Newly Made Proteins Zip Through the Cell

Many proteins apparently find the way to their specific destinations by following guides incorporated within their own structures

To the newly made protein, the inside of the cell must be as confusing as a strange city would be to us. It has to find its way, often through a maze of membranes, to its own particular destination. Some proteins, for example, are destined for export out of the cell. Others may end up in the external cellular membrane. Still others need to find their way to one of the assorted structures, such as the mitochondria or lysosomes, that dot the cell interior.

That proteins end up where they belong is of critical importance. Mistakes in transport can cause cells to malfunction, even to die, and the cumulative damage may prove fatal to the organism they compose. An example of a human disease caused by misdirected protein transport is the rare genetic condition called I-cell disease, in which enzymes destined for packaging in the lysosomes are instead shipped out of the cell. As a result the cell becomes packed with inclusion bodies (thus the "I" of I cell), which are actually bloated lysosomes filled with debris that should have been broken down by the missing enzymes. Patients with I-cell disease usually die in childhood.

All proteins are made on cell structures called ribosomes. "The problem then," says Günter Blobel of Rockefeller University, "is how proteins synthesized on the ribosomes are routed to the various addresses in the cell." The solution seems to be that proteins have within their own structures signals—which might be compared to postal zip codes—that direct them to their targets.

Much of the work on protein transport has focused on proteins destined for secretion from the cell. The overall path taken by these proteins is known, largely from research done by George Palade of Yale University Medical School and his many collaborators.

They found that secreted proteins are synthesized on ribosomes bound to the endoplasmic reticulum (ER), a convoluted network of intracellular membranes. From the ER the proteins are carried, apparently in small membrane-bound sacs called vesicles, to the Golgi apparatus, another membranous cell structure. They are then transported, again in vesicles, to the exterior cell membrane and released to the outside.

A number of problems have to be solved as secreted proteins make this journey. The first is how the ribosome with its growing protein chain attaches to the ER membrane.

Early in the 1970's, Blobel and David Sabatini of the New York University School of Medicine speculated that the signal for attachment might be incorporated in the structure of the growing protein chain itself. The logical location for this signal would be the amino-terminal end of the chain, which is the first end to be synthesized.

Evidence for this hypothesis—which is called the signal hypothesis—began accumulating almost immediately. For example, Cesar Milstein and George Brownlee of the MRC Laboratory of Molecular Biology in Cambridge, England, found that the smaller of the two protein chains comprising antibody molecules, when it is initially synthesized, is larger than the finished chain. The extra length appeared to be on the amino-terminal end of the protein, and these investigators speculated that the extension might play a role in attachment to the ER.

Another boost for the signal hypothesis came when Blobel and his colleagues found that all the major proteins secreted by the pancreas initially contain amino-terminal extensions. Analogous sequences, which are often called signal sequences because of their postulated function, are now turning up on a large number of additional proteins secreted by cells ranging in complexity from very simple ones, such as nonnucleated bacterial cells, to the most complex mammalian, including human, cells.

The signal sequences detected thus far range in length from 15 to 30 amino acid residues. Each has a stretch of hydrophobic (uncharged, "water-hating") residues in the middle and hydrophilic ("water-loving") residues on the ends. Even though the structures of all the sequences fit this pattern, there are considerable variations in their amino acid compositions. Some investigators have questioned whether these differences are compatible with the postulated role of the sequences in attaching to specific receptors on the ER membranes. But Blobel points out that the receptors may be recognizing the overall shape of the signals, rather than particular amino acids.

The mere existence of the amino-terminal extensions, no matter how widespread they are, does not necessarily prove that they act as signals, however. Evidence that they do comes from experiments in which Blobel's group reproduced the transport of secretory proteins into the ER in a simple test-tube system. They showed that the newly synthesized protein was transported into vesicles consisting of ER membranes, but only if the membrane preparation was added in the early stages of synthesis. During the transport, the signal peptide was removed. But if the vesicles were added later, transport did not occur, and the protein, still containing the signal sequence, remained outside the vesicles in the incubating solution. The lack of uptake could be the result of the protein folding up so that the signal sequence was shielded and could no longer attach to the membranes.

Some of the best evidence for the signal hypothesis was obtained from studies of bacterial proteins, which are synthesized with amino-terminal sequences resembling those of mammalian proteins.

