GPI anchoring of protein in yeast and mammalian cells or: How we learned to stop worrying and love glycophospholipids

Peter Orlean¹ and Anant K. Menon²
¹Dept of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61801 and
²Dept of Biochemistry, Weill Cornell Medical College, New York, NY 10021

Abbreviated title: GPI biosynthesis in yeast and mammals

Address correspondence to Peter Orlean at p-orlean@uiuc.edu or Anant Menon at akm2003@med.cornell.edu
Abstract

Glycosylphosphatidylinositol (GPI) anchoring of cell surface proteins is the most complex and metabolically expensive of the lipid post-translational modifications described to date. The GPI anchor is synthesized via a membrane-bound multi-step pathway in the endoplasmic reticulum requiring more than 20 gene products. The pathway is initiated on the cytoplasmic side of the ER and completed in the ER lumen, necessitating flipping of a glycolipid intermediate across the membrane. The completed GPI anchor is attached to proteins that have been translocated across the ER membrane and which display a GPI signal anchor sequence at their C-terminus. GPI-proteins transit the secretory pathway to the cell surface; in yeast many become covalently attached to the cell wall. Genes encoding proteins involved in all but one of the predicted steps in the assembly of the GPI precursor glycolipid and its transfer to protein in mammals and yeast have now been identified. Most of these genes encode polytopic membrane proteins, some of which are organized in complexes. The steps in GPI assembly, and the enzymes that carry them out, are highly conserved. GPI biosynthesis is essential for viability in yeast and for embryonic development in mammals. In this review we describe the biosynthesis of mammalian and yeast GPs, their transfer to protein, and their subsequent processing.

Supplementary key words:
glycosyltransferase, flippase, dolichol, phosphatidylethanolamine, phosphatidylinositol, endoplasmic reticulum, cell wall
Roughly 1% of all proteins encoded by eukaryotic genomes, or ~10-20% of all membrane proteins that enter the secretory pathway after being targeted to the endoplasmic reticulum (ER), are post-translationally modified at their C-terminus by glycosylphosphatidylinositol (GPI), a complex glycoprophospholipid that serves to anchor proteins to the cell surface. Although GPI proteins are functionally diverse, many have – or are predicted to have – hydrolytic activity, or serve as receptors or adhesive proteins. The importance of GPI anchoring in mammals is underscored by the facts that abrogation of GPI biosynthesis results in embryonic lethality (1) and that an inherited hypomorphic promoter mutation that lowers expression of the mannosyltransferase PIG-M and impairs GPI mannosylation leads to thromboses of the portal and hepatic veins as well as seizures (2). Further, a somatic defect in the first step of GPI biosynthesis in multipotent hematopoietic human stem cells results in paroxysmal nocturnal hemoglobinuria, an acquired hemolytic disease (3). GPIs are essential for viability in yeast and critical for maintenance of normal cellular morphology (4,5). They are also important for the cell surface display of a variety of proteins and glycoconjugates in parasitic protozoa like the African sleeping sickness parasite Trypanosoma brucei, and essential in certain life-cycle stages of these parasites (6-8).

Examples of GPI-proteins include cell surface receptors (e.g., folate receptor, CD14), cell adhesion molecules (e.g., NCAM isoforms, carboxyembryonic antigen variants, fasciclin I), cell surface hydrolases (e.g., 5'-nucleotidase, acetylcholinesterase, alkaline phosphatase), complement regulatory proteins (e.g., decay accelerating factor (CD55)), the scrapie prion, and protozoal coat proteins (e.g., Trypanosoma brucei variant surface glycoprotein). All these molecules require GPI in order to be expressed at the cell surface, and to function. Unlike conventional hydrophobic transmembrane protein domains that span the membrane bilayer, GPI-anchors penetrate only into a single membrane leaflet. This feature, as well as the saturated fatty acyl chains typically found in the GPI lipid portion, ensures the association of GPI-anchored proteins with lipid rafts, the nanoscale cholesterol and sphingolipid-rich membrane domains that are postulated to play an important role in membrane trafficking and cell signaling (9-11).

Unlike GPI, most other forms of protein lipidation are found on cytoplasmic proteins. Like myristoylation and prenylation, GPI attachment is irreversible. Unlike myristoyl, palmitoyl and prenyl groups, GPIs can be modified by the addition of side-branching groups to their glycan core, and their lipid moieties can be remodelled as well. These modifications can be species-specific. GPIs represent the sole means of membrane anchoring for GPI-proteins; in contrast, single myristoyl, palmitoyl and prenyl groups are insufficient as membrane anchors and their anchoring ability must be augmented by additional lipid modifications of the protein, or protein-protein or protein-membrane interactions, in order to provide stable membrane association (12, 13).

Because GPIs are preassembled stepwise by a series of enzymes located in the membrane of the ER, then transferred to protein in the lumen of that organelle, the GPI anchoring pathway is analogous to the dolichol pathway for asparagine-linked (N-linked) protein glycosylation (14). Both pathways are initiated on the cytoplasmic face of the ER and completed in the ER lumen necessitating transbilayer flipping of a glycolipid intermediate (15). The pathways share certain activated sugars (UDPGlcNAc, Dol-P-Man), and the same secretory protein can receive both types of posttranslational
modification; it is likely that all GPI-proteins are modified by N-linked glycans. Unlike the dolichol pathway, however, GPI assembly is also dependent on Ptd-Etn supply.

Genes for proteins involved in all but one of the predicted steps in the assembly of the GPI precursor glycolipid and its subsequent transfer to protein in mammals and yeast have now been identified. Most of these genes encode polytopic membrane proteins, some of which are organized as protein complexes. The steps in the GPI assembly pathway, and the enzymes that carry them out, are highly conserved. In this review we will describe the biosynthesis of mammalian and yeast GPIs, their transfer to protein, and their subsequent processing. For the most part we will not discuss the vast literature on protozoal GPI-anchored proteins and glycoconjugates; the reader is referred to other articles for information in this area (7, 16)

STRUCTURE OF GPIs AND SIGNALS FOR THEIR ATTACHMENT

Overview of GPI structure.

Detailed analyses of structures of protein-bound GPIs from mammals, protozoa, and yeast (17-22) reveal that they have a conserved core structure: protein-CO-NH₂-CH₂-CH₂-PO₄⁻-6-Manα₁,2Manα₁,6Manα₁,4GlcNα₁,6-myoinositol phospholipid (Fig. 1A). The GPI core in yeast and mammals is modified during its biosynthesis by the addition of a palmitoyl group to the 2-OH of the inositol moiety, which renders the GPI resistant to cleavage by PtdIns-specific phospholipase C (PI-PLC), and by the addition of up to three ethanolamine phosphate (Etn-P) moieties to specific positions on the first three mannoses (see below). Addition of a fourth, α₁,2-linked mannose (Man) to the third Man (Man₃) is obligatory in yeast but not in mammalian cells, although a fourth Man may be added in a tissue-specific manner (see below). Various side branching sugars can be added, most likely in the Golgi. The inositol acyl group is normally removed in the ER after transfer of the GPI to protein, and remodeling of the diacylglycerol moiety present on the GPI precursor that is transferred to protein likewise is initiated in the ER and continued in the Golgi (see below). In yeast, the lipid moiety of many GPIs is remodeled to ceramide (23). Protein-bound GPIs in both yeast and mammals can bear Etn-P on Man₁ (20,24), but it is not clear whether this substituent is invariably retained, or whether it can be removed from certain GPIs after anchor transfer to protein. Likewise, the fate of Etn-P moieties added to Man₂ of GPI precursors is unknown.

GPI attachment signals.

Direct biochemical demonstration that a protein bears a GPI anchor is not straightforward. Preliminary evidence for GPI anchorage can be obtained if a membrane-associated protein can be made aqueous-soluble by treatment by Ptd-Ins-specific phospholipase C (PI-PLC) (25, 26). However, although most, if not all protein-bound GPIs in S. cerevisiae are PI-PLC-sensitive, the released proteins tend to remain hydrophobic and partition into the detergent phase in the commonly used Triton X-114 fractionation procedure (27). Moreover, many yeast GPI-proteins become covalently cross-linked to cell wall polysaccharide and are therefore hard to recover for analysis. Proteomics approaches, however, are now permitting increasing numbers of GPI-proteins to be identified directly in fractions of yeast and mammalian cells (28, 29).
Biochemical verification of potential GPI modification is now largely eschewed for *in silico* analyses using predictive algorithms, which reliably identify GPI proteins (30-35). GPI attachment is inferred if a protein’s amino acid sequence contains (i) a hydrophobic NH$_2$-terminal secretion signal that targets the protein to the ER and (ii) a C-terminal GPI signal-anchor sequence. Yeast GPI proteins additionally often contain stretches of amino acids rich in serine and threonine (30, 31). The features of mammalian and *S. cerevisiae* signal anchor sequences were originally defined experimentally in site-directed mutagenesis studies using model proteins (36, 37; reviewed in ref. 38), and are shown schematically in Fig. 1B. The amino acid to which the GPI becomes linked is referred to as the “ω” residue, with amino acids N-terminal to it designated ω-minus, and those C-terminal, ω-plus. The general features of a signal-anchor sequence are i) a stretch of about 10 polar amino acids (ω-10 to ω-1) that form a flexible linker region; ii) the ω amino acid, typically G, A, S, N, D, or C; iii) the ω+2 amino acid, the most restrictive position, with predominantly G, A, and S; iv) a spacer region of moderately polar amino acids (ω+3 to ω+9 or more); v) a stretch of hydrophobic amino acids variable in length but capable of spanning the membrane (32-35, 38). A given C-terminal sequence may contain more than one set of candidates for the ω and ω+2 amino acids. Slight variations have been noted in the relative lengths of the spacer regions and hydrophobic segments of metazoan and protozoan signal anchor sequences (33).

