Overview of the Physical State of Proteins Within Cells

The word protein comes from the Greek word *proteios*, meaning primary. And, indeed, proteins are of primary importance in the study of cell function. It is difficult to imagine a cellular function not linked with proteins. Almost all biochemical catalysis is carried out by protein enzymes. Proteins participate in gene regulation, transcription, and translation. Intracellular filaments give shape to a cell while extracellular proteins hold cells together to form organs. Proteins transport other molecules, such as oxygen, to tissues. Antibody molecules contribute to host defense against infections. Protein hormones relay information between cells. Moreover, protein machines, such as actin-myosin complexes, can perform useful work, including cell movement. Thus, studying proteins is a prerequisite in understanding cell structure and function.

The physical characterization of proteins began well over 150 years ago with Mulder's characterization of the atomic composition of proteins. In the latter half of the nineteenth century Hoppe-Seyler (1864) crystallized hemoglobin and Kühn (1876) purified trypsin. A variety of physical methods have been developed over the years to increase convenience and precision in the characterization and isolation of proteins. These include ultracentrifugation, chromatography, electrophoresis, and others. In many instances our understanding of cell proteins parallels the introduction and use of new techniques to examine their structure and function.

PROTEIN CLASSIFICATIONS

All proteins are constructed as a linear sequence(s) of various numbers and combinations of ~20 α-amino acids joined by peptide bonds to form structures from thousands to millions of daltons in size. Proteins are the most complex and heterogeneous molecules found in cells, where they account for >50% of the dry weight of cells and ~75% of tissues.

Proteins can be classified into three broad groups: globular, fibrous, and transmembrane (Fig. 1.5.1; Table 1.5.1). Globular proteins are, by definition, globe-shaped, although in prac-

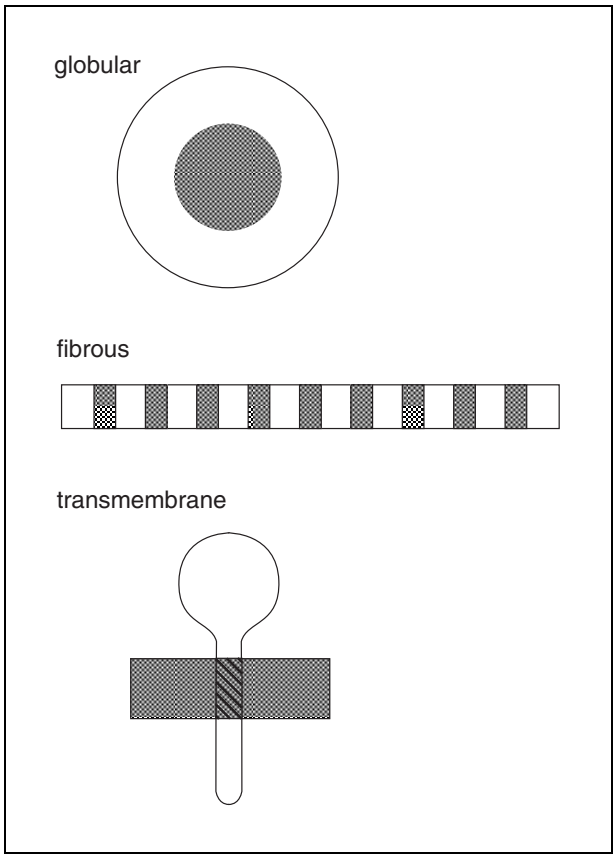

Figure 1.5.1 General classifications of proteins. In these schematic representations of globular, fibrous, and transmembrane proteins, hydrophobic regions are shaded. Note that the disposition of hydrophobic residues often reflects the protein class.

Contributed by Howard R. Petty
Current Protocols in Protein Science (2002) 1.5.1-1.5.10

Table 1.5.1 Broad Classifications for Proteins[a]

Type	Location/type	Examples
Globular	Intracellular	Hemoglobin, lactate dehydrogenase, cytochrome *c*
	Extracellular	Serum albumin, immunoglobulins, lysozyme
Fibrous	Intracellular	Intermediate filaments, tropomyosin, lamins
	Extracellular	Collagen, keratin, elastins
Transmembrane	Single pass	Insulin receptor, glycophorin, HLAs[b]
	Multipass	Glucose transporter, rhodopsin, acetylcholine receptor

[a]Additional information regarding fibrous and transmembrane proteins can be found in Squire and Vibert (1987) and Petty (1993). Information concerning globular proteins can be found in numerous books on proteins and enzymes such as Schultz and Schirmer (1979).

