CHAPTER 2 Lipid modifications of proteins

Nikola A. Baumann and Anant K. Menon

Department of Biochemistry, University of Wisconsin-Madison. 433 Babcock Drive, Madison, WI 53706-1569, USA

1. Preamble

Lipid modifications of proteins (Fig. 1) are widespread and functionally important in eukaryotic cells. For example, many intracellular proteins such as the signal-transducing heterotrimeric GTP-binding proteins (G proteins) and the Ras superfamily of G proteins are modified by 14- or 16-carbon fatty acids and/or 15- or 20-carbon isoprenoids. Also, a variety of cell surface proteins such as acetylcholinesterase, the T lymphocyte surface antigen Thy-1, and members of the cell adhesion protein family are modified by glycosylphosphatidylinositol (GPI) anchors. In most cases, the lipid moiety is crucial to protein function as it allows an otherwise water-soluble protein to interact strongly with membranes. The lipid moiety may also aid in the sorting of the protein to membrane domains that promote lateral and transbilayer protein–protein interactions that are critical for cell function. In some instances, the covalent lipid acts as a functional switch resulting in functional membrane association of certain protein conformations but not of others.

The covalent attachment of lipid to protein was first described in a study of myelin protein in 1951, but only clearly documented as important for protein biosynthesis and function in a study of the outer membrane murein lipoprotein of *Escherichia coli* by Braun and Rehn in 1969. These early discoveries were followed, in the 1970s, by the identification of fatty acids linked to viral glycoproteins and of isoprenoids covalently attached to fungal mating factors and to GTP-binding proteins. The 1980s saw the identification and characterization of *N*-myristoylated proteins and GPI-anchored proteins, and work on tissue patterning factors in the 1990s revealed a new class of autoprocessed proteins modified by cholesterol. Our purpose in this chapter is to document the structure of these various lipid modifications, describe their biosynthesis, and survey their functional significance. The chapter does not cover the structure and biosynthesis of diacylglycerol-modified proteins found in *E. coli* and other bacteria; information on this may be found in articles by Wu and colleagues (Wu, 1993).

2. Protein prenylation

Prenylated proteins constitute approximately 0.5–2% of all proteins in mammalian cells. They contain a farnesyl (15-carbon) or geranylgeranyl (20-carbon) isoprenoid attached via a thioether linkage to a cysteine residue at or near the carboxy-terminus of the protein. Protein prenylation was originally discovered when certain fungal



Fig. 1. Membrane topology of lipid modified proteins and structures of the covalently attached lipids. The structures (left to right) represent *N*-myristoyl glycine, palmitate thioester-linked to cysteine, farnesyl thioether-linked to cysteine, geranylgeranyl thioether-linked to cysteine, cholesterol ester-linked to glycine, and a minimal GPI anchor linked to the ω amino acid in a GPI-anchored protein. The GPI structure is shown with a diacylglycerol moiety containing two, ester-linked fatty acids. Other GPI anchors are based on ceramide, while yet others have monoacylglycerol, a fatty acid in ether-linkage to glycerol, and/or a fatty acid ester-linked to inositol.

peptide mating factors were shown to contain a carboxy-terminal cysteine modified by a thioether-linked farnesyl group. However, it was not until Glomset and colleagues (Schmidt, 1984) subsequently showed that animal cell proteins could be metabolically radiolabeled with radioactive mevalonate, an isoprenyl group precursor, that protein prenylation was more widely appreciated. Glomset and colleagues initially observed that the growth arrest of mammalian cells induced by compactin, an inhibitor of isoprenoid biosynthesis, could not be reversed by exogenously added sterols including cholesterol, the major product of the isoprenoid pathway (Brown, 1980). The compactininduced growth arrest could, however, be alleviated by small amounts of mevalonate, suggesting that mevalonate itself or a non-sterol metabolite of mevalonate played an important role in the growth cycle of cells. This result prompted studies in which cells were metabolically labeled with radioactive mevalonate and led to the discovery that almost 50% of the cell-associated radioactive mevalonate could not be extracted into lipid solvents as a result of post-translational (cycloheximide-insensitive) covalent association with proteins.

Protein prenylation is catalyzed by one of three different multisubunit prenyltransferases located in the cytoplasm of cells [1]. The majority of prenylated proteins, including most members of the Ras family of G proteins [2], contain a carboxy-terminal CaaX motif (CaaX box) composed of a conserved cysteine residue, two aliphatic amino acids (a) and a variable carboxy-terminal residue (X). The CaaX box is recognized by CaaX prenyltransferases that catalyze the attachment of a farnesyl or geranylgeranyl group from the corresponding isoprenyl pyrophosphates to the cysteine residue (Fig. 2). The CaaX prenyltransferases involved in these reactions are protein farnesyltransferase (FTase) and protein geranylgeranyltransferase type I (GGTase-I). FTase recognizes CaaX boxes where X = M, S, Q, A, or C, whereas GGTase-I recognizes CaaX boxes with X = L or F. Other prenylated proteins, such as the Rab proteins involved in vesicular transport, terminate in a CC or CXC motif; these proteins are substrates for protein geranylgeranyltransferase type II (GGTase-II) [3].

Subsequent to prenyl modification, Ras and most other CaaX proteins are further processed by two ER-localized, membrane-bound enzymes. The first prenyl-dependent processing step is the proteolytic removal of the –aaX tripeptide by the CaaX protease



Fig. 2. Farnesylation, proteolysis and carboxymethylation of a CaaX protein. The farnesyl donor is farnesyl pyrophosphate.

Rcel; this is followed by carboxymethylation of the now C-terminal prenylcysteine residue by the methyltransferase Icmt (Fig. 2). The result of these modifications is to produce a protein that exhibits a high affinity for cellular membranes and also to impart a unique structure at the C-terminus that can serve as a specific recognition motif in certain protein–protein interactions [4].