In yeast, *in silico* predictions that a GPI attachment sequence is functional have been verified by fusing the sequence to the C-terminus of a reporter protein and testing whether it directs cell surface expression of the protein (which is assumed to be GPI dependent), either in the plasma membrane via a lipid, or in a glycosidic link to cell wall β1,6-glucan (31) (see below).

**BIOSYNTHESIS OF THE GPI PRECURSOR**

The biosynthetic pathway for the glycan core of the GPI precursor was initially charted via *in vivo* and *in vitro* radiolabeling experiments in trypanosomes (39-42). The steps in the more complex GPI assembly pathway in yeast and mammals were identified via analyses of GPI biosynthesis, transfer, or processing defects in mutant cell lines (43, 44), or by *in vitro* radiolabeling experiments (45). Since GPI biosynthesis is dispensable for the growth of mammalian cells in culture many GPI-deficient mutant cell lines were available early on, facilitating analyses of the pathway. Others were established more recently by using innovative selection procedures such as resistance to aerolysin, a GPI-binding, pore-forming toxin (46). In contrast to cultured mammalian cells, most genes required for assembly of the GPI precursor in yeast are essential, the exceptions being genes encoding subunits of an enzyme complex that includes essential proteins, and genes responsible for two of the known modifications to the GPI core. In these cases, however, viable disruptants are temperature-sensitive or have severe growth defects. Conditional yeast mutants as well as GPI-defective mammalian cell lines were instrumental in defining proteins required for GPI biosynthesis.

The products of more than 20 genes are directly involved in the assembly of the GPI precursor and its transfer to protein. These proteins are localized in the membrane of the ER, and the vast majority are polytopic membrane proteins. None of the proteins
has been structurally characterized at the atomic level, nor is there information on the detailed enzymology of most of the individual steps. Here we summarize what is known about the individual reaction steps of GPI biosynthesis in yeast and mammals. The proteins associated with each step are listed in Table 1 and the biosynthetic pathway is illustrated in Fig. 2.

**The N-acetylglucosaminyl-PI synthetic complex and its regulation.**

GPI assembly is initiated with the transfer of GlcNAc from UDPGlcNAc to Ptd-Ins (step 1, Fig. 2A). The acyl chains of acceptor Ptd-Ins appear to be the same length as those in bulk membrane phospholipids (47, 48), although in vitro experiments with mammalian cell membranes suggest that there may be some preference for certain acceptor PtdIns molecular species over others (49, 50). The GlcNAc transfer reaction is unexpectedly complex in that many proteins are involved. The mammalian proteins and their yeast orthologues are PIG-A/Gpi3p, PIG-C/Gpi2p, PIG-H/Gpi15p, PIG-P/Gpi19p, PIG-Q(hGpi1p)/Gpi1p, PIG-Y/Eri1p (50-65). In mammalian cells, a seventh protein, Dpm2p, a non-catalytic subunit of Dol-P-Man synthase, physically associates with the PIG-A, C, and Q proteins and enhances GlcNAc-PI synthetic activity (58). *S. cerevisiae* has no Dpm2p orthologue. A defect in PIG-A causes paroxysmal nocturnal hemoglobinuria and overexpression of PIG-P has been noted in fetal Down Syndrome brain (2, 66).

Four of the six *S. cerevisiae* proteins involved in GlcNAc-PI synthesis are essential, but strains lacking the non-essential Gpi1p and Eri1p proteins are temperature-sensitive for growth and lack detectable in vitro GlcNAc-PI synthetic activity (60, 63, 64). PIG-A/Gpi3p is the catalytic subunit because Gpi3p binds a photoactivatable UDPGlcNAc analogue (67) and is a member of Glycosyltransferase Family 4 (GT4) of retaining glycosyltransferases (68, 69). Since the bulk of PIG-A is oriented towards the cytoplasm and anchored in the ER membrane by a single transmembrane domain, synthesis of GlcNAc-PI is likely to take place at the cytoplasmic face of the membrane (62, 70). Consistent with this, newly synthesized GlcNAc-PI is accessible to membrane topological probes on the cytoplasmic face of the ER (71). The mammalian proteins involved in GlcNAc-PI synthesis form a multiprotein complex (“GPI-GnT”) (50, 58, 65) that has been modeled (62, 72). Genetic interactions between GPI1, GPI2, and GPI3 are also consistent with the notion that their products function in a complex (52). The roles of the other subunits in the GPI-GnT complex are as yet unclear, but they may mediate regulatory interactions. Yeast GPI-GnT can be joined by a further subunit, Ras2p, which, in its GTP-bound form, physically associates with the complex (64). Membranes from RAS2-deleted *S. cerevisiae* strains had 8-10 times the activity of membranes from wild type cells, whereas membranes from cells expressing the activating RAS2-Val19 mutation made virtually undetectable amounts of GlcNAc-PI (64), indicating that Ras2p-GTP functions as a negative regulator of GlcNAc-PI synthesis. Indeed, depending on the degree of its activation, yeast Ras2p could permit about a 200-fold range in GPI-GnT activity (64). The significance of Ras2-regulation of GPI-GnT is not fully understood. However, because GPI-GnT mutants display the hyperactive Ras phenotypes of filamentous growth and invasion of agar when grown at a semi-permissive temperature, it was suggested that Ras 2-mediated down regulation of GPI-GnT may contribute to the changes in cell wall architecture that
accompany the dimorphic transition to filamentous growth (59, 63, 64). Ras GTPases have so far not been found to regulate the activity of mammalian GPI-GnT (65).

**GlcNAc-PI de-\(N\)-acetylation.**

Following its transfer to PI, GlcNAc-PI is de-\(N\)-acetylated to GlcN-PI by Pig-L/ Gpi12p (step 2, Fig. 2A). These orthologues are catalytic because bacterially expressed PIG-L has GlcNAc-PI de-\(N\)-acylase activity (73). The GlcNAc-PI de-\(N\)-acetylation reaction occurs on the cytoplasmic side of the ER membrane, because PIG-L is a type I membrane protein with a large cytoplasmic domain and very few amino acids disposed towards the ER lumen (73-75). Consistent with this, newly synthesized GlcN-PI can be hydrolyzed by PI-PLC on the cytoplasmic face of intact ER vesicles (71).

**Flipping.**

As depicted in Fig. 2A, the GPI biosynthetic pathway is topologically split. The first two reactions occur on the cytoplasmic face of the ER while inositol acylation and mannosylation reactions likely occur in the ER lumen as described below. For GlcN-PI to be inositol acylated and mannosylated, it must be flipped across the ER membrane (step 3, Fig. 2A). Flipping of glycerophospholipids is energetically expensive and does not occur spontaneously: \(\sim 20-50\) kcal/mol must be expended to transfer the polar headgroup of an amphipathic phospholipid such as PtdCho through the hydrophobic interior of the bilayer (15). Since the weight of the available evidence suggests that a lipid translocator - or flippase - is needed to flip PtdCho and other glycerophospholipids across the ER membrane in the process of membrane biogenesis, it is likely that a flippase would also be needed to flip GlcN-PI from the cytoplasmic leaflet of the ER to the luminal side during GPI biosynthesis. Although a GPI flippase has yet to be identified, a recent report describes the successful biochemical reconstitution of GPI flipping in proteoliposomes generated from a detergent extract of rat liver ER vesicles (76). Flipping was demonstrated to be bi-directional and independent of ATP; both GlcNAc-PI and GlcN-PI were flipped. Curiously, both the GPI flippase and the ER phospholipid flippase appeared to be similarly abundant in the detergent extract used for reconstitution, raising the possibility that they could be one and the same protein (76). This result would explain why, despite extensive efforts, flipping is the only step of GPI assembly for which a mammalian cell mutant has not been identified; this could be because the GPI flippase also operates in a process - such as flipping of bulk phospholipids - that is essential for mammalian cell viability in culture.