[b]Human histocompatibility leukocyte antigens.

tice they can be spherical or ellipsoidal. Globular proteins are generally soluble in aqueous environments. Examples of globular proteins are hemoglobin, serum albumin, and most enzymes. Fibrous proteins are elongated linear molecules that are generally insoluble in water and resist applied stresses and strains. Collagen is a physically tough molecule of connective tissue. Just as collagen gives strength to connective tissues, intermediate filaments linked to desmosomes give strength to cells in tissues. The third general class of proteins, transmembrane proteins, contain a hydrophobic sequence buried within the membrane; these proteins are discussed more fully below (see Membrane Proteins).

These protein categories are not mutually exclusive. For example, the nominally fibrous intermediate filament proteins also have globular domains. Similarly, transmembrane proteins almost always possess globular domains. Thus, these definitions serve as a useful guide but should not be rigidly applied.

HYDROPATHY PATTERNS OFTEN REFLECT A PROTEIN'S CLASSIFICATION

A key physical feature of proteins is their hydropathy pattern (i.e., the distribution of hydrophobic and hydrophilic amino acid residues). Indeed, hydrophobic interactions provide the primary net free energy required for protein folding. Figure 1.5.1 illustrates the disposition of hydrophobic amino acids in proteins. In an intact globular protein, hydrophobic amino acids are generally shielded from the aqueous environment by coalescing at the center of the molecule, with the more hydrophilic residues exposed at its surface. However, the linear arrangement of hydrophobic residues fluctuates in an apparently random fashion. The α helices within globular proteins may express a hydrophobic face oriented toward the center of the protein. (Within these helices hydrophobic residues are nonrandomly positioned every three or four amino acids to yield a hydrophobic face.) For coiled-coil α helix–containing fibrous proteins, such as tropomyosin and α-keratin, hydrophobic residues at periodic intervals allow close van der Waals contact of the chains and potentiate assembly as hydrophobic residues are removed from the aqueous environment (Schulz and Schirmer, 1979; Parry, 1987). Secondarily, regularly spaced charged groups can also contribute to the shape of fibrous proteins (Schulz and Schirmer, 1979; Parry, 1987). Transmembrane proteins provide a rather different physical arrangement of hydrophobic residues in which hydrophobic residues are collected primarily into a series of amino acids that is embedded within a cell membrane.

One important means of analyzing the hydropathy of a sequenced protein is a hydropathy plot (Kyte and Doolittle, 1982). In this method, each amino acid residue is assigned a hydropathy value, an ad hoc measure that largely reflects its relative aqueous solubility; these values are plotted after being averaged. The successful interpretation of hydropathy plots

depends on the parameters chosen for averaging. The parameters are the number of residues averaged (amino acid interval or "window") and how many amino acids are skipped when calculating the next average (step size). Using this approach with a window of ~10 residues, it is often possible to find the positions of hydrophobic residues coalescing near the interior of globular proteins. The method is particularly useful in predicting transmembrane domains of proteins, generally with a window of ~20 amino acids. To detect the repetitious pattern of coiled-coil fibrous proteins, however, windows smaller than the repeat length would be required.

MEMBRANE PROTEINS

In addition to their presence in the extracellular and intracellular milieus, proteins are also found in association with biological membranes. Proteins constitute one-half to three-quarters of the dry weight of membranes. Membrane proteins perform a broad variety of functions including intermembrane and intercellular recognition, transmembrane signaling, most energy-harvesting processes, and biosynthesis in the endoplasmic reticulum (ER) and Golgi complex.

Membrane proteins have been traditionally characterized as integral (or intrinsic) or peripheral (or extrinsic) on the basis of operational criteria. Peripheral membrane proteins are associated with membrane surfaces and can be dislodged from membranes using hypotonic or hypertonic solutions, pH changes, or chelation of divalent cations. Components of the erythrocyte membrane skeleton, for example, are peripheral membrane proteins. Although most peripheral proteins are removed by washing a sample with buffers, integral proteins cannot be removed by such treatments. To isolate integral membrane proteins, which are embedded within the lipid bilayer, one must use detergents that disrupt the bilayer and bind to the proteins, thus solubilizing them. In general, integral membrane proteins have a portion of their peptide sequence buried in the lipid bilayer whereas peripheral proteins do not. However, the discovery of glycosylphosphatidylinositol (GPI)-linked membrane proteins added to the ambiguity of the situation. GPI-linked proteins are globular proteins with no membrane-associated peptide sequence, yet they require harsh conditions for solubilization.