The importance of prenylation in CaaX protein function, most notably as a regulator of the oncogenic potential of the Ras proteins [2], has led to considerable efforts to identify inhibitors of the prenyltransferases involved for evaluation as therapeutic agents [5,6]. The majority of these studies have focused on FTase, since this enzyme modifies Ras proteins, and early preclinical studies indicated significant anticancer potential for FTase inhibitors. A wide variety of FTase inhibitors have been developed, including some very promising ones that possess antitumor activity in animal models and are now in clinical development [6]. In addition, the success of FTase inhibitors in preclinical models of tumorigenesis, the increasing realization that proteins modified by GGTase-I play important roles in oncogenesis, and the finding that post-prenylation processing by Rce1 is important in the function of Ras and other CaaX proteins has led to the current situation in which all of the enzymes involved in CaaX protein processing are viewed as potential therapeutic targets.

2.1. The CaaX prenyltransferases FTase and GGTase-I

FTase is a heterodimer consisting of 48-kDa (α) and 46-kDa (β) subunit polypeptides. GGTase-I also consists of two subunits, a 48-kDa α subunit shared with FTase, and a 43-kDa β subunit. The isoprenoid substrates for the two enzymes are farnesyl pyrophosphate and geranylgeranyl pyrophosphate. Protein substrates for FTase in mammalian cells include Ras GTPases, lamin B, several proteins involved in visual signal transduction and at least three protein kinases and phosphatases. Known targets of GGTase-I include most γ subunits of heterotrimeric G proteins and Ras-related GTPases such as members of the Ras and Rac/Rho families. Both FTase and GGTase-I recognize short peptides containing appropriate CaaX motifs, and tetrapeptide substrates were instrumental in purifying the enzymes to homogeneity.

Both FTase and GGTase-I are zinc metalloenzymes in which the single bound zinc atom participates directly in catalysis. FTase additionally requires high concentrations (>1 mM) of magnesium for catalysis. FTase proceeds via a functionally ordered kinetic mechanism, with farnesyl pyrophosphate binding first to create an FTase-farnesyl pyrophosphate binary complex that then reacts rapidly with a CaaX substrate to form a prenylated product. In the absence of excess substrate, the dissociation rate is so slow that FTase-product complexes can be isolated. A wealth of structural information has emerged for FTase in the past several years beginning with the first X-ray crystal structure of unliganded FTase solved at 2.2 Å resolution (Park, 1997), greatly enhancing the ability of investigators to conduct structure-function analyses on the enzyme and investigating the roles of specific residues in substrate binding and catalysis.

3. Myristoylation

N-Myristoylated proteins comprise a large family of functionally diverse eukaryotic and viral proteins. Myristate, a relatively rare 14-carbon, saturated fatty acid, is transferred from myristoyl-CoA and linked via an amide bond to the N-terminal glycine residue of the target protein. The reaction is catalyzed by the enzyme myristoyl-CoA; protein Nmyristoyltransferase (NMT) which recognizes and modifies N-terminal glycine residues presented in a particular sequence context. Although myristate can be attached posttranslationally to N-terminal glycine in synthetic peptides of the appropriate sequence, in vivo myristoylation is an early co-translational event occurring in the cytoplasm as soon as ~ 60 amino acids of the nascent peptide emerge from the ribosomal tunnel and after the N-terminal glycine residue is made available by cellular methionyl-aminopeptidases that remove the initiator methionine residue. The myristoyl-CoA pools used by NMT appear to be supplied by de novo synthesis and by activation of exogenous myristate. Of 6220 open reading frames surveyed in the Saccharomyces cerevisiae genome data base for an appropriately positioned glycine residue, 70 (1.1%) are known to be or predicted to be N-myristoylated proteins. Some of these proteins play critical roles in cell survival since NMT is an essential enzyme in S. cerevisiae. NMT is also essential in Candida albicans, and Cryptococcus neoformans, the most common causes of systemic fungal infections in immunocompromised individuals.

3.1. N-Myristoyltransferase (NMT)

NMT was first purified from *S. cerevisiae* and characterized as a \sim 55-kDa monomeric protein with no apparent cofactor requirements. The crystal structures of the *S. cerevisiae* and *C. albicans* NMTs have been determined, the former as a co-crystal with a non-hydrolyzable myristoyl-CoA derivative and a dipeptide inhibitor of the enzyme [7].

NMT uses an ordered Bi–Bi reaction mechanism: myristoyl-CoA binds first, followed by the peptide substrate. After catalysis, CoA is discharged first, followed by the *N*-myristoyl peptide. The enzyme is highly selective for myristoyl-CoA and for polypeptide substrates with an N-terminal glycine, an uncharged residue at position 2, and neutral residues at positions 3 and 4. Serine is found at position 5 of all known yeast *N*-myristoyl proteins, while lysine is commonly found at position 6. In vitro analyses of human and fungal NMTs indicate that while their reaction mechanism and acyl-CoA substrate specificities are the same, their peptide specificities are different — this difference has been exploited to develop species-selective NMT inhibitors that act as fungicidal agents.

3.2. Myristoyl switches to regulate protein function

The myristoyl group in *N*-myristoylated proteins frequently acts as a key regulator of protein function. In some cases, the myristate residue provides a constitutive source of membrane affinity that needs to be supplemented by a second interaction between the protein and the membrane in order for the protein to stay membrane associated. For the MARCKS protein (myristoylated alanine-rich C-kinase substrate) as well as the tyrosine kinase Src, this second interaction is provided by electrostatic affinity

between a polybasic region of the protein and the negatively charged headgroups of phospholipids in the cytoplasmic leaflet of cell membranes. When serine residues in the polybasic region of MARCKS or Src are phosphorylated, the electrostatic contribution to membrane binding is reduced, and the protein moves off the membrane into the cytoplasm.