The bacterium *Escherichia coli* is surrounded by two membranes separated by the "periplasmic space." Jonathan Beckwith of Harvard Medical School, Thomas Silhavy of the Frederick Cancer Research Center, Philip Bassford of the University of North Carolina, and Maxime Schwartz of the Institut Pasteur in Paris have been studying the synthesis and transport of some of the proteins destined either for one or the other of the *E. coli* membranes or for the periplasmic space. These proteins are synthesized on ribosomes bound to the inner bacterial membrane, just as proteins secreted by mammalian cells are synthesized on

0036-8075/80/0111-0164$00.50/0

ribosomes bound to the ER membranes.

In an ingenious series of experiments, the investigators showed that transport of the newly synthesized proteins to the appropriate destinations requires a signal sequence. They fused the gene coding for β-galactosidase, an enzyme normally located in the *E. coli* cytoplasm, with portions of the genes coding for the other proteins in question. The β-galactosidase, which is easy to detect, permitted the investigators to determine the ultimate location of the proteins whose synthesis was directed by the hybrid genes. They found that if the hybrid protein contained a sufficiently long segment of the amino acid end of the membrane proteins it was transported to the appropriate membrane. The amino acid segment must contain the signal sequence for proper localization, although the evidence suggests that additional amino acids may also play a role.

The results for hybrids bearing amino-terminal segments from the periplasmic protein are less clear-cut than those for the membrane proteins. The periplasmic hybrids do not actually enter the periplasmic space but instead end up in the inner bacterial membrane. Their transport is at least begun normally, however, even if it is not completed.

A major advantage of the bacterial systems is that they allow the selection of mutants defective in protein transport. Beckwith, Silhavy, Bassford, and Scott Emr, also at Frederick, have produced such mutations of the hybrid genes. The protein products of the mutants accumulate in the cytoplasm, with little, if any, being transferred to the inner or outer membranes, the destinations of the nonmutant hybrids. The investigators mapped the mutations to the portion of the gene coding for the signal sequence of the proteins.

With Maurice Hofnung and his colleagues at the Institut Pasteur, Beckwith, Silhavy, and Emr have now determined the nucleotide order in that portion of the mutant and nonmutant genes. (Genes are composed of sequences of nucleotides.) In most cases, the alternation is such that one of the hydrophobic residues of the signal sequence would be replaced by a charged amino acid in the mutant proteins. According to Beckwith, "the cytoplasmic accumulation of the mutant proteins is direct evidence for the signal hypothesis."

The ER-bound ribosomes of mammalian cells synthesize a number of other proteins in addition to those secreted by the cell. These include proteins destined for insertion into cellular membranes or for incorporation into the lysosomes,

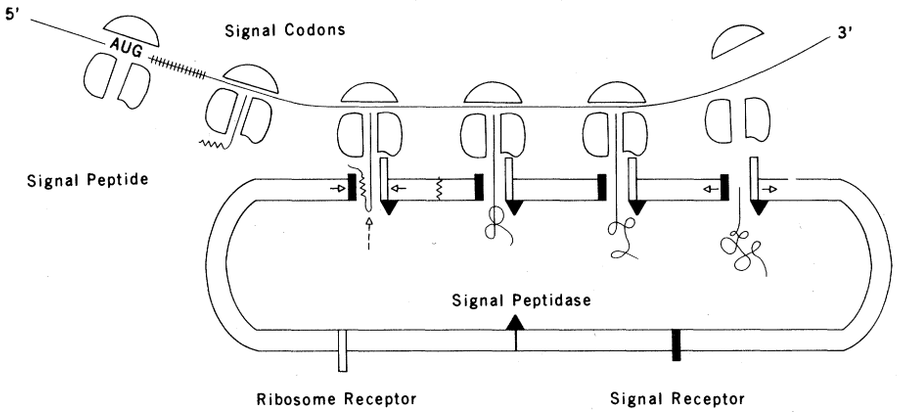

Diagram of the signal hypothesis for the transport of secreted proteins across the ER membrane. The ribosomes synthesizing a protein move along the messenger RNA specifying the amino acid sequence of the protein. (The messenger is represented by the line between 5' and 3'.) The codon AUG marks the start of the message for the protein; the hatched lines that follow AUG represent the codons for the signal sequence. As the protein grows out from the larger ribosomal subunit, the signal sequence is exposed and binds to its receptor (represented by the solid bar) on the ER membrane. There is also a receptor (open bar) for the ribosome itself. The interaction of the ribosome and growing peptide chain with the ER membrane results in the opening of a pore through which the protein is transported to the interior space of the ER. During transport, the signal sequence of most proteins is removed by an enzyme called the signal peptidase. The completed protein is eventually released by the ribosome, which then separates into its two components, the large and small ribosomal subunits. The protein ends up inside the ER. [*Source: Günter Blobel, Rockefeller University*]

small vesicles filled with 40 or so different enzymes, whose job is to break down a wide variety of large molecules.