**Inositol acylation.**

*S. cerevisiae* and mammalian mutant cell lines defective in the formation of dolichol phosphate (Dol-P)-Man, the donor of the mannoses in the GPI precursor, accumulate GlcN-PI bearing an acyl group on the inositol's hydroxyl, which makes the lipid resistant to PI-PLC (77-79). All mannose-containing GPI precursors characterized in *S. cerevisiae* and mammalian cells are inositol-acylated. Inositol acylation therefore most likely occurs at the level of GlcN-PI *in vivo* (step 4, Fig. 2A). Both yeast and mammalian membranes have an acyl CoA-dependent GlcN-PI: inositol acyltransferase
activity (78, 80), although CoA-independent acyltransferase activity has also been reported (81, 82).

The orthologous mammalian PIG-W and yeast Gwt1 proteins are responsible for inositol acylation. GWT1 was identified because its high level confers resistance to the anti-fungal compound 1-[4-butylbenzyl] isoquinoline, which blocks surface expression of GPI proteins (83). Immunoprecipitated, epitope-tagged PIG-W catalyzes palmitoyl CoA-dependent inositol acyl transfer (84), and membranes from pig-w and gwt1 cells lack in vitro palmitoyl CoA-dependent GlcN-PI: inositol acyltransferase activity (84, 85). These results indicate that the major, and perhaps only in vivo inositol acylating activity uses acyl CoA as donor. The acyl group transferred in vivo in S. cerevisiae appears to be palmitate, although acyl chains of a range of different lengths can be transferred from the corresponding CoA derivatives by S. cerevisiae and Cryptococcus neoformans membranes in vitro (78, 86).

The catalytic site of PIG-W is predicted to be located on the lumenal side of the ER (84). Although it has been reported that the ER presents a barrier to the transport of fatty acyl CoAs (87) that are required for the PIG-W-catalyzed acyl transfer reaction, it is known that a number of secretory proteins such as Hedgehog and Wnt are acylated in the ER lumen by members of the acyl-CoA-dependent membrane-bound O-acyl transferase (MBOAT) family of proteins (12, 88). Similarly (see later), fatty acid remodeling of the GPI anchor in the ER lumen requires the acyl-CoA-utilizing MBOAT family members Per1p and Gup1p. These points suggest that acyl-CoAs must somehow be available in the lumen of organelles of the secretory pathway and so it seems likely that the inositol acylation reaction occurs at the luminal face of the ER (step 4, Fig. 2A).

What is the role of the inositol acyl group? Deletion of yeast GWT1 is either lethal or leads to very slow, temperature-sensitive growth depending on the strain (83), indicating the importance of inositol acylation in yeast. PIG-W is critical for expression of GPI proteins at the surface of mammalian cells. In vitro, the presence of the inositol acyl group is obligatory for transfer of α1,4-linked Man-1 to inositol (80). However, radiolabeling studies with pig-w mutants reveal that they can still synthesize mannosylated GPIs lacking the inositol acyl group (84), and it was concluded that inositol acylation is not obligatory for translocation of GlcN-PI into the ER lumen, or for subsequent mannosylation steps. Because PIG-W mutants accumulate GlcN-PI, and this accumulation is eliminated in cells transfected with PIG-W cDNA, inositol acylation clearly enhances the efficiency of the first mannosylation step. In pig-w mutants, the absence of the inositol acyl group did not affect modification of Man-1 with Etn-P, but GPIs bearing Etn-P on Man-3 were not detected, suggesting that the inositol acyl chain is also important for addition of the Etn-P through which the GPI becomes linked to protein (84).

Mannosylation.

Dolichol-P-mannose is the donor of the core mannosyl residues (Man-1, Man-2 and Man-3) of the GPI precursor (89), and is presumed also to contribute the fourth mannosyl residue that is essential in the yeast GPI biosynthetic pathway (Fig. 2B). Mammalian Dol-P-Man synthesis-defective mutants (termed Class E), and the temperature-sensitive yeast dpm1 mutants are blocked in GPI anchoring (77, 90, 91),
as is a mammalian cell mutant with a putative defect in Dol-P-Man usage (92). GPI mannosylation can be blocked in mammalian cells by treatment with mannosamine, a compound that may act as a chain terminator (93). Four Dol-P-Man-utilizing, inverting GPI-mannosyltransferases (GPI-Man-Ts) have been identified. These multi-spanning membrane proteins are highly conserved functionally, for the mammalian genes encoding GPI-Man-T-II, III, and IV all rescue lethal null mutations in their S. cerevisiae orthologues. By analogy with the Dol-P-Man-dependent steps in the dolichol-linked pathway for protein N-glycosylation (94), mannosyltransfer to GPIs is expected to occur on the luminal face of the ER. This means that Dol-P-Man, a lipid that is synthesized from Dol-P and GDP-Man on the cytoplasmic face of the ER (Fig. 2A), must flip across the ER membrane. Although a transporter/flippase for Dol-P-Man remains to be found, ATP-independent dolichol-P-mannose transport activity was recently reconstituted, raising the possibility that this transporter could be identified biochemically in the near future (95).

**GPI-Man-T-I.** PIG-M/Gpi14p is responsible for addition of Man-1, linked α1,4 to GlcN (step 5, Fig. 2A) (96). These proteins have been assigned to CAZy glycosyltransferase family 50 (68), and are members of a large family of proteins with 12 predicted transmembrane domains (97), in a subgroup with Dol-P-Man-utilizing α1,3 Man-T Alg3p, which acts in the dolichol pathway for N-glycosylation. A lumenerally oriented loop between transmembrane domains I and II contains a “DXD motif” common to many glycosyltransferases (98), and necessary for PIG-M function. The lumeneral location of the functionally important DXD motif is one of the pieces of evidence in support of the lumeneral orientation of GPI mannosyltransfer reactions. However, whether the DXD motif serves a ligand binding function in PIG-M/Gpi14p remains to be determined since it was previously demonstrated to play a role in coordinating nucleotide sugars rather than lipid-linked sugars such as Dol-P-Man.

A second protein, PIG-X, is also required for GPI-Man-T-I function in mammalian cells. This membrane protein, which physically associates with and stabilizes PIG-M, does not appear to be involved in Dol-P-Man synthesis or its utilization for N-glycan synthesis (99). The essential 416 amino acid Pbn1 protein, is the functional homologue of PIG-X in yeast. Demonstration that PBN1 complements the GPI anchoring defect in pig-x cells required that GPI14 be cotransfected with PBN1, suggesting that the functional interactions between PIG-X/PIG-M and Pbn1p/Gpi14p are specific ones (99). Analyses of Pbn1p-deficient yeast strains, however, indicate that this protein is involved in processing diverse proteins that pass though the ER, including a protease destined for the vacuole, a vacuolar membrane protein, as well as the GPI-protein Gas1p, attachment of whose anchor was not apparently affected (100). Pbn1p therefore has additional function in yeast, and it’s role in GPI biosynthesis may be a minor one.

**GPI-Man-T-II.** PIG-V/Gpi18p is required for addition of α1,6-linked Man-2 (step 7, Fig. 2A) (101, 102). Mammalian pig-v mutants accumulate a GPI intermediate with a single mannose, as well as a GPI with a single mannose additionally modified with Etn-P (lipid H5, Fig. 2A). Likewise, yeast gpi18 mutants accumulate a Man1-GPI bearing Etn-P on its mannose, but an unmodified Man1-GPI was not resolved unambiguously (101). Thus PIG-V/Gpi18p, could, in principle, use Man1-GPIs, with or without Etn-P on their mannose, as acceptor, and this possibility is illustrated in Fig. 2B.
The PIG-V/Gpi18p proteins have been assigned to single-member CAZy glycosyltransferase family GT-76. In contrast to GPI-Man-T I, II, and IV, PIG-V/Gpi18p are predicted to have 8 transmembrane domains and no obvious sequence similarity to Family 22 or 50 glycosyltransferases or obvious homologues in the mammalian or yeast proteomes. However, sequences of protein: O-mannosyltransferases and the Stt3p subunit of oligosaccharyltransferases are among the many proteins identified after multiple reiterations of Psi-BLAST using the PIG-V sequence as a probe (102), raising a possibility that these glycan transferring proteins are distantly related.