Myristate can also provide a regulated source of membrane affinity. Some proteins such as ARF (ADP ribosylation factor) and recoverin exist in alternate conformations in which the myristoyl group is exposed and available for membrane binding, or sequestered within a hydrophobic pocket in the protein. On ligand binding (GTP for ARF, and Ca^{2+} for recoverin), the myristoyl group is exposed and becomes available to promote interactions with target membranes and or protein partners.

4. Protein thioacylation

Thioacylated proteins contain fatty acids in thioester linkage to cysteine residues [8,9]. This class of lipid modification of proteins was first identified in studies of brain myelin protein, but only firmly established in the late 1970s when Schlesinger, Schmidt and co-workers reported the palmitoylation of Sindbis virus and vesicular stomatitis virus glycoproteins. Protein thioacylation is frequently referred to as palmitoylation, although fatty acids other than palmitate may be found on thioacylated proteins. Membrane proteins as well as hydrophilic proteins are thioacylated, the latter, in many cases, acquiring the modification when they become associated with a membrane compartment as a result of N-myristoylation or prenylation. Thioacylated cysteine residues are found in a variety of sequence contexts and are invariably located in portions of the protein that are cytoplasmic or within a predicted transmembrane domain. Unlike the other known lipid modifications of proteins, thioacylation is reversible: the protein undergoes cycles of acylation and deacylation, and as a result, the half-life of the acyl group is much shorter than that of the polypeptide (~ 20 min for the acyl group versus ~ 1 day for the polypeptide in the case of N-Ras (Magee, 1987)). Several protein acyltransferases have been isolated, but it is unclear whether they are required for thioacylation in living cells: non-enzymatic thioacylation has been observed in vitro suggesting that acyl transfer may occur through an autocatalytic process. In contrast, a number of thioacyl protein thioesterases have been identified and these appear to be responsible for the deacylation of thioacylated proteins.

4.1. Examples of thioacylated proteins

There are three classes of thioacylated proteins [9]: polytopic membrane proteins such as some G-protein coupled receptors (β -adrenergic receptor, rhodopsin), monotopic membrane proteins including viral glycoproteins, the transferrin receptor, and the cation-dependent mannose-6-phosphate receptor, and hydrophilic proteins such as members of the Src family of protein tyrosine kinases (e.g., p59^{fyn} and p56^{lck}) [10], as well as H-Ras, N-Ras, and the synaptic vesicle protein SNAP-25. The functional significance of thioacylation of polytopic and monotopic membrane proteins is unclear, but the acyl

modification may dictate how the protein is trafficked between membranes and whether it partitions into sterol/sphingolipid-rich membrane domains (lipid rafts) (Chapter 1) (Melkonian, 1999). For intrinsically hydrophilic proteins such as p59^{fyn} and SNAP-25, the function of thioacylation is clear — the acyl group functions together with other acyl groups, or other lipid modifications, to provide membrane anchoring in the cytoplasmic leaflet of the membrane bilayer for an otherwise soluble protein.

4.2. Membrane anchoring of thioacylated proteins: the need for multiple lipid modifications and the role of dynamic thioacylation

Thioacylated proteins that lack transmembrane spans must have more than one covalently bound lipid chain in order to be stably associated with membranes. This is also true for *N*-myristoylated proteins and prenylated CaaX proteins that, when newly synthesized, are modified by only a single lipid moiety. The association of monoacylated or monoprenylated proteins with the lipid bilayer appears to be rapidly reversible and thus, for stable membrane association, lipid modified proteins require at least two lipid chains or must rely on some other interaction with membranes such as recognition by a membrane-bound receptor or electrostatic interaction between charged amino acids in the protein and charged phospholipids in the membrane [10].

The *N*-myristoylated Src family protein tyrosine kinases are frequently thioacylated at one or more cysteine residues near the myristoylated glycine, and it is these doubly or triply lipid modified proteins that are found associated with the cytoplasmic face of the plasma membrane [10] (the N-terminal sequences of p59^{fyn} and p56^{lck} are ¹MGCVC- and ¹MGCVQC-, respectively, where the *N*-myristoylated glycine is shown in bold and the thioacylated cysteine residues are underlined). A similar situation is seen for the prenylated Ras proteins which must be thioacylated before they associate firmly with membranes (the C-terminal sequences of H-Ras and N-Ras are -GCMSCKCVLS-COOH and -GCMGLPCVVM-COOH, respectively (the farnesylated cysteine is shown in bold and the thioacylated cysteines are underlined)). A third example is provided by proteins such as SNAP-25 which are exclusively thioacylated, but display at least four thioacyl chains through which they become stably associated with the synaptic vesicle membrane.

An interesting and persuasive model for the targeting of lipid modified proteins to particular membranes was suggested by Shahinian and Silvius [11] (Fig. 3), based on the notion that single lipid modifications allow the protein to undergo transient interactions with a variety of intracellular membranes whereas tandem modifications promote stable membrane association. Thus, a protein with a single lipid modification such as the cytoplasmically synthesized *N*-myristoylated $p59^{fyn}$ becomes stably associated with the cytoplasmic face of the plasma membrane only when it becomes thioacylated at this location. In this scenario, thioacylation would not only provide for stable membrane association of a protein with a single lipid modification, but it would also ensure targeting of that protein to the membrane where thioacylation occurs. Thus, within 5 min of the completion of peptide synthesis, $p59^{fyn}$ becomes *N*-myristoylated, thioacylated and located to the cytoplasmic face of the plasma combrane association. Thus, respectively associated to the cytoplasmic face of the plasma combrane where thioacylation occurs. Thus, within 5 min of the palmitoylation sites slows the kinetics of membrane (van 't Hof, 1997). Removal of the palmitoylation sites slows the kinetics of membrane association,