The initial stages of the synthesis and transport of proteins destined for cell membranes are quite similar to those of secreted proteins, according to Harvey Lodish of the Massachusetts Institute of Technology and James Rothman of Stanford University. They have been studying the synthesis of certain viral proteins as models for cellular proteins.

The viruses in question shut down protein synthesis by the cells they infect and take over the cells' synthetic machinery in order to make viral components. Because the viral proteins are made in large quantities and are different from the normal cellular proteins, they are easy to detect. And because the viruses use the cells' own enzymes, the protein products presumably follow the same paths as the cellular ones.

One of the proteins studied by Lodish and Rothman is the major membrane protein of vesicular stomatitis virus (VSV). During VSV replication, this protein, designated the G protein, becomes embedded in the exterior cell membrane. The completed VSV particle is released from the cell by budding of the cell membrane, during the course of which it becomes encapsulated in a portion of membrane bearing the G protein.

The Lodish and Blobel groups have found that the G protein has a signal sequence 16 amino acids in length. During its synthesis, this protein becomes attached to the ER membrane by the signal sequence just as secreted proteins do. But the G protein, instead of being transported through the membrane, gets stuck in it for reasons as yet unclear. Blobel has hypothesized, however, that membrane proteins bear a "stop transfer" sequence in addition to their signal sequences.

No one knows exactly how the proteins enter the membranes of the ER and, in the case of secreted and lysosomal enzymes, pass through them, but membranes must present a formidable barrier to protein passage. Membranes contain large quantities of lipids, which give them a hydrophobic character, whereas proteins are very large and contain numerous charged amino acid residues. They should find the membrane environment very inhospitable. Blobel postulates that attachment of the growing protein chains to the ER membranes by their signal sequences opens a pore or gate that permits protein passage. There is currently little evidence for the existence of this gate beyond the fact that proteins do, in fact, cross the membrane. In any event, during their passage the signal sequences are cleaved from most proteins.

There is evidence for the existence of a receptor for the proteins, however. For example, the growing chain of one protein carrying a signal sequence can prevent the binding of another to the ER. This type of inhibition usually means that the two proteins are competing for the same binding site—that is, for the same receptor. Proteins that do not bear

the leaders have no effect on the binding of those that do.

The ER is only a way station through which the proteins made on the bound ribosomes must pass on the journey to their final destinations. "Moreover," says Rothman, "now that they are in the ER, they are all in the same pot and must be sorted."

During their sojourn in the ER, the proteins all have more or less the same carbohydrate structure, called the core carbohydrate, attached to one of their amino acid residues. (The enzymes for making this addition appear to be in the lumen of the ER, where the secreted and lysosomal proteins are now located. The ends of the membrane proteins, such as the G protein, project into the ER lumen, and these also acquire the core carbohydrate.)

After picking up the core carbohydrate in the ER, the proteins may be transported to the Golgi apparatus, where the core carbohydrate is modified, with some sugars being removed and others added, according to Stuart Kornfeld of Washington University Medical School and Phillips Robbins of Massachusetts Institute of Technology. How the carbohydrates are changed may differ from one protein to another.

In the case of the lysosomal enzymes, a particular carbohydrate residue, modified by addition of a phosphate group, appears to be the signal directing the enzymes to their destination. Much of the evidence for this hypothesis comes from studies of the mucopolysaccharide storage (MPS) diseases, a group of genetic conditions that have been characterized as "garbage disposal" disorders by Elizabeth Neufeld of the National Institute of Arthritis, Metabolism, and Digestive Diseases. In these disorders, large quantities of complex carbohydrates (the mucopolysaccharides) accumulate in the lysosomes. The buildup of this cellular debris has severe consequences for the patients affected. They suffer neurological impairments, including mental retardation, blindness, and deafness, and also skeletal abnormalities. The conditions are usually fatal.