As the technology for studying membrane proteins improved, it became necessary to develop a more precise vocabulary to describe membrane proteins. Transmembrane integral membrane proteins have at least one stretch of amino acids spanning a membrane. Membrane proteins are classified as type I, II, III, or IV depending on the nature of their biosynthesis and topology in membranes (Spiess, 1995; Table 1.5.2 and Fig. 1.5.2). The biosynthetic insertion of these proteins in membranes is, in turn, dependent on the presence or absence of a cleavable signal peptide, the relative positions of the hydrophobic transmembrane domain and positively charged topogenic signals, and/or

Table 1.5.2 Definitions of Integral Transmembrane Proteins

Type	Definition	Examples
I	An N-terminal–cleavable signal peptide is removed at the luminal face yielding a luminal N terminal during biosynthesis. (Positive charges are found on C-terminal side of first long hydrophobic sequence after the signal peptide.)	LDL receptor, insulin receptor, glycophorin A, thrombin receptor
II	An N-terminal–uncleaved signal peptide leads to a cytoplasmic N terminus. (Positive charges are generally found on N-terminal side of first long hydrophobic sequence.)	Transferrin receptor, sucrase/isomaltase, band 3
III	A long N-terminal hydrophobic sequence is followed by a sequence of positive charges. This leads to a luminal N terminus in the absence of a cleavable signal peptide.	β-Adrenergic receptor, cytochrome P450
IV	A short C terminus is present at the luminal side of membrane. A large N terminus is exposed at the cytoplasmic face.	Synaptobrevin, UBC6

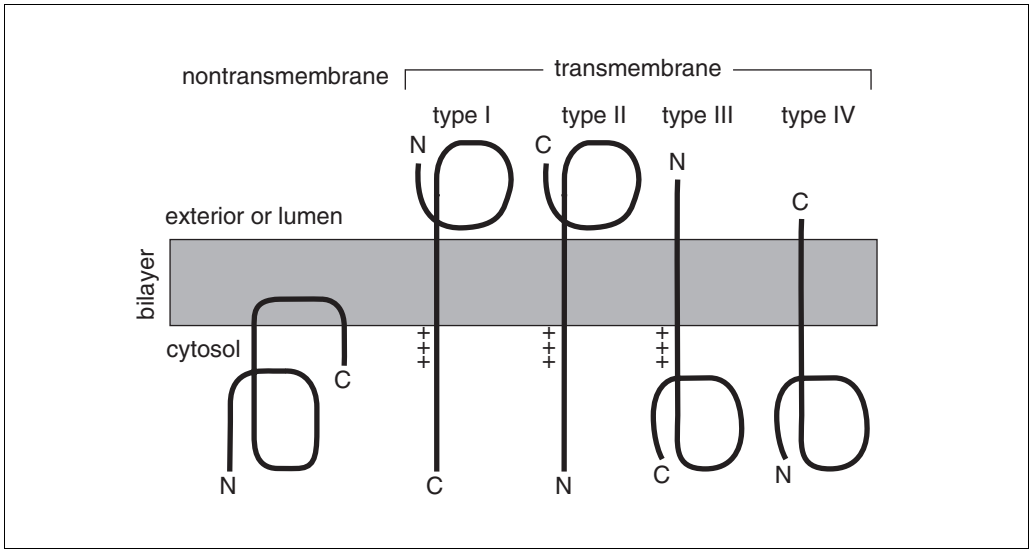

Figure 1.5.2 Membrane proteins containing hydrophobic anchors. A nontransmembrane or monotopic membrane protein is anchored to the membrane via a hydrophobic amino acid sequence. Transmembrane proteins are classified as types I, II, III, and IV (Table 1.5.2). The first transmembrane segment of a multispanning membrane protein can be inserted as in type I, II, or III proteins. This segment functions as a start-transfer peptide. Subsequent transmembrane segments will function as stop-transfer and start-transfer sequences, resulting in a multispanning membrane topology.

the mechanism of nascent protein delivery to the ER.

Type I membrane proteins are synthesized with an amino-terminal signal sequence that is inserted into the ER membrane. When the signal sequence is proteolytically removed in the ER lumen, a new luminal amino terminus is exposed. A series of positively charged residues at the C-terminal side of the first hydrophobic transmembrane domain following the signal sequence generally denotes the end of the first transmembrane domain (von Heijne and Gavel, 1988). Although a hydrophobic transmembrane domain followed by a positive sequence of amino acids is sufficient to act as a stop-transfer signal, this motif is not required for stop-transfer events and other, less well-understood regulatory mechanisms are also involved (Andrews and Johnson, 1996).

Membrane proteins types I, II, and III are delivered to the ER membrane via a signal recognition particle (SRP)-dependent mechanism. In contrast to type I proteins, type II and III membrane proteins do not have a cleavable N-terminal signal sequence. Instead, they have an internal hydrophobic signal that acts as both a signal sequence for ER delivery and a transmembrane domain in the mature protein. Type II proteins have a cytoplasmic amino terminus and a luminal (or extracellular) carboxyl terminus. In this case a positively charged sequence of amino acids at the N-terminal side of the first hydrophobic sequence causes the amino terminus to be retained at the cytoplasmic face of the ER membrane. Thus the internal uncleaved signal peptide becomes the transmembrane domain of the mature protein.