Fig. 3. Bilayer trapping mechanism for membrane targeting of lipid-anchored proteins lacking transmembrane spans (redrawn from [11]). A singly lipid-modified protein associates transiently with a membrane containing a membrane targeting receptor, possibly a thioacyltransferase. Thioacylation at the membrane yields a dual-anchored species that is stably associated with the membrane until it is deacylated.

and reduces the proportion of $p59^{fyn}$ that is membrane associated. The dynamic nature of thioacylation would suggest that the duration of association of $p59^{fyn}$ with the plasma membrane would be dictated by the half-life of the thioacyl chain. However, since $p59^{fyn}$ is doubly thioacylated, it is unlikely that it would revert to its solely *N*-myristoylated state and re-enter the cytoplasm.

Thioacylation frequently dictates plasma membrane targeting of proteins lacking transmembrane spans. In the case of $p59^{fyn}$, targeting occurs directly, with the *N*-myristoylated protein becoming thioacylated and plasma membrane associated rapidly upon completion of synthesis. In contrast, $p56^{lck}$ appears to be thioacylated on intracellular membranes and arrive at the plasma membrane via vesicular transport (bound to the cytoplasmic face of secretory vesicles) (Bijlmakers, 1999). In yet another targeting variation, newly synthesized *N*-myristoylated $G_{z\alpha}$, a dually acylated trimeric G protein α subunit, associates with all cellular membranes but accumulates eventually at the plasma membrane: the plasma membrane form is the only one that is both *N*-myristoylated and thioacylated.

4.3. Thioesterases

Three thioacyl protein thioesterases have been identified [9]. Two of these, PPT1 (Camp, 1994) and PPT2 (Soyombo, 1997) (palmitoyl-protein thioesterases 1 and 2) are localized to the lysosomes and are thus likely to be involved in the catabolism of thioacylated proteins or peptides. The thioacylated molecules are presumed to gain access to the lysosomal lumen by an autophagic pathway in which membrane fragments are captured into a vacuole that subsequently fuses with lysosomes. A defect in PPT1 leads to a severe neurodegenerative disorder termed infantile neuronal ceroid lipofuscinosis characterized by the accumulation of autofluorescent material (including lipid thioesters) in all tissues. No diseases have been linked to PPT2.

A third thioesterase was purified from rat liver cytosol using palmitoylated G-protein α subunit as a substrate (Duncan, 1998). This thioesterase, a 25-kDa monomeric protein, is likely to be the one involved in turnover of thioacyl groups on proteins. It

displays both acylprotein thioesterase activity as well as lysophospholipase activity, but thioacylproteins are by far the preferred substrate.

5. Cholesterol modification

In addition to its numerous roles in membrane architecture and steroid and bile acid synthesis (Chapters 15 and 16), cholesterol has recently been discovered as a post-translational protein modification on a family of signaling proteins referred to as hedgehog (Hh) proteins [12]. This modification was identified in studies on the processing of the *Drosophila* Hh protein (Porter, 1996). During biosynthesis, the Hh protein undergoes an autocatalytic cleavage event during which the carboxy-terminal domain is removed and cholesterol is covalently attached to the amino-terminal domain yielding an active signaling molecule (Fig. 4). The Hh proteins, found in insects, vertebrates, and other multicellular organisms, are part of a family of secreted signaling molecules involved in the patterning of diverse tissues during development. These proteins function



Fig. 4. Modification of hedgehog with ester-linked cholesterol and an N-terminal amide-linked fatty acid (redrawn from [12]). (A) Overall processing steps leading to lipid-modified Hh. (B) Mechanism of cholesterol addition. Residues indicated as B1 and B2 in panel B are catalytic bases presumably contained in the carboxy-terminal domain of the Hh precursor.

in a graded manner, emanating from cells in which the protein is produced and acting upon target cells several cell diameters away, thus specifying different cell fates by altering programs of gene expression. Cholesterylation of the Hh signaling protein is crucial for its proper targeting to specific tissues during development [13,14].

5.1. Addition of cholesterol to hedgehog proteins

Newly synthesized Hh precursor proteins of approximately 45 kDa contain an Nterminal signal sequence which targets the protein to the ER and is subsequently trimmed, followed by an N-terminal signaling domain (~19 kDa), an absolutely conserved Gly–Cys–Phe tripeptide motif, and a C-terminal processing domain which is removed to yield the active signaling molecule (Fig. 4). Addition of cholesterol to Hh proteins proceeds via an autoproteolytic internal cleavage reaction at the Gly–Cys–Phe tripeptide sequence. The cleavage is initiated through a nucleophilic attack by the thiol side chain of Cys on the Gly carbonyl, replacing the Gly–Cys peptide bond with a thioester linkage. A second nucleophilic attack on the same carbonyl by the hydroxyl group of cholesterol results in removal of the C-terminal processing domain and produces the active Hh signaling molecule containing a C-terminal cholesterol modification (Fig. 4). Although other sterols have been shown to be able to substitute for cholesterol in in vitro Hh processing assays, the free 3β -hydroxyl group is essential for processing [12].

The autocleavage reaction is mediated by the C-terminal processing domain of the Hh precursor polypeptide. The first 145 amino-terminal residues of C-terminal processing domain are sufficient for thioester formation as indicated by the ability of Hh proteins truncated after this point to undergo cleavage in the presence of dithiothreitol. At least some part of the remaining carboxy-terminal residues is required for cholesterol addition and contains a sterol recognition region (Hall, 1997).

