cys, Phe, Leu, Phe, Leu, Ala, Ala, Ile, Gly, Lys

Pre-prolactin
Met, Asn, Ser, Gin, Val, Ser, Ala, Arg, Lys, Ala, Gly, Thr, Leu, Leu, Leu, Met, Met, Ser, Asn, Ala, Leu, Phe, Cys, Gin, Asn, Val, Gin, Thr, Leu

Pre-procalcitonin (E. coli)
Met, Ser, Ile, Gin, His, Phe, Arg, Val, Ala, Leu, Ile, Pro, Phe, Ala, Phe, Cys, Ile, Pro, Val, Phe, Ala, His

Pre-prolactin (E. coli)
Met, Lys, Asp, Thr, Lys, Leu, Val, Leu, Gly, Ala, Val, Ile, Leu, Gly, Ser, Thr, Thr, Leu, Leu, Ala, Gly, Cys

CLEAVAGE

Figure 1. Amino acid sequence of pre-proteins of several secretory and membrane proteins. The vertical dotted line denotes the site of cleavage of these residues, which 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. Likewise, proinsulin can be converted efficiently to insulin in vitro by the combined action of trypsin and carboxypeptidase B. 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 biosynthesis of the peptide hormones ACTH, β-lipotropin, β-endorphin,

Figure 2. Schemes for combined actions of tryptic (T) and carboxypeptidase D-like (CP) activities in the conversion of proproteins (prohormones) to the biologically active 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 82; proprolactin, ref 53; protrenalin, ref 54; proglucagon, ref 31; proinsulin, ref 50. Solid lines denote the mature form of the protein; ligase lines denote the 'pre' sequences.

Figure 3. Amino acid sequence of pre-proinsulin-2 (pre-proIPTH). Residues in hatched circles constitute the NH2-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 arrow denotes the peptide bond cleaved in removal of, first, the pre, and later, the pro extensions. Redrawn from ref 47.
Reviews

The N-terminal glycopeptide of ACTH and β-LPH is cleaved in the pituitary to form smaller peptides, which are further cleaved to form smaller fragments. These fragments are then reassembled into the intermediate lobe of the pituitary to form the mature ACTH and β-LPH.

**Figure 7** Structure of the common precursor molecule of ACTH and β-LPH. The precursor is cleaved to form smaller peptides, which are further cleaved to form smaller fragments. These fragments are then reassembled into the intermediate lobe of the pituitary to form the mature ACTH and β-LPH.

**Figure 8** Proteolytic processing of the influenza HA glycoprotein. The partial sequence of HA from two influenza strains is presented. In both cases, two proteolytic cuts are required to generate the mature HA, which is essential for infectivity of the virus.

The targets of immunological response to the infecting virus include the haemagglutinin and the neuraminidase of the virus. These glycoproteins are cleaved by proteolytic enzymes to form smaller fragments, which are recognized by the immune system.

**Figure 9** Cleavage of the influenza HA glycoprotein. The glycoproteins are cleaved by proteolytic enzymes to form smaller fragments, which are recognized by the immune system.
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 NH₂ terminal methionine residue if the initiator methionine is not removed. This approach has been employed for the production of human growth hormone. A related method was used for human fibroblast interferon. Here, methionine is the natural NH₂-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.

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 obtained by chemical or enzymatic means. An example is of human somatostatin, which was synthesized as a precursor containing, at its NH₂-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 could not be said to be a general procedure. Related procedures have been used to produce human insulin and murine β-endorphins from such hybrid polypeptides.

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 residues 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 hormonal action. In higher cells, S-S bonds are generally confined to secreted proteins and certain membrane proteins. Possibly, because of a greater reducing potential, cytoplasmic proteins in higher cells generally do not utilize the more oxidized (2SH → 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). In the case of the secreted immunoglobulin light chain polypeptide, which is portrayed in Figure 9, the S-S bonds are formed on the luminal surface of the ER, and bond formation occurs sequen-
mannose and two N-acetylglucosamine residues, is performed 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.\textsuperscript{31,32}

This oligosaccharide chain is transferred, \textit{en bloc}, to the nascent polypeptide. Many studies of the biosynthesis of glycoproteins use viral membrane proteins as model systems. In particular, it was possible to show by synchronized \textit{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.\textsuperscript{33} 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.\textsuperscript{37,38} The oligosaccharides invariably face the luminal side of the ER, and it is believed that the enzyme which transfers the saccharide to the polypeptide is localized on the luminal surface of the ER (Figure 2).\textsuperscript{39}

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-acetylgalactosamine (three residues per chain), galactose

- **Figure 10** Structures of asparagine-linked oligosaccharides. (a) Structure of the two high mannose carbohydrate chains found on the monoclonal form of VSV G protein (drawn from refs 75 and 15); including the site of cleavage by endo-
\textit{H} oligosaccharidase \textit{H}. (b) Structure of the asparagine-linked oligosaccharide found on the VSV glycoprotein.\textsuperscript{71} (c) Structure of one of the related high mannose glycopeptides of hen ovalbumin.\textsuperscript{72} (d) Structure of the carbohydrate portion of yeast mannan.\textsuperscript{73}

- **Figure 11** Proposed sequence for the synthesis of complex-type oligosaccharides (refs 31, 53). Dol = dolichol, the lipid carrier of the oligosaccharide. Symbols: $\alpha$, N-acetylgalactosamin, $\beta$, N-acetylglucosamin, $\gamma$, mannose; $\delta$, glucose; $\varepsilon$, galactose; $\zeta$, sialic acid; $\omega$, fucose

(three residues), salic acid (one to three residues per chain) and lycose (one residue) are added (Figures 10 and 11). Oligosaccharides processing is completed ~10 min before the protein reaches the cell surface or is secreted.