GPI-Man-T-III. The PIG-B/Gpi10p mannosyltransferase adds the third α1,2-linked Man to GPs (step 8, Fig. 2A) (103-105). Mammalian pig-b and yeast gpi10 mutants accumulate a Man2-GPI with an Etn-P substituent on Man-1 (104-108). The preferred acceptor for Gpi10p appears to be a Man2-GPI with Etn-P on Man-1, but, as discussed below, the requirement for Etn-P on Man-1 is not absolute. PIG-B/Gpi10p is a member of CAZy glycosyltransferase Family 22, along with Smp3p, as well as Alg9p and Alg12p, respectively α1,2- and α1,6-Man-T that function in the assembly of the dolichol-linked precursor in N-glycosylation). No Family 22 protein has a “DXD” motif. Protease protection experiments on PIG-B suggested the bulk of the protein is lumenally oriented and membrane-anchored by a single transmembrane domain near its amino terminus (103). However, in silico analyses predict 12 transmembrane domains (97).

GPI-Man-T-IV. The human and yeast Smp3 proteins, also members of Glycosyltransferase Family 22, add a fourth, α1,2-linked Man to Man-3 of the GPI (Fig. 2B) (109, 110). Smp3p function is essential in yeast (109) because addition of this side branching fourth Man to the 2-OH of Man-3 obligatorily precedes addition of Etn-P to the 6-OH of Man-3 (109). Mammalian cells do not have this stringent requirement for a fourth mannose, for unlike yeast, they transfer Man3-GPIs to protein (18) and make trimannosyl GPI precursors bearing Etn-P on Man-3 (108, 111-114). Further, the major GPI precursors that accumulate in GPI transamidase-defective mammalian cell lines are trimannosyl structures (112, 115). However, a fourth mannose can be present on the GPs of some mammalian proteins, notably in brain (18, 116, 117), and traces of putative tetramannosyl GPs have been detected in radiolabeling experiments with cultured mammalian cells (108, 113, 118). An mRNA for an Smp3p orthologue is expressed in human tissues, highest levels being in the brain and colon, and expression of hSMP3 cDNA in human cells results in formation of Man4-GPIs (110). hSmp3p may therefore be expressed in a tissue-specific manner.

Phosphoethanolamine addition

Proteins involved in Etn-P addition. In addition to the Etn-P moiety at the 6-position of Man-3, through which the GPI becomes linked to protein, Etn-P moieties can be added to the 2-OH of Man-1 and the 6-OH of Man-2 as well (Fig. 1). These side branches are added during precursor biosynthesis. Protein-bound GPs of both mammalian and yeast proteins can bear side-branching Etn-Ps on their first mannosyl residue (18, 20, 24). There is no evidence, however, that side-branching Etn-Ps are involved in linkages to protein (105). The Etn-P moieties on Man-1 and Man-3 are transferred from Ptd-Etn, not CDP-Etn (119, 120). Since all Etn-P additions are likely to occur in the ER lumen, this requires flipping of PtdEtn from the cytoplasmic face of the
ER to the luminal face (Fig. 2A). Three related proteins, PIG-N/Mcd4p, hGpi7p/Gpi7p, and PIG-O/Gpi13p (GPI-Etn-P-T-I, II, and III), appear to be responsible for addition of an Etn-P moiety to each of the three mannoses. These large proteins are predicted to have 10-14 transmembrane domains and a luminal hydrophilic loop between transmembrane segments 1 and 2 containing sequences characteristic of members of the alkaline phosphatase superfamily (113, 114, 121-123). The alkaline phosphatase domain is critical for protein function because the G227E mutation encoded by the mcd4-174 allele that leads to temperature-sensitivity and a conditional block in GPI anchoring (121) is in one of the two metal binding sites in alkaline phosphatase family proteins (123). The presence of the alkaline phosphatase domain is consistent with a role for these proteins in the addition or removal of a phosphodiester-like substituent, and indeed, deficiencies in these proteins prevent addition of Etn-P moieties to mannoses 1, 2, or 3 of GPI precursors. The small, hydrophobic PIG-F/Gpi11 protein functions as an auxiliary subunit of at least two of the GPI-Etn-P-T, but its relative importance varies between yeast and mammalian cells. The phenotypes of yeast mutants defective in Etn-P addition are complex, suggesting that the substituents they add are important for diverse processes involving the GPI; there is also evidence that the GPI-Etn-P-T themselves may have additional functions in yeast.

**GPI-Etn-P-I.** The PIG-N/Mcd4 proteins are involved in addition of Etn-P to Man-1 (step 6, Fig. 2A, B), but this modification is of different relative importance in mammalian and yeast cells. Mammalian cells lacking PIG-N synthesize GPIs lacking Etn-P on Man-1, and are significantly affected in surface expression of GPI proteins (113), consistent with their inability to synthesize a full complement of mature GPIs. In contrast, a conditionally lethal yeast mcd4 allele blocks GPI anchoring (121) and leads to accumulation of a Man$_2$-GPI lacking Etn-P (124). *In vitro* GPI synthesis by mcd4 membranes also does not proceed beyond Man$_2$-GPI (24). The finding that Mcd4p defects lead to accumulation of Man$_2$-GPI lacking Etn-P on Man-1 suggests that Mcd4p normally transfers Etn-P to a dimannosyl GPI, but the accumulation of a Man$_1$-GPI with Etn-P on its lone mannose in both yeast and mammalian GPI-Man-T-II mutants (see above) indicates that singly mannosylated GPIs can serve as acceptors for PIG-N/Mcd4p as well. The accumulation of a Man$_2$-GPI lacking Etn-P on Man-1 in Mcd4p-deficient cells further suggests that an Etn-P on Man-1 is required for GPI-Man-T-III-dependent addition of the third mannose. However, mcd4 disruptants can be partially bypassed by high-level expression of GPI10, and a Man$_3$-GPI probably bearing one EtnP on Man-2 is still formed in an mcd4/gpi11 double mutant (124). Gpi10p’s requirement for Etn-P on Man-1 is therefore not an absolute one.

A screen for inhibitors of GPI-dependent processing of the model yeast protein Gas1p identified the terpenoid lactone YW3548. Treatment of yeast cells with this compound led to accumulation of an unsubstituted Man$_2$-GPI, consistent with the idea that YW3548 blocks Mcd4p-dependent addition of Etn-P to Man-1 (104, 125). YW3548 was shown to elicit a similar block in GPI biosynthesis in mammalian cells (109). Yeast mcd4 disruptants partially bypassed by native or heterologous GPI-Man-T-III genes are slow-growing (124, 126), indicating that in addition to possibly enhancing the efficiency of Man-2 addition by Gpi10p, this side branch may be required for later functions of the GPI. Indeed, mcd4Δ cells partially rescued by *Trypanosoma brucei* GPI10 transfer GPIs to protein less efficiently, exhibit retarded export of GPI proteins
from the ER, no longer remodel the lipid moiety of GPI anchors to ceramide, and are
defective in axial bud site selection, suggesting that recognition of Etn-P on Man-1 is
important in diverse contexts (126). An Etn-P moiety on Man-1 is important for
recognition of GPIs by the human GPI transamidase (118) (see below).

Mcd4p may have additional functions in yeast. The mcd4-P301L allele, but not
the G227E mutant, exhibits a defect in Ptd-Ser transport to the Golgi and vacuole for
subsequent decarboxylation, yet is not obviously defective in GPI anchoring, suggesting
Mcd4p may have a role in transport-dependent metabolism of Ptd-Ser (127). Further,
high-level expression of MCD4 results in extracellular ATP release and enrichment of
both Mcd4p and ATP uptake activity in the Golgi, leading to the suggestion that Mcd4p
may normally mediate symport of ATP and Ptd-Etn into the lumen of the ER (128).

GPI-Etn-P-T-II. The human and yeast Gpi7 proteins are involved in Etn-P
addition to Man-2 (step 10, Fig. 2A). Viable, though temperature-sensitive yeast gpi7
disruptants accumulate a Man₄-GPI with Etn-P on Man-1 and Man-3, but unmodified on
Man-2 (122), and analogously, hGpi7p-depleted human cells accumulate a Man₃-GPI
bearing Etn-P on Man-1 and Man-3, but not on Man-2 (129).

The hydrophobic PIG-F protein binds to and stabilizes hGpi7p (129) (as well as
PIG-O (see below)). Although an interaction between Gpi11p and Gpi7p in yeast has
not been demonstrated, a role for PIG-F/Gpi11 in Gpi7p function was proposed
because the profile of Man₄-GPI precursors that accumulated in gpi7 disruptants was
similar to that accumulated in gpi11 mutants (130).

In yeast, the gpi7 deletion leads to a range of defects. These include impaired
ER to Golgi transport of GPI proteins and ceramide remodeling (122), and inefficient
cross-linking of GPI proteins into the cell wall (131). In addition, gpi7 disruptants have a
cell separation defect attributable to mis-targeting of Egt2p, a daughter cell-specific
protein that normally is involved in degradation of the division septum (132). These
phenotypes suggest that the Etn-P moiety on Man-2 of protein-bound GPIs may be a
signal recognized by components of the intracellular transport machinery, by lipid
remodeling enzymes, or by proteins involved in cross-linking GPI to the cell wall.