In addition to the cholesterol modification, Hh proteins can also be palmitoylated at the N-terminal Cys residue (part of a CGPGR motif) after cholesterol is added. An acyltransferase, named Skinny Hh, was discovered in *Drosophila*. Skinny Hh mutant cells express Hh proteins containing the cholesterol modification but lacking aminoterminal palmitate (Chamoun, 2001). Available information indicates that the palmitate residue is amide-linked to the N-terminus of the protein and not thioesterlinked to the cysteine side chain.

5.2. Biological significance of cholesterol modification

The role of cholesterol-anchoring of the hedgehog signaling molecule is not fully understood. Cholesterylated Hh must travel to other cells to elicit its effect, even though the hydrophobic cholesterol molecule is presumed to act as a membrane anchor. Hh proteins truncated at the cleavage site are not modified by cholesterol, but nevertheless produce active signaling molecules; however, studies in *Drosophila* have shown that these proteins are not targeted properly and cause mispatterning and lethality in embryos. The subcellular location of the autocatalytic cleavage and processing events is not clear. Cholesterol-modified Hh in *Drosophila* partitions into detergent insoluble membranes, or rafts, which are enriched in specific molecules such as sterols, sphingolipids, thioacylated proteins and GPI-anchored proteins and usually considered predominantly plasma membrane domains (Chapter 1) (Rietveld, 1999). Several proteins, each containing a sterol-sensing domain, have been identified that function in the release of Hh from Hh-producing cells and proper sequestering of Hh in target cells. For example, the proteins Patched and Dispatched are involved in sequestration of the Hh signal within tissues and release of functional signal from Hh-producing cells, respectively. A soluble form of mammalian Hh (sonic Hh) that is cholesterol-modified, multimeric and biologically potent has been isolated from developing chick limb (Zeng, 2001). The protein Patched is required for release of this soluble complex. A proposed model is that cholesterol-modified Hh protein is recruited

to lipid rafts (exoplasmic leaflet of the plasma membrane) where multimerization occurs

and the soluble complex can then travel to target cells and elicit its effect.

6. GPI anchoring of proteins

Glycosylphosphatidylinositols (GPIs) are a ubiquitous family of eukaryotic glycolipids. GPIs were originally discovered covalently linked to cell surface glycoproteins and recognized to be an important alternative mechanism for anchoring proteins to the cell surface [15]. After synthesis of GPI anchors and attachment to protein in the endoplasmic reticulum (ER), GPI-anchored proteins are transported via the secretory pathway to the cell surface. GPI anchors are found on a variety of functionally diverse proteins and glycoconjugates including cell surface receptors (folate receptor, CD14), cell adhesion molecules (NCAM isoforms, carcinoembryonic antigen variants), cell surface hydrolases (5'-nucleotidase, acetylcholinesterase), complement regulatory proteins (decay accelerating factor), and protozoal surface molecules (*Trypanosoma brucei* variant surface glycoprotein, *Leishmania* lipophosphoglycan). Along with serving to attach proteins to the cell surface, GPI-anchored proteins appear to be markers and major constituents of 'detergent-resistant' lipid rafts, the sphingolipid- and sterol-rich domains in membranes that are postulated to play an important role in the activation of signaling cascades [16].

6.1. Biosynthesis of GPI

GPIs are assembled via a biosynthetically and topologically complex metabolic pathway that is comprised of at least 10 steps and requires the participation of at least 20 distinct gene products (reviewed in [17–19]). In basic terms, synthesis occurs in the ER and involves the sequential addition of monosaccharides to phosphatidylinositol (PI) yielding the core GPI structure of ethanolamine-*P*-6Man α 1–2Man α 1–6Man α 1–4GlcN α 1–6*myo*-inositol-P-lipid which is then added to proteins. Many of the genes encoding enzymes involved in synthesis have been identified. Fig. 5 illustrates the sequence of reactions involved in the assembly of a GPI protein anchor precursor in mammalian cells; variations of this general sequence are found in all organisms studied thus far.



Fig. 5. GPI biosynthesis in the ER (redrawn from [17]). The pathway is initiated on the cytoplasmic face of the ER and proceeds clockwise in the figure to yield the lumenally oriented mature GPI structure H7 (mammalian cells synthesize another mature GPI termed H8 (not depicted) which is identical to H7 except that it contains an additional EtNP residue linked to the second mannose). H7 is attached to ER-translocated proteins bearing a C-terminal, GPI-directing signal sequence. MAM, mitochondria-associated membrane. See text for other abbreviations.

Initiation of GPI biosynthesis occurs on the cytoplasmic leaflet of the ER via the addition of *N*-acetylglucosamine (GlcNAc) from UDP-GlcNAc to PI yielding GlcNAc-PI (Fig. 5, step 1). The enzyme which mediates this first step is a complex consisting of at least six proteins in mammalian cells: PIG-A, PIG-C, PIG-H, GPI1, PIG-P, and DPM2; however, the subunit stoichiometry and functions of most of the subunits are not yet known. PIG-A which has homology to many glycosyltransferases is most likely the catalytic component. PIG-C, PIG-H, and PIG-P are essential components of the complex but their precise functions are unknown. GPI1 is important but not essential for activity and DPM2 appears to be a positive regulator stimulating GlcNAc transferase activity approximately 3-fold.

GlcNAc-PI is subsequently de-*N*-acetylated to generate GlcN-PI by the GlcNAc-PI de-N-acetylase, PIG-L (Fig. 5, step 2). The third step in biosynthesis is the acylation, predominantly palmitoylation, of the inositol residue of GlcN-PI at the 2 position (Fig. 5, step 3). The acylation step is required prior to the addition of three mannose residues (four in yeast) to GlcN-(acyl)PI and modification of the mannose residues with phosphoethanolamine (EtNP) side chains. In contrast to the first two steps of synthesis which occur on the cytoplasmic leaflet of the ER, the mannosylation reactions are thought to occur lumenally based on the predicted topology of the mannosyl transferases. This indicates that the substrate (GlcN-PI or GlcN-(acyl)PI must be flipped across the ER membrane bilayer (Fig. 5, step 4).