Note that the 'high mannose' oligosaccharides which are found on the mature form of certain glycoproteins resemble an intermediate (Man$_3$GlcNAc$_2$) 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 in the action of the α-mannosidase on the (Man$_3$GlcNAc$_2$) oligosaccharides. Since both 'complex' and 'high mannose' oligosaccharides are occasionally found on the same protein molecule, the differences in processing may reflect the necessity of a particular sugar side chain for cell function. 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 10h). Linked to the common 'core' disaccharide GlcNAc-GlcNAc-Asn are 70–40 Man residues, in a highly specific configuration. Interestingly, the oligosaccharide which is transferred initially to the polypeptide is composed of only 6–12 Man residues identified to the polypeptide Asn residue is different from that found in mammalian cells — Glc$_3$ManGlcNAc$_2$. 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 cerevisiae contains several specific GDP-mannose α1→6, α1→2, and α1→2 glycosyl transferases which are involved in mannan synthesis.

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.

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 of the cell surface of membrane glycoproteins. Many of these studies utilize tunicaivcycin, an antibiotic which blocks the first stage in formation of the oligosaccharide — lipid donor; in its presence, the polypeptide is synthesized but contains no N-linked sugar chains. Secretion of some but not all proteins will take place even in the presence of tunicamycin. For instance, the rate and extent of secretion of glycosylated and unglycosylated fibronectin by fibroblasts or transferrin by a cultured line of rat hepatoma cells is the same. Secretion of some classes of immunoglobulins, but not others, will occur if addition of the asparagine-linked oligosaccharide is blocked, as will secretion of immune interferon. Unglycosylated HLA-A (histocompatibility antigen) protein matures to the cell surface at the same rate, and to the same extent, as does the normally glycosylated form. A similar result has been obtained with glycoporphin. 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; in some cases, this alteration may be essential for subsequent movement.

Post-translational modification of proteins: Harvey F. Lodish

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. Human leucocyte interferon probably contains attached carbohydrates, but the exact nature of this substituent is still not clear.

Oligosaccharides may act to stabilize protein structure. 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 aggregating sheep erythrocytes, and 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.

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.

Some proteins in nature are only partially glycosylated. About 35% 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 milieu 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. Aswell and coworkers showed that serum glycoproteins from which the terminal sialic acid residues have been removed are rapidly (half-life < 30 min) cleared from the circulation. Enzymic readuction of the sialic acid residues restores normal stability (half-life of several days) (Figure 12). Rapid clearance of asialo-glycoproteins is due to a surface receptor protein on hepatocytes which binds to, and directs internalization and destruction of, proteins which contain exposed galactose residues.

Other receptors, localized in other cells, particularly phagocytes and macrophages, specifically bind glycoproteins with terminal GlcNAc or mannose residues. 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 mannose side chain (Figure 10f) 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. In principle, it should be possible enzymatically to convert the yeast mannans oligosaccharide to one
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 bacteriologically 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 glycosyltransferases 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, GM 27989, and GM 20897 from the NIH, and NSF grant 79-08397. The author thanks Miriam Boucher for her patience and expertise in preparing the manuscript, and Dr. Asher Zilberstein, Anthony Susky and Rex Chisholm for their helpful comments.

References

1 Snaps, R. and Sutherland, B. Biochemistry 1980, 19, 4201–4208
4 Hunt, T. and W. J. J. Biol. 1979, 10, 801–804
12 Holtzman, E. L. Sources: Springer-Verlag, 1975
15 Palade, G. Science 1975, 189, 347–358
25 Lawrence, J. M. and Thack, R. J. Biol. Chem. 1975, 250, 918 – 928
28 Leader, D. P. Trends in Biochemical Sciences 1979, 4, 205 – 208
40 Zwindelius, C. and Wicks, W. J. Biol. Chem. 1980, 255, 7973 – 7977
45 Precursor Processing in the Biosynthesis of Proteins (Zimmerman, M., Munford, R. A. and Steiner, D. F., eds.) Ann. NY Acad. Sci. 1980, 343
49 Elper, B. A., Mason, R. E. and Gos, D. J. Biol. Chem. 1976, 251, 1421 – 1426
50 Steiner, D. F., Kornell, W., Clark, J. L., Oyer, F. E. and Rubenstein, A. H. The Biosynthesis of Insulin in Handbook of Physiology – Endocrinology (Steiner, D. F. and Frerken, N., eds) Williams and Wilkins, Baltimore, 1972, pp. 175 – 196
51 Steiner, D. F. Diabetes 1977, 26, 522 – 540
52 Elper, B. A., Mason, R. E. and Gos, D. J. Biol. Chem. 1974, 250, 393 – 300

References

91 Tabas, I. and Kornfeld, S. J. Biol. Chem. 1979, 254, 11655–11663
96 Nakajima, T. and Ballou, C. E. J. Biol. Chem. 1974, 249, 7685–7694
102a Morabito, A., O'Malley, J., Carter, W., Takatsuki, A., Tamura, G. and Solokowski, E. J. Biol. Chem. 1978, 253, 7612–7615