GPI-Etn-P-III. Human PIG-O and yeast Gpi13p are required for addition of Etn-
P to α1,2-linked Man-3 (step 9, Fig. 2A) (114, 130, 133). This Etn-P is the one that
becomes amide-linked to the COOH terminus of a GPI protein, and as would be
expected, GPI13 is essential in yeast. The Zn²⁺ chelator 1,10-phenanthroline inhibits
GPI anchor synthesis in mammalian cells by blocking the PIG-O catalyzed Etn-P
transfer reaction (134); metabolic labeling experiments with phenanthroline-treated cells
show an accumulation of H6 (Fig. 2A), a triply mannosylated GPI intermediate with Etn-
P on Man-1. As is the case for hGpi7p, PIG-O is also bound and stabilized by the PIG-F
protein (114), and indeed, the GPI accumulation phenotype of pig-f mutants indicates
they are blocked in Etn-P addition to Man-3 (107, 108, 135, 136). In yeast, however, the
PIG-F orthologue Gpi11p does not have a major role in Etn-P addition to Man-3,
because the GPIs accumulated in strains lacking Gpi11 or PIG-F function include a
Man₄ species bearing Etn-P on its third mannose (130). Yeast and human cells may
therefore differ in the extent to which their GPI-Etn-P-II and III proteins depend on a
partnership with PIG-F/Gpi11p. Yeast Gpi11p’s role, however, is not solely to partner
Gpi7p, for unlike GPI7, GPI11 is an essential gene. Indeed, both Gpi11p and PIG-F
may also act earlier in GPI assembly, because in addition to the two Man₄-GPIs that
have been characterized, gpi11 mutants accumulate two more non-polar mannolipids (130), and pig-f mutants also accumulate multiple mannolipids (108, 135-137). PIG-F, however, does not appear to have a role in PIG-N function, because these two proteins do not interact physically (114).

What is the normal GPI assembly pathway?

Structural characterization of the headgroups of GPIs that accumulate in mammalian and yeast GPI mutants has identified a series of potential intermediates that can be arranged in a linear pathway for assembly of the complete GPI precursor, a Man$_3$- or possibly a Man$_4$-GPI with EtN-Ps on Man-1, 2, and 3 in mammals, and a Man$_4$-GPI with Etn-Ps on Man-1, 2, and 3 in yeast (Fig. 2A).

In addition to these complete precursors, it is possible that counterparts of these GPIs lacking Etn-P on Man-2, but probably not Man-1, are also transferred to protein. However, transfer of these structures may be sub-optimal because the double mutant between gpi7 and transamidase mutant gpi8 has a synthetic growth defect (122). This synthetic phenotype, though, could also reflect a need for some threshold level of protein-bound GPIs with Etn-P on Man-1.

A feature of the linear pathway is the early addition of Etn-P to Man-1. However, although the "[(Etn-P)Man$_1$-GPI]" pathway is a direct route to a transfer-competent precursor, both yeast and mammalian mutants also synthesize mannosylated GPIs that cannot be placed in it. Some of these GPIs have no EtN-P’s whereas others lack Etn-P on Man-1, but receive this moiety on Man-2 instead (109, 110, 113, 129, 130, 133; Fig. 2B). These GPIs can arise in both yeast and mammals because neither PIG-N nor Mcd4p are obligatory for addition of Man-3 (113, 124), although GPI-Man-T-III likely has a strong preference for acceptors bearing Etn-P on Man-1.

It is not known, however, whether the GPI isoforms lacking EtN-P on Man-1 are physiologically relevant and represent alternative routes to GPI transamidase substrates. These glycolipids may remain as free GPI and have additional functions of their own, but their accumulation could also be toxic to cells.

TRANSFER OF GPIs TO PROTEIN

As discussed earlier, genes encoding GPI-anchored proteins specify two signal sequences in the primary translation product: an N-terminal signal sequence for ER-targeting and a C-terminal sequence that directs the attachment of a GPI anchor (Fig. 1B). Both sequences are removed during processing of the pre-pro-protein to its mature GPI-anchored form. The removal of the C-terminal GPI signal sequence and its replacement with GPI on the lumenal face of the ER (step 11 in Fig. 2A) are catalyzed by GPI transamidase (GPIT) (Fig. 3A), a multisubunit membrane-bound enzyme.

Protein translation and translocation must be complete in order for the C-terminal GPI signal sequence of the pro-protein to emerge on the lumenal side of the ER where it is recognized by GPIT. GPIT activates the carbonyl group of the $\omega$ amino acid in the pro-protein by displacing the GPI signal sequence (residues $\omega+1$ through to the C-terminus; the metabolic fate of the cleaved signal sequence is unknown but it may be processed in a manner analogous to that described for cleaved ER-targeting N-terminal signals (138)). Nucleophilic attack on the activated carbonyl by the amino group of an
Etn-P-capped GPI substrate (such as mammalian H7 or H8, or yeast CP2, a Man₄-GPI with Etn-P on Man-1,2 and 3; Fig. 2A) yields a GPI-anchored protein and regenerates GPIT (139-141) (Fig. 3B). Small nucleophiles like hydrazine and hydroxylamine can replace GPI in microsome-based assays of GPI anchoring (140, 141), providing a simple test of carbonyl activation. Nucleophilic attack by water is also seen in these assays (142) raising the possibility that a small percentage of pro-protein products is simply secreted without receiving a GPI anchor. The proportion of secreted proteins may increase when GPI supply is short (93).

GPIT is a complex of five membrane proteins; in mammals and yeast these are PIG-K (or GPI8)/Gpi8p, GAA1/Gaa1p, PIG-S/Gpi17p, PIG-T/Gpi16p and PIG-U/Gab1p (Fig. 3A). Gaa1p and Gpi8p were the first to be identified through genetic approaches (143-146), and were subsequently shown to form a complex (147); the other subunits were identified mainly because they specifically co-immunoprecipitated with the GPI8/Gpi8p-GAA1/Gaa1p complex (148-151). GPITs from D. melanogaster, C. elegans and A. thaliana are similar to the mammalian/yeast enzyme. In contrast, GPITs from trypanosomatids such as Trypanosoma brucei, share 3 subunits with mammalian/yeast GPIT (homologs of GPI8, GAA1 & PIG-T termed TbGPI8, TbGAA1 and TbGPI16/PIG-T, respectively) but have 2 novel subunits (TTA1 & TTA2) in lieu of PIG-S and PIG-U (142, 153). All 5 subunits are essential in both human and trypanosomatid-type GPIT, and all are needed for the nucleophilic attack on the ω residue that produces the activated carbonyl intermediate to initiate transamidation (147, 148, 150).

GPI8/Gpi8p is presumed to be the catalytic center of the enzyme since (i) it shares sequence homology with a family of cysteine proteases, one member of which has transamidase activity in vitro (154), (ii) mutagenesis of a putative Cys-His catalytic dyad in PIG-K/GPI8 inactivates GPIT (147, 154), and (iii) it is in physical proximity to the pro-protein substrate (155, 156). Mammalian/yeast GPI8/Gpi8p is a type I membrane protein with the majority of its sequence located in the ER lumen; GPI8 in many lower eukaryotes such as nematodes and trypanosomes is a soluble protein of the ER lumen suggesting that the TM domain is functionally dispensable (152, 157). Consistent with this, a human GPI8 construct lacking the TM domain can rescue GPI-anchoring in GPI8-K562 cells (147). GPI8 is disulfide-linked to PIG-T (158); although this association is not essential, it promotes GPIT activity in vivo. E. coli-expressed T. brucei GPI8 cleaves a tetrapeptide substrate (159) indicating that TbGPI8 - in the absence of other GPIT subunits - is catalytically active, at least as a protease. Genetic and biochemical analyses of the interactions among the subunits of the yeast GPT transamidase complex suggest that the enzyme may consist of two subcomplexes, one containing Gpi8p, Gpi16p, and Gaa1p, and the other, Gab1p and Gpi17p (149, 151, 160; Fig. 3A).