Dolichol-P-mannose serves as the mannose donor for all three mannosylation steps. The first mannosylation is catalyzed by PIG-M, the GPI a1-4-mannosyltransferase (Fig. 5, step 5). Phosphoethanolamine is then added to the 2 position of the first mannose residue (Fig. 5, step 6); PIG-N is involved in this reaction and is probably the EtNP transferase. A second mannose residue is then added by an α 1–6-mannosyltransferase that has not yet been identified (Fig. 5, step 7). An $\alpha 1-2$ -mannosyltransferase, most likely the protein PIG-B, mediates transfer of the third mannose residue to the GPI moiety (Fig. 5, step 8). EtNP residues can be added to both the second and third mannose residues as well (Fig. 5, steps 9 and 10). PIG-O and PIG-F are the gene products responsible for transferring EtNP from phosphatidylethanolamine (PE) to the 6 position on the third mannose (Fig. 5, step 10). These proteins form a complex, with PIG-O most likely serving the catalytic role. This modification is crucial as it is this EtNP residue on the third, or terminal mannose that is involved in the linkage of GPI to the carboxy-terminus of proteins destined to be GPI-anchored (Fig. 5, step 11). Fully assembled GPI structures, or the GPI moiety in GPI-anchored proteins, are frequently subject to lipid re-modeling reactions in which fatty acids or the entire lipid structure are replaced with different fatty acids or lipids.

6.2. Subcellular location and membrane topology of GPI biosynthesis

GPI biosynthesis occurs in the ER, however the ER is an extremely heterogeneous organelle and GPI biosynthetic activities are not uniformly distributed throughout the bulk ER [19]. The capacity to synthesize GlcNAc-PI appears to be uniformly distributed in the ER in mammalian cells while subsequent biosynthetic steps leading to the formation of (Etn-P)Man-GlcN-PI are highly enriched in an ER domain that is associated with mitochondria (mitochondria-associated ER membrane domain) (Fig. 5). The reason for segregation of some GPI biosynthetic reactions to the mitochondria-associated membrane is not known.

In addition to spatial segregation of GPI biosynthetic activities within subdomains of the ER, the biosynthetic pathway is also topologically complex with the first two steps occurring in the cytoplasmic leaflet of the membrane and later mannosylation steps and attachment of GPI to protein occurring in the lumenal leaflet (Vidugiriene, 1993, 1994). Thus, for GPI synthesis to go to completion GPI precursors must be flipped across the ER bilayer to access the enzymes involved in the late steps of synthesis [19]. The point at which this flipping occurs as well as the mechanism for flipping GPI precursors remain to be found.

6.3. Attachment of GPIs to proteins

Newly synthesized proteins are attached to pre-existing GPIs in the lumenal leaflet of the ER by a GPI: protein transamidase complex [20,21]. Proteins destined to be GPI-anchored contain an N-terminal signal sequence for targeting to the ER and a C-terminal signal sequence that directs GPI-anchoring. The N-terminal signal sequence is cleaved by signal peptidase during or after translocation of the nascent polypeptide into the ER lumen. The C-terminal GPI signal sequence is then



Fig. 6. GPI anchoring of proteins. (A) The processing steps involved in the conversion of a pre-pro-protein to a GPI-anchored protein. (B) The transamidation reaction in which a GPI anchor is attached to a pro-protein. The transamidase enzyme, which contains an enzymatically critical cysteine residue, is depicted as Enz-S⁻. The ω amino acid is depicted with a side chain labeled R1; the $\omega + 1$ amino acid has a side chain labeled R2.

cleaved and replaced with a GPI anchor through the action of GPI transamidase (Fig. 6).

The C-terminal GPI-signal peptide, which is cleaved at what is referred to as the ω site, is necessary and sufficient for designating that a protein becomes GPIanchored. The sequence motif for GPI-anchoring includes four sequence elements: (1) an unstructured linker region of about 11 residues ($\omega - 11$ to $\omega - 1$); (2) four preferably small amino acids ($\omega - 1$ to $\omega + 2$) including the cleavage site; (3) a moderately polar spacer region ($\omega + 3$ to $\omega + 9$); and (4) a hydrophobic tail from $\omega + 9$ or $\omega + 10$ to the C-terminal end [21].

The GPI transamidase complex responsible for attaching the protein to the GPIanchor consists of at least five subunits: Gpi8p, Gaa1p, PIG-S, PIG-T and PIG-U (Vidugiriene, 2001; Ohishi, 2001; Fraering, 2001; Vainauskas, 2002). The enzyme recognizes a pro-protein containing a GPI-anchoring signal sequence, cleaves the signal sequence and attaches the new carboxy terminus to the terminal EtNP of a pre-existing GPI. Gpi8p is a membrane protein with a large lumenal domain and has homology to a family of cysteine proteases, one of which is a plant endopeptidase with transamidase activity. Cysteine and histidine residues conserved between the cysteine proteases and Gpi8p are essential for GPI transamidase activity, thus it is proposed that Gpi8p cleaves the signal sequence and functions in the transamidation reaction. The functions of the remaining subunits are unknown.