Type III membrane proteins have the same overall topology as type I proteins, but they are inserted into membranes by a different mechanism. In type III proteins the first hydrophobic sequence of amino acids is immediately followed by a series of positively charged amino acids. Thus, the first hydrophobic sequence becomes the transmembrane domain of the protein, with the amino terminus at the luminal face of the membrane.

Type IV membrane proteins are characterized by a large, cytoplasmically exposed amino-terminal domain and a short carboxyl-terminal domain facing the lumen. Importantly, these proteins are delivered to the ER by an unknown SRP-independent mechanism.

In addition to the single-pass membrane proteins just described, integral membrane proteins can display zero, two, three, or more transmembrane domains. Some membrane proteins, such as cytochrome b_5, have protein segments buried in the hydrophobic core of membranes but do not cross the membrane. Membrane proteins with multiple membrane-spanning domains are classified as type I, II, or

III depending on the topogenic signals in the first transmembrane domain. For example, a multispan membrane protein with a cleavable signal sequence, luminal amino terminal, and a positively charged sequence following the first transmembrane domain from the amino terminal, such as the thrombin receptor, is a type I membrane protein. The remaining transmembrane domains are inserted into the bilayer depending on the orientation of the first transmembrane domain. Multispanning type II and III proteins are similarly defined according to the properties of their single-spanning counterparts.

In addition to hydrophobic protein sequences acting as membrane anchors, membrane proteins may also carry bilayer-associated hydrophobic lipid components. These hydrophobic lipid anchors define three broad groups of lipid-modified proteins: fatty acylated, isoprenoid-linked, and GPI-linked (Fig. 1.5.3). Several cytosolic transmembrane proteins have been identified that contain a covalently attached hydrophobic fatty acyl residue. For example, fatty acids, including palmitic, palmitoleic, *cis*-vaccenic, and cyclopropylene-hexadecanoic, are covalently linked to the amino terminus and the amino-terminal glycerylcysteine of *E. coli* lipoprotein. Moreover, palmitate- and myristate-labeled transmembrane proteins have been observed in eukaryotic cells (e.g., Schlesinger et al., 1980).

In both isoprenoid-linked and GPI-linked proteins, globular proteins become membrane-bound due to the addition of a hydrophobic lipid moiety. Certain proteins containing conserved cysteine residues at or near the C-terminus are modified by prenylation, in which a farnesyl or geranylgeranyl isoprenoid tail is added (Zhang and Casey, 1996). This hydrophobic moiety promotes protein association with the cytoplasmic face of cell membranes. Notably, cytosolic G proteins and protein kinases that participate in signal transduction are prenylated.

GPI-linked proteins are a major class of membrane proteins (Cardoso de Almeida, 1992; Englund, 1993). In contrast to isoprenoid-modified proteins, GPI-linked proteins are attached to the luminal or extracellular face of membranes via a glycosylphosphatidylinositol anchor of variable structure (e.g.,