It is unclear why GPIT consists of five essential subunits when, in principle, the GPI8/Gpi8p subunit should suffice. With the possible exception of the PIG-U/Gab1p subunit that shares a minor resemblance to fatty acid elongase proteins and may therefore have a lipid binding role within GPIT (150), the function of the other ‘non-catalytic’ subunits of GPIT cannot be predicted from their primary sequence. These subunits resemble only their counterparts in other eukaryotes and share no sequence similarity with proteins of known function. A similar conundrum exists for the ER-localized oligosaccharyltransferase (OST) complex that has 8 subunits, many of which are essential (161). OST, like GPIT, engages protein (ER-translocated nascent chains
with an N-X-S/T- sequon) and lipid (dolichol-PP-oligosaccharide) to generate an N-glycosylated protein. Since a single OST subunit, STT3, is able to carry out oligosaccharyl transfer alone in a bacterial system (162), the question of the variety and essentiality of the other OST subunits in eukaryotes remains. For both GPIT and OST, it seems likely that the ‘non-catalytic’ subunits select and recruit the two very different substrates that are handled by the enzymes, or regulate substrate access to the catalytic site. Recent data support this idea.

Site-specific photo-crosslinking experiments indicate that GAA1 is physically close to pro-proteins that are bound to GPIT (155), and co-immunoprecipitation analyses show that it is required for GPI8 to recognize its protein substrate (163). Capitalizing on the observation that interactions between GPIT and its pro-protein and GPI substrates can be preserved under suitable detergent extraction conditions (164, 165), recent work established that human GPIT complexes containing a defective GAA1 subunit (either lacking the last transmembrane span, or with a proline residue in this span mutated to another amino acid) could bind pro-proteins but not GPI (164, 166). These results indicate that GAA1 plays a key role in substrate recognition by GPIT.

The co-immunoprecipitation methodology used to identify a role for GAA1 in GPI binding was also used to identify the molecular features of GPI that are recognized by human GPIT. Surprisingly, the Etn-P cap linked to Man-3 was shown not to be a prerequisite for GPI recognition, but the Etn-P residue linked to Man-1 was critical (165). Thus, in an in vitro assay, human GPIT was able to pull-down the minimal GPI (Etn-P)2Manα1-4GlcN-acylPI (165) and its more elaborate variants. However, Etn-P-capped GPls lacking the Etn-P modification on Man-1 were poorly recognized. This is a point of some interest for future work since trypanosomatids do not modify GPIs with Etn-P on Man-1 and, as described above, trypanosomatid GPIT complexes are compositionally distinct from the mammalian/yeast family of GPITs.

Components of the GPI transamidase complex have additional, novel functions. Depletion of PIG-U/Gab1p or Gpi8p in yeast, but not of Gaa1p, PIG-T/Gpi16p, or PIG-S/Gpi17p, results in the accumulation of bar-like structures of actin that are closely associated with the perinuclear ER and decorated with the actin binding protein cofilin (151). Since formation of actin bars is not a characteristic of yeast GPI assembly mutants, and therefore not a general consequence of a defect in GPI anchoring, it was suggested that this phenotype reflects perturbation of a functional interaction between resident proteins of the ER membrane and the actin cytoskeleton, leading to collapse of the ER network around the nucleus (151). It can be speculated that Gab1p and Gpi8p are involved in this interaction, although it is not clear whether the interaction is a direct one. Curiously, recent results indicate that the gene for PIG-U, the human Gab1 orthologue, is amplified in human bladder cancer, and overexpression of PIG-T and GAA1 has been linked to breast cancer (166, 168). In the case of PIG-U, it is not yet known whether this gene amplification affects the actin cytoskeleton in human cells.

**LIPID MODIFICATIONS OF PROTEIN-BOUND GPls**

AND CELL WALL CROSS-LINKING

Both the glycan and lipid portions of protein-bound GPls are modified as the GPI-protein transits the secretory pathway. In yeast, an α1,2- or α1,3-linked mannose can
be added to Man-4 in the Golgi (169) by as yet unknown transferases. In the case of mammals, however, neither the subcellular sites at which Man-1 of GPIs receive additional sugars (Fig. 1A), nor the transferases involved, have been identified. It is plausible, though, that Golgi-localized transferases are involved. Our focus is on the modifications to the lipid portion of protein-bound GPIs that are initiated even before the GPI-protein exits the ER, and continue as the protein travels to the cell surface. These modifications are removal of the inositol acyl chain and introduction of longer-chain lipids, and are important for transport of GPI proteins and for modulating the association of GPI proteins with membranes. A further processing event in yeast involves the glycan portion of the GPI, which participates in a transglycosylation reaction in which the entire GPI protein becomes cross-linked into cell wall β1,6 glucan.

**Inositol deacylation**

In what so far appears to be the first post-transfer GPI processing event, the inositol group introduced prior to mannosylation of the GPI precursor is removed by an ER-localized deacylase (170). The orthologous PGAP1/Bst1 proteins, which contain a lipase consensus motif, are responsible (171, 172).

The phenotypes of yeast and mammalian mutants suggest roles for inositol deacylation in incorporation of GPI-proteins into transport vesicles and in quality control of GPI proteins. In mammalian cells, loss of PGAP1 function delays transport of GPI proteins from the ER to the Golgi (171). *S. cerevisiae bst1* mutants were isolated as bypass suppressors of the lethality of a mutation in *SEC13*, which encodes a component of the coat protein complex of COPII-coated vesicles (173). *bst1* mutants are defective in a quality control mechanism in which misfolded proteins are transported to the Golgi by COPII vesicles, then retrieved to the ER (174). Maturation of the GPI-protein Gas1p, which normally exits the ER in a class of vesicles distinct from those used by non-GPI-proteins (175, 176; reviewed in ref. 177), is slowed too. These findings suggest a role for Bst1p in sorting cargo proteins (173, 174); interference in this process presumably allows proteins to exit the ER some other way and bypass the *sec13* defect. It has been speculated that the lethality of *sec13* may be due to accumulation of inositol-deacylated GPI-proteins in association with Emp24p and Erp1p cargo receptors that normally facilitate incorporation of GPI-proteins into transport vesicles (178, 179); this interaction is prevented in *bst1* (171). Loss of Bst1p function not only impacts forward transport of normal GPI proteins, but also their degradation. *BST1* disruptants are viable, but are delayed in ER-associated degradation of a mutant form of Gas1p that is normally misfolded and rapidly degraded in the ER, indicating a role for inositol deacylation in quality control of GPI proteins (172).

**Lipid remodeling.**

In yeast, remodeling of a GPI’s lipid moiety is initiated in the ER, whereas the homologous remodeling activities in mammalian cells seem to start in the Golgi. Analyses of the lipid moieties of the anchors of Gas1p, as well as bulk GPI-proteins, indicate the existence in yeast of pathways for remodeling the shorter acyl chains of the diacylglycerol of freshly transferred GPIs to either base-labile C_{26:0}/C_{26:0} diacylglycerols or to a base-stable ceramide consisting of C_{18:0} phytosphingosine and a hydroxy-C_{26:0} fatty acid (23). The remodeling events start in the ER with removal of the acyl chain at
the sn-2 position of the diacylglycerol to form a lyso-GPI. Per1, an ER membrane protein is required for this activity in yeast (180). Next, a C_{26:0} acyl chain is introduced at sn-2 by the O-acyltransferase Gup1p, an MBOAT family member and resident of the ER membrane (181). The Per1p- and Gup1p-dependent remodeling events are required for optimal ER to Golgi transport of GPI proteins and for their association with lipid rafts (180, 181). The GPI remodeled by Gup1p at sn-2 is a plausible intermediate in the formation of C_{26:0}/C_{26:0} diacylglycerol-containing GPs as well as ceramide-based anchors because the gup1 deletion severely impacts, if not altogether blocks formation of base-stable, ceramide-containing GPs (181). Formation of the latter by replacement of the diacylglycerol with ceramide is also an ER event (48, 182), but the phytosphingosine C_{26:0} ceramide that is initially attached in the ER can in turn be replaced in the Golgi with a phytosphingosine hydroxy-C_{26:0} ceramide (182) to generate the major ceramide-based GPI species found in yeast. The proteins involved in ceramide remodeling are as yet unknown. The Etn-P side-branches introduced by Mcd4p and Gpi7p appear to be important for optimal remodeling (122, 126). It is not yet clear, however, whether the diacylglycerol on any given GPI-protein can be remodeled to ceramide; if only specific proteins are targeted for this modification then it remains to be determined what features of the GPI or protein dictate remodeling.

A mammalian counterpart of Gup1p has not yet been identified, but Per1p’s orthologue is the PERLD1/PGAP3 protein (180; 183). PGAP3-dependent removal of unsaturated fatty acyl chains at the sn-2 position, however, occurs predominantly in the Golgi. The PGAP2 protein, likewise Golgi-localized, is involved in the subsequent introduction of a saturated (C_{18:0}) fatty acid at sn-2 (47). Mutations in the yeast gene that encodes a homologue of PGAP2 – albeit a much larger protein – cause cell wall abnormalities consistent with defects in cell surface anchorage of GPI proteins (47, 184). In mammals, remodeling at sn-2 requires prior inositol deacylation by PGAP1 (183). The PGAP3- and PGAP2-dependent remodeling activities, in turn, are necessary for the GPI anchored proteins to associate with lipid rafts (183).