Most of the MPS diseases result from a deficiency of a single lysosomal enzyme that would normally break down the accumulating material. Consequently, for almost 15 years investigators have been trying to remedy the conditions by administering the appropriate enzymes. They have occasionally tried to treat human patients in this way, but most of the studies have been performed with cells taken from the patients and grown in culture.

Most of the early efforts at replacement therapy produced disappointing results. The cells took up the enzymes inefficiently, and the enzymes that were taken up did not remove the MPS debris from the cells.

The turning point came, Neufeld explains, when investigators realized that there are specific receptors on cell membranes that facilitate uptake of the enzymes. In addition, only a certain form of the enzymes can bind to these receptors.

> " . . . the primary function of the carbohydrate residue is to direct intracellular traffic. I think it is an address marker."

Neufeld and her colleagues found, for example, that when cells from patients with two different MPS diseases (Hunter's and Hurler's syndromes) were mixed together, each produced a factor that corrected the enzyme deficiency of the other. As expected, the corrective factors had the activities of the appropriate missing enzymes, but they were not identical to the enzymes as usually isolated from the cell. Of particular significance was the observation that the factors were readily taken up by cells but the cellular enzymes were not.

Neufeld's group showed that the uptake depended on the presence on the factors of a recognition marker consisting of carbohydrate. Subsequently, William Sly of the Washington University School of Medicine and Arnold Kaplan of the Saint Louis University School of Medicine determined that the recognition marker for a different lysosomal enzyme, β-glucuronidase, is a phosphorylated sugar residue, namely, mannose-6-phosphate. This was somewhat surprising because phosphorylated mannose had not previously been identified in mammalian cells, but investigators have now shown that it is the recognition marker for some half-dozen lysosomal enzymes. Says Sly, "I am convinced that the primary function of the carbohydrate residue is to direct intracellular traffic. I think it is an address marker." Neufeld compares the marker to a zip code that directs the proteins to their destination just as postal zip codes direct the mail. (She has no comment about which system is more effective, however.)

Another link in the chain of evidence suggesting that the carbohydrates on the lysosomal proteins are like zip codes comes from studies of cells from patients with I-cell disease. At first, these cells did not seem to fit the normal pattern for an MPS disease because they lack several lysosomal enzymes, not just one. Moreover, the enzymes, although missing from the patients' cells, especially cells of connective tissue, are found in high concentrations in the body fluids. The explanation for this, according to Neufeld, is that the lysosomal enzymes synthesized by I cells do not acquire the mannose-6-phosphate recognition factor because I cells cannot attach phosphate to the sugar. As a result, the proteins are treated like secretory proteins and exported from the cell.

Neufeld originally thought that the lysosomal proteins are actually secreted from all cells and the normal proteins are then recaptured by the receptors on the cell surface and sent back into the cell. Sly says, however, that work from his and several other laboratories indicates that large-scale export of the enzymes is unlikely. He suggests that the lysosomal enzymes and secreted proteins may be separated from one another in the ER or the Golgi apparatus and transported independently to their destinations. Alternatively, they may be transported together to the outer membrane in vesicles that bud off the membranes of the Golgi apparatus. The secreted proteins would then be released as the transporting vesicles fuse to the outer cell membrane, but the lysosomal enzymes would bind to the receptors there and be carried back to the lysosomes in the cell interior. In this event, the enzymes would be transitorily outside the cell but never actually released by it. Neufeld favors the latter hypothesis.

Exactly how the various proteins are transported from the ER to the Golgi apparatus and thence to their destinations is unclear, although most investigators think that they are carried in membrane-bound vesicles, as proposed by Palade. One advantage of enclosure in vesicles is that it protects the cell contents from the degradative action of lysosomal enzymes. "You do not want the lysosomal

enzymes running loose in the cell," as Neufeld puts it.

Recently, Rothman, with Richard Fine of Boston University School of Medicine, acquired evidence suggesting that membrane proteins are carried to the cell surface in small vesicles coated with the protein clathrin. Coated vesicles of this type had been shown earlier to work in the opposite direction—that is, to carry large molecules picked up from outside the cell to the inside.

According to Rothman and Fine, two waves of coated vesicles carry the G protein of VSV to the cell membrane. The first wave apparently carries the protein from the ER to the Golgi apparatus, and the second wave carries it from the Golgi to the membrane. They base this conclusion on differences in the carbohydrate content of the protein in the two waves. In the first wave, G protein contains carbohydrate like that added in the ER, whereas in the second wave, it contains carbohydrate characteristic of material that has passed through the Golgi apparatus. The evidence suggests that the vesicles are reused; they continually cycle back and forth between the cell membrane and the interior structures.