6.4. GPI-anchoring in mammals, parasitic protozoa and yeast

GPI-deficient mammalian cells are viable in tissue culture, but a GPI defect has clear consequences for multicellular organisms. Transgenic mouse embryos lacking the ability to initiate GPI biosynthesis (defective in PIG-A) do not develop beyond the ninth day of gestation [18]. The inability of certain blood cells to express GPI-anchored proteins results in the rare human disease, paroxysmal nocturnal hemoglobinuria (PNH), characterized by intravascular hemolysis, thrombosis and bone marrow failure. The disease is caused by a somatic mutation of the X-linked PIG-A gene (encoding the catalytic component of the enzyme responsible for the first step in GPI biosynthesis, see above) in hematopoietic stem cells. Cells defective in PIG-A are either unable to synthesize GPIs or exhibit a significant decrease in the synthesis of GPIs and thus decreased expression of GPI-anchored proteins. Red blood cells no longer expressing GPI-anchored complement regulatory proteins (e.g. CD55 and CD59) become susceptible to complement-mediated lysis, resulting in the release of heme and hemoglobin into the blood, filtering by the kidney, and excretion in the urine. Interestingly, while the GPI-deficient phenotype of hematopoietic stem cells makes the cells more sensitive to complement-mediated lysis, the defective PNH clone still persists. This implies that there must be other factors which select for clonal expansion and maintenance of GPI-deficient blood cells. One proposal is that PNH patients possess autoreactive T cells that target GPI on the surface of hematopoietic stem cells, thus PNH cells would evade damage because they lack surface GPI molecules [22].

GPI anchoring is the most prominent mode of attachment for cell surface proteins and glycans in parasitic protozoa [23]. Pathogenic protozoa, including species of the genera *Trypanosoma*, *Leishmania*, and *Plasmodium*, display abundant GPI-anchored cell surface macromolecules that play crucial roles in parasite infectivity and survival. An example is the GPI-anchored variant surface glycoprotein (VSG) of bloodstream forms of *Trypanosoma brucei*, the causative agent of African sleeping sickness.

GPIs serve a unique function in yeast (*S. cerevisiae*). In addition to anchoring secretory proteins to the cell surface, GPIs play an additional role in yeast cell wall biosynthesis [24]. The yeast cell wall consists of a fibrous lattice of mannoproteins, β 1,3-glucan, β 1,6-glucan and chitin. During cell wall biosynthesis GPI-anchored mannoproteins are transported through the secretory pathway to the cell surface. After arrival at the plasma membrane, a transglycosylation reaction (catalyzed by an unknown enzyme) results in cleavage of the GPI moiety between GlcN and the first mannose residue and formation of a glycosidic linkage between the mannoprotein-GPI-remnant and β 1,6-glucan. Mutations in GPI biosynthesis are lethal in yeast and decreased levels of GPI biosynthesis cause growth defects and aberrant cell wall biogenesis. However, not all GPI-anchored proteins become crosslinked to β 1,6-glucan; the amino acids within four or five residues upstream of the ω -site determine whether the protein becomes incorporated into the cell wall or remains anchored to the plasma membrane. A unique characteristic of yeast GPIs is the addition of a fourth mannose residue to the core GPI structure (via an α 1–2

linkage to the third mannose). Most GPI-anchored proteins in yeast also undergo lipid remodeling replacing the glycerolipid backbone with ceramide. The remodeling occurs after the attachment of protein to the GPI and can occur in both the ER and the Golgi.

6.5. Functions of GPI anchors

GPIs function as membrane anchors for secretory proteins and are believed to provide targeting signals that influence the intracellular trafficking of these proteins. Recent results indicate that GPI-proteins are packaged into unique transport vesicles, distinct from those carrying other secretory proteins, for export from the ER (Muñiz, 2001). GPI-anchored proteins have been shown to coalesce with sphingolipids and cholesterol into detergent insoluble membrane domains, or lipid rafts. Association of molecules in rafts at the plasma membrane, including GPI-anchored proteins (in the exoplasmic leaflet) and acylated signaling molecules (in the cytoplasmic leaflet), is postulated to play an important role in the activation of signaling cascades [16]. In some polarized epithelial cells, GPI-anchored proteins and many glycosphingolipids are sorted in the trans Golgi network and specifically targeted to the apical membrane. Once at the plasma membrane GPI-anchored proteins can undergo endocytosis and recycling back to the cell surface. However, uptake of GPI-anchored proteins is approximately fivetimes slower than that of receptor-mediated endocytosis and recycling to the plasma membrane is about three-times slower than that of recycling receptors (Sabharanjak, 2002). The current view is that GPI-anchoring of proteins directs their segregation into lipid rafts and affects their sorting in both the exocytic and endocytic pathways.

A GPI-anchor may also allow a protein to be selectively released from the cell surface upon hydrolysis by a GPI-specific phospholipase (e.g., PI-PLC or GPI-PLD). This has been shown to occur for certain GPI-anchored proteins in mammalian cell culture. One example is GPI-anchored membrane dipeptidase which is released from the adipocyte cell surface by a phospholipase C in response to insulin (Movahedi, 2000). Interestingly, other GPI-anchored proteins are not released indicating a level of regulation in insulin-stimulated hydrolysis of GPI-anchored proteins. GPI-anchored molecules have also been to shown to transfer between cells and stably insert in the external leaflet of the acceptor cell's plasma membrane (Low, 1998). The biological significance of this is unclear; however, the ability of GPI-anchored proteins to transfer between cells has implications for the expression of foreign proteins on the cell surface.

7. Future directions

This chapter describes post-translational lipid modifications of proteins that represent functionally critical elaborations of protein structure. These modifications, with the possible exception of thioacylation, can be anticipated at the level of primary protein sequence, enabling predictions about the localization and likely behavior of the modified proteins within cells. Although the functional diversity of lipid-modified proteins makes it difficult to arrive at generalizations about the evolutionary impetus for lipid anchoring compared to the use of 'conventional' protein transmembrane domains, the observation that many lipid modified proteins are associated with cholesterol and sphingolipid-rich membrane domains (rafts) is suggestive. Analyses of the biophysical characteristics of membrane association of lipid modified proteins, the role of raft-associated proteins in cell signaling, and the regulation of protein function by lipid modification are likely to remain fruitful areas of investigation.