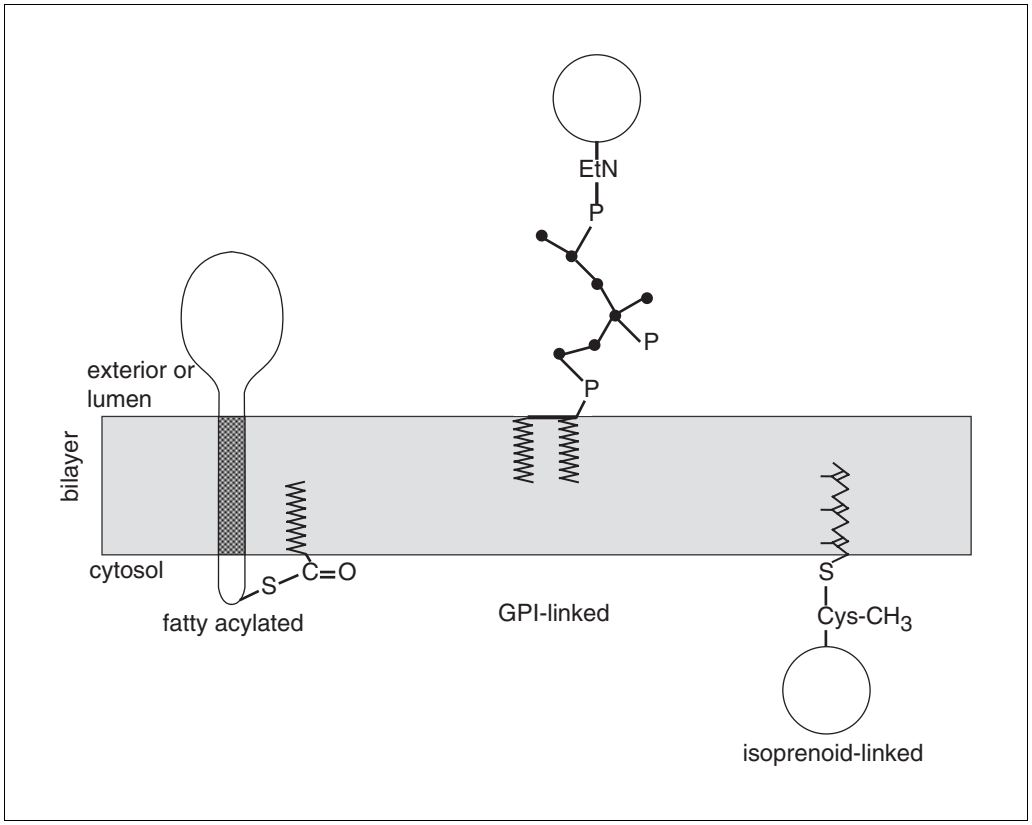

Figure 1.5.3 Membrane proteins containing lipid moieties. In the simplest case, fatty acids can be covalently attached to transmembrane proteins. Hydrophobic tails are also attached to proteins to form isoprenoid-linked proteins. A third class of lipid-attached proteins are the GPI-linked proteins. Hydrophobic regions are shaded.

Fig. 1.5.3). Well over 100 GPI-linked proteins have been identified in cells, where they perform numerous functions including acting as enzymes and receptors. The ability of GPI-linked proteins, which possess no transmembrane or cytosolic sequences, to elicit transmembrane signals seems paradoxical. However, studies have suggested that interactions with other proteins, including transmembrane integrins (Petty et al., 1996), contribute to transmembrane signaling of these proteins. In addition, GPI-linked proteins can collect in microdomains called lipid rafts within cell membranes (Rietveld and Simons, 1998). Although GPI-linked proteins must collaborate with other membrane proteins to elicit signals, they do possess certain functional advantages. First, GPI-linked proteins (and isoprenoid-linked proteins as well) diffuse in membranes much faster than transmembrane proteins and thus relay information faster. Second, certain cells, such as leukocytes, can rapidly shed their GPI-linked proteins, thus altering their functional properties in seconds. Although the importance of lipid-linked membrane proteins has only recently been appreciated, the impact of these structures on our understanding of cell properties is growing rapidly.

ADDITIONAL FACTORS AFFECTING THE PHYSICAL HETEROGENEITY OF PROTEINS

Additional factors contributing to the physical-chemical heterogeneity of proteins are size, charge, chemical modifications, and assembly. A typical amino acid has a molecular mass of ~110 Da, and a small protein has a molecular mass of a few thousand daltons (e.g., for insulin, M_r = 5733). Large proteins have molecular masses of several hundred thousand daltons. When proteins are assembled to form large multiprotein complexes such as ribosomes, molecular masses are well into the millions. The diameters of these structures range from 4 Å for an individual amino acid to ~30 nm for a ribosome.

Electrostatic charge is of major importance in protein structure and function. Charged proteins are more soluble than uncharged proteins. The large number of positive charges on histones allow them to bind DNA. The spatial arrangement of charges on cytochrome *c* allows it to bind the complementary charges of its oxidase and reductase, thereby orienting the proteins prior to electron transfer. Similarly, the arrangement of charges on the apoprotein and receptor for low-density lipoprotein (LDL) allows for lock-and-key–like interactions (Petty, 1993). In addition to structural and binding considerations, electrostatic interactions play a regulatory role. For example, the phosphorylation and dephosphorylation of insulin receptors alter electrostatic interations between the active site and a regulatory loop of the kinase domain, thereby changing its three-dimensional shape (Hubbard et al., 1994). This changes the V_{max} of the kinase, thus triggering intracellular signals.

In addition to the types of physical heterogeneity listed above, >100 distinct chemical modifications of proteins have been observed. These include, for example, glycosylation, ubiquitin attachment, phosphorylation, acetylation, and hydroxylation (Table 1.5.3). Thus, proteins undergo extensive physical-chemical modification.