**Cell wall crosslinking.**

For many fungal GPI proteins, the outer face of the plasma membrane is not the final destination. Rather, an additional processing step occurs in which some, perhaps the majority, of GPI-proteins becomes cross-linked to the cell wall via a remnant of their GPI anchor. This post-secretory pathway processing event seems to be characteristic of fungi and has no counterpart in mammals. Pulse-chase and cell fractionation experiments with the S. cerevisiae Agα1p mating agglutinin led to the proposal that the GPI of a membrane-anchored protein is cleaved within its glycan, with loss of the inositol-containing lipid moiety and transient formation of a soluble intermediate, whose GPI remnant is then transferred to cell wall β1,6-glucan, yielding a GPI-Cell Wall Protein (GPI-CWP) (185, 186). The glucan portion of GPI-CWP-β1,6-glucan is in turn covalently cross-linked to β1,3-glucan and chitin (187-189).

Analyses of the linkage region between GPI-CWPs and β1,6-glucan suggest that the GPI is cleaved between its GlcN moiety and the first mannose, and that the latter’s reducing end is glycosidically linked to the non-reducing end of a β1,6-glucan chain or to
an internal β1,6-linked glucose (189, 190) (Fig. 4). The exact nature of the GPI-Glc linkage is not yet known, nor have the enzymatic activities involved been characterized.

Many yeast GPI-proteins are predominantly cell wall-localized, a few mainly localized in the plasma membrane, and yet others can be detected in both locations (31, 191-195). Although GPI could serve as a generic cell wall targeting signal whose efficiency varies widely between proteins (193), the ability of a protein to be directed to one or the other location can be critical for it to fulfill its function (194, 196). The importance of protein localization, together with the differential distribution of GPI-proteins, suggests the existence of signals in the GPI-protein that promote or prevent the cross-linking reaction at the cell surface.

Studies in which the COOH-terminal sequences of putative GPI-proteins, including amino acids immediately N-terminal to the protein’s predicted GPI attachment site (the “ω-minus” region), are exchanged between GPI-proteins or fused to heterologous proteins, indicate that the ω-minus residues are key determinants of plasma membrane or cell wall localization (191-193, 195). These analyses, together with in silico analyses of the sequences of predicted S. cerevisiae GPI proteins (30), indicate that, in general, if the ω-minus acids include two basic amino acids, the protein will be retained in the plasma membrane and if basic residues are absent, or replaced with hydrophobic amino acids, the protein is shifted to the cell wall (191, 192, 194). However, the dibasic motif alone is not the sole determinant of plasma membrane or cell wall localization (193, 195), for it can be overridden, for example, if a long, serine and threonine-rich stretch of amino acids is also present, which results in cell wall anchorage. Whether any structural feature of the GPI anchor itself also plays a role in determining cross-linkage is unknown.

CONCLUDING REMARKS:
WHITHER GPI PATHWAY RESEARCH?

In this review we describe the almost exponential progress that has been made recently in providing a genetic and biochemical definition of the pathway for GPI anchoring, a highly conserved glycosylation pathway dependent on aminophospholipid supply and critical for modification, transport and surface localization of a subset of proteins. Efforts to chart the GPI pathway have not only identified genes and proteins associated with most of the committed biochemical steps but have also highlighted fascinating issues for future exploration, such as the importance of specific modifications to the protein-bound GPI, the mechanism of GPI flipping, and the enzymology and structure of the membrane proteins involved.

Most of the proteins of the GPI biosynthetic pathway are polytopic membrane proteins, and some function in membrane-bound complexes, making purification a challenge and three-dimensional structural analysis a distant prospect. It is also possible that there is a higher-order organization of some or all of the biosynthetic enzymes, enabling substrate channeling. This possibility has been suggested for the dolichol pathway of protein N-glycosylation (197). The structural complexity of substrate GPIs has restricted enzymological studies. However, advances in synthetic chemistry have allowed the synthesis of selected GPIs, with which GPI-Man-T-I specificity and
GPI flipping have already been investigated, and a chemical-biological analysis of the entire pathway, including the acceptor specificity of pathway enzymes, is likely to follow (76, 80, 198). The availability of defined GPI structures may also aid in the detection of so far elusive enzyme activities such as lipid remodelases, the glycosyltransferases that modify protein-bound GPs, and cell wall cross-linkers.

The pathway for assembly of GPs and subsequent remodeling of their lipid moieties has the potential for generating considerable diversity in anchor forms, a diversity already apparent in vivo from the occurrence of species and tissue-specific variations in GPI structure. Analyses of mutants defective in Etn-P addition and in lipid remodeling are revealing the functional importance of GPI modification at the level of bulk GPI protein in yeast and cultured mammalian cells. What remains to be learned is whether individual proteins receive GPs with specific headgroup variations, whether GPI headgroup structure influences lipid remodeling, and whether the structure of the headgroup and the lipid moiety of the GPI on an individual protein impacts transport, localization, and function in a protein-specific way. Alternatively, are tissue-specific structural variations a consequence of altered expression of specific GPI biosynthetic genes? For example, does upregulation of SMP3 expression cause a wholesale increase in the abundance of Man₄-GPIs on brain glycoprotein? Answers to these questions will require determination of detailed structures of the GPs of individual proteins in different tissues and cell types, as well as during development and in disease. Advances in proteomic analyses focusing on posttranslational modifications may soon make these studies possible (29).

The biological significance of GPI structural diversity in multicellular organisms remains to be explored. For example, what developmental defects might mice exhibit if they are unable to attach a fourth mannose or add Etn-P to Man-2? Might any human genetic diseases be associated with subtle alterations in GPI structure or diminished ability to transfer GPs to protein, in the same way that perturbations in assembly of the dolichol-linked precursor oligosaccharide lead to the range of developmental and neurological defects seen in congenital disorders of glycosylation (199)?

Not only may structural features of GPs be important, but individual proteins in the GPI anchoring pathway may themselves contribute to normal cell growth in novel ways. The findings that the genes for GPI transamidase subunits GAA1 and PIG-U are amplified in human cancers, and that the gene for the PIG-GnT component of GPI-GnT is overexpressed in fetal Down Syndrome brain (66), raise questions of whether and how elevated expression of these non-catalytic subunits impacts GPI anchoring, and, ultimately, gives rise to clinical symptoms. Whether the actin bar phenotype seen when the PIG-U orthologue, Gab1p, and GPI transamidase subunit Gpi8p are depleted in yeast cells (151) will give insights into PIG-U’s novel role in mammalian cells is not yet clear, but the striking actin phenotype in yeast highlights the possibility that individual components of the GPI anchoring machinery have functions beyond their known roles in a glycosylation pathway.

From the outset of work on GPs in parasitic protozoa, the notion has been pursued that there are species-specific steps in GPI assembly that could be exploited as targets for new antiparasitic (7, 8) drugs. The identification of the genes involved in so many of the steps in GPI anchoring in humans and pathogenic eukaryotic microbes
is highlighting new potential drug targets, which are beginning to be validated by gene disruption in pathogenic protozoa and fungi (6, 200).

Further studies of GPIs and the proteins that make them are clearly needed to reveal all the functions of this complex lipid modification in normal cells and in normal development and to determine what goes awry when GPI anchoring is perturbed. There is much to be done and we hope that others, like us, will learn to stop worrying and love glycophospholipids.
Acknowledgements

Work in the authors' laboratories was supported by NIH GM55427 (A.K.M.) and GM46620 (P.O.). We thank S. Kubrick and B. Dylan for the usual reasons, and Sam Canis for his unique insights.
References


complete precursor glycolipids. *EMBO J.* **13:** 2789-9276.


118. Vainauskas, S., and A. K. Menon. 2006. Ethanolamine phosphate linked to the first mannose residue of glycosylphosphatidylinositol (GPI) lipids is a major feature of the GPI structure that is recognized by human GPI transamidase. *J. Biol. Chem.* **281**: 38358-38364.


165. Vainauskas, S., and A. K. Menon. 2006. Ethanolamine phosphate linked to the first mannose residue of glycosylphosphatidylinositol (GPI) lipids is a major feature of the GPI structure that is recognized by human GPI transamidase. *J. Biol. Chem.* **281**: 38358-38364.


agglutinin and β1,6-glucan in the *Saccharomyces cerevisiae* cell wall. *J. Cell Biol.* **128**: 333-340.


Figure legends

Figure 1

(A) Structure of yeast and mammalian GPI-anchored proteins. The core structure of the GPI anchor is shown in chemical detail on the right and in schematic form on the left. The schematic illustration - used in other figures in this review - makes use of symbology recommended by the Consortium of Glycobiology Editors, La Jolla, California, in ‘Essentials in Glycobiology’ (Cold Spring Harbor Laboratory Press).