Rothman and Fine hypothesize that there may be several species of coated vesicles, each with its own specifically packaged contents and destination.

In addition to the proteins made on the membrane-bound ribosomes of the ER, there are a number of cell proteins that are synthesized on ribosomes floating free in the cell cytoplasm. The latter proteins, in contrast to those produced in the ER, are usually soluble and do not contain carbohydrate. Nevertheless, investigators have detected sequences of amino acids on the amino-terminal ends of some cytoplasmic proteins that may be signal sequences.

For example, Nam-Hai Chua of Rockefeller University, in collaboration with Blobel, has detected such a sequence on a chloroplast protein. The arrangement of amino acids in the potential signal is now being determined and is turning out to be very different from that in the signal sequences of the proteins made on bound ribosomes. Blobel points out that this is not surprising since the two signals are directing the proteins bearing them to very different locations. Some mitochondrial proteins may also have signals on their amino-terminal ends.

Not all mitochondrial proteins have them, however, according to Robert Poyton of the University of Connecticut Health Center. Paul Lazarow of Rockefeller University says that the protein catalase, which is the principal enzyme found in peroxisomes, also appears to lack a signal sequence. (Peroxisomes are organelles that break down peroxides, potentially dangerous compounds formed in some biochemical reactions.)

These findings do not necessarily mean that the proteins lack directional information of some kind. After all, they find their way to the appropriate locations. But the signals for these proteins, if they exist, still remain to be identified.

—JEAN L. MARX

Double Hubble, Age in Trouble

Measurements of distances to galaxies hint that the universe may be only half as old as was thought

On the basis of data from infrared and radio astronomy, a team of scientists estimates that the universe is merely 10 billion years old, not 20, as many astronomers believe. Their new measurements of the distances to galaxies suggest that the value for the Hubble constant—one indicator of age of the universe—should be revised significantly. "We just made the universe a lot younger than it was thought to be," says one member of the team. Although the new value for the Hubble constant challenges the accepted one, it agrees with a previous but doubted finding.

Moreover, the recent measurements suggest that the neighborhood of our galaxy may not be a representative sample of the universe. If true, this observation implies that cosmology has its own "Catch-22"—whereas the vicinity of our galaxy may not be typical of the universe as a whole, only that region can be studied accurately.

The distance measurements were made with a technique introduced last spring (*1*) by Marc Aaronson of Steward Observatory in Tucson, Jeremy Mould of Kitt Peak near Tucson, and John Huchra of Harvard-Smithsonian Center for Astrophysics in Cambridge. With the new technique, distances to galaxies known to be close to each other were found to be very nearly equal. Furthermore, the method is thought to be based on physical principles. The older ways of measuring distance are based for the most part on chains of assumptions and extrapolations. Thus, many astronomers welcome the technique as a valuable additional way to measure extragalactic distances.

The technique is directly useful only for measuring distances that are quite small on a cosmic scale, although large when compared with the ranges of several of the established ways of measuring distance. The measurement of larger distances is based on the concept that the universe is expanding, and the farther apart any two objects are, the faster they recede from each other. It is assumed that distance is linearly proportional to velocity. The constant of proportionality is known as the Hubble constant, after the late Edwin Hubble of Mount Wilson Observatory (now part of Hale Observatories) near Los Angeles. Due in large part to Hubble's work 50 years ago, the concept of the expanding universe became generally accepted.

Astronomers get a clue to the distance of a faraway galaxy by measuring how fast it is receding from us. They can estimate the distance accurately if they know the Hubble constant. Consequently, a major quest of cosmologists has been to determine the value of the Hubble constant on the basis of the velocities of galaxies which are close enough to the earth that their distances can be measured directly. As yet, astronomers making such meaurements have not been able to agree on the value of the Hubble constant appropriate for the universe as a whole.

The new data of Aaronson, Mould, and Huchra suggest that the Hubble constant changes with distance. This observation supports the idea that there might be a local blemish in the expansion of the universe. The Hubble constant calculated from data on the galaxies in the nearby Virgo cluster of galaxies is much

0036-8075/80/0111-0167$00.50/0