PROTEIN ASSEMBLIES

Proteins can be assembled in a variety of states in both aqueous media and within membranes. Protein assembly into complex supramolecular structures plays vital roles in enzyme regulation, cell skeleton formation, and transmembrane signaling. Both covalent bonds and noncovalent bonds participate in protein assembly. One frequently encountered covalent mechanism of protein assembly is the formation of disulfide bonds. These covalent linkages often form during protein maturation. They can link two separate proteins together or two portions of the same protein. For example, the two chains of insulin molecules are held together by disulfides, as are the two chains of its membrane receptor. However, disulfide bond formation is mostly limited to oxidative environments such as the ER lumen and the exterior face of the cell surface.

One of the best-known examples of noncovalent assembly is the formation of hemoglobin tetramers. Polymerization is another frequently encountered mechanism for protein assembly in cells. The globular protein actin polymerizes to form microfilaments in the absence of covalent bond formation. Intermediate filaments are formed by the polymerization of fibrous proteins. Under certain circumstances transmembrane proteins polymerize as well; bacteriorhodopsin, for example, forms two-dimensional pseudocrystals called purple membranes. Protein assemblies formed from various numbers of similar units are homodimers, homooligomers, and homopolymers.

Assembly of protein structures from dissimilar subunits is more common than assem-

Table 1.5.3 Common Physical-Chemical Modifications of Proteins[a]

Modification	Example
Homodimerization	Transferrin receptor
Homooligomerization	*S. typhimurium* glutamine synthetase
Homopolymerization	Actin
Heterodimerization	Integrins
Heterooligomerization	Histones, proteasomes
Heteropolymerization	Ribosomes
Proteolytic cleavage	Signal peptide cleavage in ER
Prosthetic group addition	Heme addition to cytochromes and hemoglobin
Oxidation-reduction	Disulfide bond formation in ER
Glycosylation	Glycoprotein maturation
Phosphorylation	Regulation of protein function, such as the tyrosine kinase activity of insulin receptors
Acetylation	Blockage of N-termini of certain membrane proteins
Ubiquitination	Ubiquitin-dependent proteolysis via proteasomes, histones
Hydroxylation	Proline hydroxylation on collagen
Fatty acylation	Insulin receptors, *E. coli* lipoprotein
Isoprenylation	G proteins
GPI addition	Alkaline phosphatase, urokinase receptors

[a]For details, see Freedman and Hawkins (1980, 1985), Schlesinger et al. (1980), Englund (1993), and Zhang and Casey (1996).

bly from identical subunits. For example, heterodimers are formed from the α and β chains of integrins within cell membranes. Complex heterooligomeric and heteropolymeric structures vary from relatively small structures such as histone octamers, which bind to DNA in the nucleus, to large particles such as ribosomes, found both in the cytosol and attached to nuclear and ER membranes. The signal recognition particle is a relatively small heterooligomeric structure, composed of one RNA subunit and six proteins, that potentiates the delivery of secretory and most membrane proteins to the ER membrane. Membrane-associated heterooligomeric structures have also been observed. One of the best examples of such structures is the components of the electron transport systems in chloroplasts and mitochondria (Petty, 1993). For example, the ubiquinone-cytochrome *c* reductase is composed of eleven different subunits. Thus, proteins can be assembled in a variety of manners within cells.

Although some protein assemblies, such as intermediate filaments, are static structures, many are dynamic structures which provide functional flexibility. For example, microfilaments can rapidly assemble and disassemble. In addition to the physical changes in assembly state, compositional dynamics is also observed. For example, interferon γ treatment alters the composition of proteasomes. Developmental changes in protein composition are also observed. As an example, fetal and newborn forms of a component of cytochrome *c* reductase are expressed in humans. Thus, protein assemblies can be characterized by both physical and compositional dynamics.

ALTERING THE SOLUBILITY OF PROTEINS: PROTEIN EXTRACTION

The in vitro characterization of cellular proteins begins with their extraction from tissues or cells into a buffer. With the exception of globular secretory proteins, such as those found in plasma, proteins are generally not easily accessible for experimental manipulation. For example, many fibrous proteins are not soluble in aqueous buffers. Cellular proteins are entrapped within or on a cell and therefore must be extracted from the cell in a soluble form.

A variety of methods including osmotic lysis, enzyme digestion, homogenization using a blender or mortar and pestle, and disruption by French press and sonication have been employed to disrupt cells. For a cytosolic protein such as hemoglobin, no further extraction from the sample is necessary. However, many important cellular proteins, such as those associated

with membranes, cytoskeletal components, and DNA, remain insoluble. To further solubilize cell proteins, both nonionic (e.g., Triton X-100) and ionic (e.g., sodium dodecyl sulfate) detergents are often employed. Detergents are small amphipathic molecules that interact with both nonpolar and polar environments. Detergents disrupt membranes. They also bind to hydrophobic regions of proteins, such as their transmembrane domains, thereby replacing the unfavorable contacts between hydrophobic protein regions and water with the more favorable hydrophilic domains of the detergent. Thus, instead of the hydrophobic regions of the insoluble protein forming an aggregate in the bottom of a test tube, the protein becomes soluble and can be employed in most in vitro analyses.