Although many biochemical and cell biological aspects of the various lipid modifications remain to be elucidated, it is worth highlighting the scope for investigation offered by the GPI biosynthetic pathway. GPI anchoring is arguably the most biosynthetically and topologically complex of the lipid modifications described in this chapter. Despite the identification of numerous genes and gene products associated with the GPI biosynthetic pathway, the enzymology of GPI biosynthesis, including analyses of enzyme structure, basis for enzyme localization to the ER and ER domains, and transbilayer distribution of the pathway remains open to new investigation.

Abbreviations

ER	endoplasmic reticulum
EtN	ethanolamine
FTase	protein farnesyltransferase
FTI	FTase inhibitor
GGTase	protein geranylgeranyltransferase
GlcN	glucosamine
GlcNAc	N-acetylglucosamine
GPI	glycosylphosphatidylinositol
G proteins	GTP-binding proteins
Hh	hedgehog protein
Man	mannose
NMT	N-myristoyltransferase
PE	phosphatidylethanolamine
PI	phosphatidylinositol
PNH	paroxysmal nocturnal hemoglobinuria

References

- Zhang, F.L. and Casey, P.J. (1996) Protein prenylation: molecular mechanisms and functional consequences. Annu. Rev. Biochem. 65, 241–269.
- 2. Magee, T. and Marshall, C. (1999) New insights into the interaction of ras with the plasma membrane. Cell 98, 9–12.
- 3. Pereira-Leal, J.B., Hume, A.N. and Seabra, M.C. (2001) Prenylation of rab GTPases: molecular mechanisms and involvement in genetic disease. FEBS Lett. 498, 197–200.
- Gelb, M.H. (1997) Protein prenylation signal transduction in two dimensions. Science 275, 1750– 1751.
- 5. Gelb, M.H., Scholten, J.D. and Sebolt-Leopold, J.S. (1998) Protein prenylation: from discovery to prospects for cancer treatment. Curr. Opin. Chem. Biol. 2, 40–48.

- Johnston, S.R. (2001) Farnesyltransferase inhibitors: a novel targeted therapy for cancer. Lancet Oncol. 2, 18–26.
- 7. Farazi, T.A., Waksman, G. and Gordon, J.I. (2001) The biology and enzymology of protein Nmyristoylation. J. Biol. Chem. 276, 39501-39504.
- Mumby, S.M. (1997) Reversible palmitoylation of signaling proteins. Curr. Opin. Cell Biol. 9, 148– 154.
- Linder, M.E. (2000) Reversible modification of proteins with thioester-linked fatty acids. In: F. Tamanoi and D.S. Sigman (Eds.), The Enzymes, Volume XXI 'Protein Lipidation'. Academic Press, New York, pp. 215–240.
- 10. Resh, M.D. (1999) Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. Biochim. Biophys. Acta 1451, 1–16.
- 11. Shahinian, S. and Silvius, J.R. (1995) Doubly-lipid-modified protein sequence motifs exhibit long-lived anchorage to lipid bilayer membranes. Biochemistry 34, 3813–3822.
- 12. Mann, R.K. and Beachy, P.A. (2000) Cholesterol modification of proteins. Biochim. Biophys. Acta 1529, 188-202.
- 13. Incardona, J.P. and Eaton, S. (2000) Cholesterol in signal transduction. Curr. Opin. Cell Biol. 12, 193-203.
- 14. Ingham, P.W. (2000) Hedgehog signaling: how cholesterol modulates the signal. Curr. Biol. 10, R180--R183.
- Low, M.G., Ferguson, M.A.J., Futerman, A.H. and Silman, I. (1986) Covalently attached phosphatidylinositol as a hydrophobic anchor for membrane proteins. Trends Biochem. Sci. 11, 212– 215.
- 16. Simons, K. and Toomre, D. (2000) Lipid rafts and signal transduction. Nat. Rev. (Mol. Cell Biol.) 1, 31–39.
- 17. Kinoshita, T. and Inoue, N. (2000) Dissecting and manipulating the pathway for glycosylphosphatidylinositol-anchor biosynthesis. Curr. Opin. Chem. Biol. 4, 632–638.
- Tiede, A., Bastisch, I., Schubert, J., Orlean, P. and Schmidt, R.E. (1999) Biosynthesis of glycosylphosphatidylinositols in mammals and unicellular microbes. Biol. Chem. 380, 503–523.
- 19. McConville, M.J. and Menon, A.K. (2000) Recent developments in the cell biology and biochemistry of glycosylphosphatidylinositol lipids. Mol. Membr. Biol. 17, 1–16.
- Udenfriend, S. and Kodukula, K. (1995) How glycosyl-phosphatidylinositol-anchored proteins are made. Annu. Rev. Biochem. 64, 563–591.
- Eisenhaber, B., Bork, P. and Eisenhaber, F. (2001) Post-translational GPI lipid anchor modification of proteins in kingdoms of life: analysis of protein sequence data from complete genomes. Protein Eng. 14, 17–25.
- 22. Karadimitris, A. and Luzatto, L. (2001) The cellular pathogenesis of paroxysmal nocturnal hemoglobinuria. Leukemia 15, 1148–1152.
- 23. Ferguson, M.A.J. (2000) Glycosylphosphatidylinositol biosynthesis validated as a drug target for African sleeping sickness. Proc. Natl. Acad. Sci. USA 97, 10673–10675.
- 24. Lipke, P.N. and Ovalle, R. (1998) Cell wall architecture in yeast: new structure and new challenges. J. Bacteriol. 180, 3735–3740.