In addition to detergents, several other solubilization strategies are useful for the extraction and in vitro characterization of proteins (Table 1.5.4). Chaotropic agents enhance the transfer of nonpolar molecules to aqueous environments by their disrupting influence on water structure. Chaotropic agents are generally large molecular ions such as thiocyanate (SCN^-), perchlorate (ClO_4^-), and trichloroacetate (CCl_3COO^-). Hydrophobic interactions are also reduced by exposure to organic solvents and low salt concentrations. Electrostatic interactions are reduced by high salt conditions; this decreases the Debye-Hückel screening length and coulombic attraction. To disrupt hydrogen bonds, high concentrations of urea or guanidine are often employed. More vigorous methods of sample denaturation using very low pH or harsh detergents such as sodium dodecyl sulfate are also used to diminish intermolecular contacts.

Once proteins are extracted, their size can be characterized by ultracentrifugation on sucrose gradients (*UNIT 4.2*), gel filtration chromatography (*UNIT 8.3*), SDS-PAGE (*UNIT 10.1*), and other methods (Table 1.5.5). The charge characteristics of proteins can be assessed using isoelectric focusing and ion-exchange chromatography. Specific interactions, such as antigen-antibody and biotin-avidin interactions, can also be employed in the characterization and isolation of proteins. These are useful in immunoblotting (*UNIT 10.10*) and affinity chromatography methods (Chapter 9).

Table 1.5.4 Physical Bases of Common Protein Extraction and/or Elution Methods

Physical property perturbed	Agents
Hydrogen bonds	Urea or guanidine·HCl, pH changes
Ion pair interactions	High salt, pH changes
Hydrophobic interactions	Detergents, chaotropic agents, organic solvents, low salt

Table 1.5.5 Physical Bases of Common Protein Characterization and Isolation Methods

Physical property	Method	References to other units
Solubility	Extraction with salts, detergents, and enzymes	Chapter 4, Racker (1985)
Size	Ultracentrifugation on sucrose gradients	*UNIT 4.2*
	Gel filtration	*UNIT 8.3*
	SDS-PAGE	*UNIT 10.1*
Charge	Isoelectric focusing	*UNIT 10.2*
	Ion-exchange chromatography	*UNIT 8.2*
Biospecific interaction	Immunoblotting	*UNIT 10.10*
	Immunoaffinity chromatography	Chapter 9
Hydrophobicity	Hydrophobic chromatography	*UNIT 8.4*
	Reversed-phase HPLC	*UNIT 8.7*

LIMITATIONS OF THE IN VITRO MANIPULATION OF PROTEINS

The very act of isolating proteins perturbs their physical environment. Although this is not often a major problem, a few cautionary notes should be made. The most primitive compartment of a cell, the cytosol, is a chemically reducing environment. Consequently, free sulfhydryl groups are observed in the cytosol; in fact, multiple cytosolic pathways help in preserving the proper redox conditions. On the other hand, the extracellular milieu and the luminal side of the ER are oxidative environments. The oxidizing condition within the ER is presumably due to the unidirectional transport of glutathione and cystine. Consequently, disulfides are frequently observed in the ER and extracellular environments. Thus, to prevent disulfide formation during manipulation, sulfhydryl blocking reagents such as iodoacetamide are included in extraction buffers. The cytosol is also a K^+-rich and Ca^{2+}-poor solution. These parameters should be considered in designing physiologically relevant experiments.

The experimental manipulation of membrane proteins is decidedly more difficult. The exterior face exists in a high Na^+ and Ca^{2+} solution that is oxidative; just the opposite is true for the cytoplasmic face. Since no appropriate solvent exists for such isolated proteins, experimental questions can be directed at properties associated with just one face of the molecule. A second limitation common to all in vitro studies of transmembrane proteins is that they must be solubilized using detergents. In addition to solubilizing a transmembrane protein, detergents can also bind to hydrophobic regions in the globular domain(s) of the protein, thus affecting the properties under study. One means of countering this problem is to test several detergents in the hope of finding one that retains the full biological activity of the purified protein.

Protein solubilization can also lead to loss of physiologically relevant protein-protein interactions. This can occur by simple dilution or by disruption of noncovalent interactions among proteins. For example, hemoglobin exists as a supersaturated solution in vivo which cannot be duplicated in vitro. Furthermore, protein-protein associations are generally stronger in the restricted confines of membranes than after solubilization into a buffer. Thus, protein assemblies found in cells may disappear during solubilization. One means of countering these potential difficulties is to covalently cross-link protein assemblies prior to disruption and to solubilize proteins using mild detergents (e.g., Brij-58).

CONCLUSIONS

Structural motifs, especially stretches of hydrophobic amino acids, contribute to the shape of a protein and its classification as globular, fibrous, or transmembrane. Proteins are heterogeneous at many different levels including physical attributes, covalent modifications, and supramolecular assembly. The physical properties of proteins are used to characterize and isolate these molecules. For example, the size of a protein is examined by sedimentation on sucrose gradients (*UNIT 4.2*), gel filtration (*UNIT 8.3*), and polyacrylamide gel electrophoresis (*UNIT 10.1*). Its charge is the key physical parameter in isoelectric focusing and ion-exchange chromatography. The units that follow contain detailed protocols describing the characterization of cellular proteins.

LITERATURE CITED

Andrews, D.W. and Johnson, A.E. 1996. The translocon: More than a hole in the ER membrane? *Trends Biochem. Sci.* 21:365-369.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.) 1998. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Cardoso de Almeida, M.L. 1992. GPI Membrane Anchors. Academic Press, New York.

Englund, P.T. 1993. The structure and biosynthesis of glycosyl phosphatidylinositol protein anchors. *Annu. Rev. Biochem.* 62:121-138.

Freedman, R.B. and Hawkins, H.C. 1980. The Enzymology of Post-Translational Modifications of Proteins, Vol. 1. Academic Press, New York.

Freedman, R.B. and Hawkins, H.C. 1985. The Enzymology of Post-Translational Modifications of Proteins, Vol. 2. Academic Press, New York.

Hoppe-Seyler, F. 1864. Über die chemischen und optischen Eigenschaften des Blutfarbstoffs. *Virchows Arch.* 29:233-235.

Hubbard, S.R., Wei, L., Ellis, L., and Hendrickson, W.A. 1994. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* 372:746-754.

Kühn, W. 1876. Über das Verhalten verschiedner organisirter und sogenannter ungeformter Fermete. Über das Trypsin (Enzym des Pankreas) [Reprint, *FEBS Lett.* 62:E3-E7 (1976).].

Kyte, J. and Doolittle, R.F. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132.

Parry, D.A.D. 1987. Fibrous protein structure and sequence analysis. *In* Fibrous Protein Structure (J.M. Squire and P.J. Vibert, eds.) pp. 141-171. Academic Press, New York.

Petty, H.R. 1993. Molecular Biology of Membranes: Structure and Function. Plenum, New York.

Petty, H.R., Worth, R.G., and Todd, R.F. III. 2002. Interactions of integrins with their partner proteins in leukocyte membranes. *Immunol. Res.* 25:75-95.

Racker, E. 1985. Reconstitutions of Transporters, Receptors, and Pathological States. Academic Press, New York.

Rietveld, A. and Simons, K. 1998. The differential miscibility of lipids as the basis for the formation of functional membrane rafts. *Biochim. Biophys. Acta.* 1376:467-479.

Schlesinger, M.J., Magee, A.I., and Schmidt, M.F.G. 1980. Fatty acylation of proteins in cultured cells. *J. Biol. Chem.* 255:10021-10024.

Schultz, G.E. and Schirmer, R.M. 1979. Principles of Protein Structure. Springer-Verlag, New York.

Spiess, M. 1995. Heads or tails—what determines the orientation of proteins in the membrane. *FEBS Lett.* 369:76-79.

Squire, J.M. and Vibert, P.J. (eds.) 1987. Fibrous Protein Structure. Academic Press, New York.

von Heijne, G. and Gavel, Y. 1988. Topogenic signals in integral membrane proteins. *Eur. J. Biochem.* 174:671-678.

Zhang, F.L. and Casey, P.J. 1996. Protein prenylation: Molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* 65:241-269.

KEY REFERENCES

Racker, 1985. See above.

A wonderful little book on membrane protein manipulation which disproves the hypothesis that scientists can't write.

Tanford, C. 1961. Physical Chemistry of Macromolecules. Academic Press, New York.

A rigorous introduction to the physical properties of proteins, which remains useful several decades later.

Tanford, C. 1980. The Hydrophobic Effect. John Wiley & Sons, New York.

A very readable introduction to the hydrophobic effect.

INTERNET RESOURCES

http://www.expasy.ch

A user-friendly protein database including two-dimensional PAGE data and 3D protein structures.

ftp://ftp.pdb.bnl.gov/

Contains protein crystallography data.

http://ww.ncbi.nlm.nih.gov

An important protein database.

Contributed by Howard R. Petty
Wayne State University
Detroit, Michigan