GPI consists of a trimannosyl core, glucosamine (GlcN) and PtdIns (PI). The mannose residues are labeled Man-1, Man-2 and Man-3, starting with the residue linked to GlcN. The lipid moiety is typically diacylglycerol with saturated fatty acyl chains, although alkylacylglycerol and ceramide structures are possible as well. The 6-OH of the third mannose residue is modified with ethanolamine-phosphate (Etn-P): the amino group of this Etn residue forms an amide bond with the α-carboxyl of the C-terminal amino acid of the protein; this amino acid, termed ω, is an internal amino acid that is exposed during the GPI anchoring reaction.

The schematic illustration shows the core structure of GPI as well as modifications that may or may not be found in the different GPI anchors that have been characterized in yeast and mammals. The 2-OH of inositol in GPI biosynthetic intermediates and the GPI anchor precursor is acylated, usually with palmitic acid. This acyl chain is typically removed before the GPI-protein leaves the ER en route to the cell surface. The mannose residues are decorated with Etn-P side-chains during GPI biosynthesis; carbohydrate modifications of the mannose residues, presumably acquired in the Golgi, are also seen in some GPI-anchored proteins. Man-3 can be modified by an α1,2 linked mannose to yield a Man4-GPI; this mannosylation is an essential biosynthetic step in the assembly of yeast GPIs.

(B) General features of the GPI signal anchor sequences of yeast and mammalian proteins. The C-terminal amino acid sequences of the well studied rat Thy-1 (18) and yeast Gas1 proteins (36) are presented to illustrate the features of the amino acid sequences surrounding sites of GPI attachment. The GPI is transferred to the ω-amino acid with concomitant cleavage of the peptide bond between the ω and ω+1 amino acids. The sequence of the 10 amino acids N-terminal to the ω site contains generally polar residues and forms a flexible linker region. Typical ω amino acids are G, A, S, N, D, or C. The ω+1 position is relatively permissive, whereas ω+2 amino acids are predominantly G, A, and S. A moderately polar spacer region of six or more amino acids starts at ω+3, and is followed by a hydrophobic segment amino acid (32-34, 38). The alanine in the ω+1 position of the Gas1p sequence could serve as an alternative, though less optimal, GPI attachment site.

Figure 2

(A) General scheme for GPI biosynthesis in the ER of yeast and mammals. The ER is depicted as a topologically defined compartment. Biosynthesis begins at the top of the figure with a PtdIns acceptor (grey box). In step 1, PtdIns is glycosylated to generate
GlcNAc-PI on the cytoplasmic face of the ER. GlcNAc-PI is then de-N-acetylated (step 2) to yield GlcN-PI. GlcN-PI is flipped (step 3) into the lumenal leaflet of the ER where it is inositol acylated (step 4), mannosylated and modified by Etn-P (steps 5-10). Man is derived from Dol-P-Man (synthesized from Dol-P and GDP-Man in a reaction catalyzed by DPM1 on the cytoplasmic face of the ER) and Etn-P is derived from PtdEtn (synthesized typically by decarboxylation of PtdSer via the action of PSD (PtdSer decarboxylase); synthesis from CDP-Etn via the Kennedy pathway is also possible). Both Dol-P-Man and PtdEtn must be flipped into the lumenal leaflet of the ER to participate in GPI biosynthesis. The Etn-P-capped GPIs that are synthesized by this pathway (H7 and H8 in mammals; CP2 in yeast (23) (see panel B)) are attached (step 11) to ER translocated proteins displaying a C-terminal GPI signal sequence. Step 11 is catalyzed by GPI transamidase. Man₄-versions of the lipids H6, H7, and H8 may also be synthesized (110).

(B) Additional steps in the yeast GPI assembly pathway. The figure shows that a fourth Man residue must be added before Etn-P modification of Man-3 can occur (109) and illustrates the possibility that GPI-Man-T-II (Gpi18p) and GPI-Etn-P-T-I (Mcd4p) need not act in a defined order. Also depicted are the different possible orders of addition of Etn-P and Man substituents that are inferred from the structures of the GPI structural isoforms that accumulate in yeast mutants (109, 130, 133). Analyses of the GPIs that can be radiolabeled in mammalian cells have led to proposals for analogous pathways with branches defined by the presence or absence of Etn-P on Man-1 have (108, 110, 113). It is not clear, however, whether the GPI isomers lacking Etn-P on Man-1 represent an alternative route for the assembly of the final Etn-capped yeast GPI anchor precursor CP2.

Figure 3

(A) GPI transamidase. The enzyme consists of the catalytic center, GPI8/Gpi8p, and four other subunits (PIG-T/Gpi16p, GAA1/Gaa1p, PIG-S/Gpi17p and PIG-U/Gab1p) that are all essential for function. The yeast subunits can be divided into two subcomplexes, one containing Gpi8p, Gpi16p, and Gaa1p, and the other, Gab1p and Gpi17p (149, 151, 160). GPI8 and PIG-T are disulfide-linked. PIG-T stabilizes the complex. GAA1 has a role in GPI recognition since GPIP containing a mutant GAA1 subunit lacking the last TM span binds pro-proteins but not GPIs. PIG-U has a sequence motif that suggests that it may also participate in GPI recognition. PIG-S has no known function. Although none of the subunits bears a recognizable ER retention-retrieval motif, the TM domain of PIG-T may function to localize the complex to the ER (201). The figure shows predicted or tested topological models of the five mammalian subunits; the grouping of the subunits GPI8/PIG-T/GAA1 and PIG-S/PIG-U reflects the sub-complex architecture suggested for the yeast enzyme.

(B) GPI anchoring of proteins. ER-translocated proteins with a C-terminal GPI signal sequence are recognized by GPIIT on the luminal face of the ER. The C-terminal signal sequence is cleaved between residues ω and ω+1 and replaced with GPI. GPI is attached to the ω residue by an amide bond between the α-COOH of ω and the NH₂ of the capping EtN-P residue in GPI.
Figure 4

Postulated transglycosylation reaction in which a GPI-CWP becomes crosslinked to cell wall β1,6-glucan via its GPI glycan.
The GPI is likely to be cleaved between Man-1 and GlcN (189, 190) and transferred to a terminal or an internal β1,6-linked Glc in cell wall β1,6-glucan. Features of a GPI glycoprotein that correlate with cell wall anchorage include the presence of serine and threonine-rich regions in the protein, and the absence of two basic amino acids from the ω-minus region (30, 31, 191-195). It is not known whether the presence of Etn-P side-branches, or of long acyl chains or a ceramide on the GPI, are important for transglycosylation to occur, or whether Etn-P side-branches are retained after cross-linking.
Table 1
Protein machinery for GPI anchoring in yeast and mammals

<table>
<thead>
<tr>
<th>Step</th>
<th>Reaction</th>
<th>Mammalian protein</th>
<th>Yeast protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GlcNAc-PI synthesis (GPI-Gnt)</td>
<td>PIG-A&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Gpi3p&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-C</td>
<td>Gpi2p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-H</td>
<td>Gpi15p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-P</td>
<td>Gpi19p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-Q</td>
<td>Gpi1p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-Y</td>
<td>Eri1p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DPM2</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>GlcNAc-PI de-N-acetylation</td>
<td>PIG-L&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Gpi12p&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>3</td>
<td>GPI flipping</td>
<td>not identified</td>
<td>not identified</td>
</tr>
<tr>
<td>4</td>
<td>Inositol acylation</td>
<td>PIG-W</td>
<td>Gwt1p</td>
</tr>
<tr>
<td>5</td>
<td>α&lt;sub&gt;1,4&lt;/sub&gt; mannosyltransfer</td>
<td>PIG-M</td>
<td>Gpi14p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-X</td>
<td>(Pbn1p?)</td>
</tr>
<tr>
<td>6</td>
<td>Etn-P transfer to Man-1</td>
<td>PIG-N</td>
<td>Mcd4p</td>
</tr>
<tr>
<td>7</td>
<td>α&lt;sub&gt;1,6&lt;/sub&gt; mannosyltransfer</td>
<td>PIG-V</td>
<td>Gpi18p</td>
</tr>
<tr>
<td>8</td>
<td>α&lt;sub&gt;1,2&lt;/sub&gt; mannosyltransfer</td>
<td>PIG-B</td>
<td>Gpi10p</td>
</tr>
<tr>
<td>9</td>
<td>Etn-P transfer to Man-3</td>
<td>PIG-O</td>
<td>Gpi13p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-F</td>
<td>(Gpi11p?)</td>
</tr>
<tr>
<td>y</td>
<td>α&lt;sub&gt;1,2&lt;/sub&gt; mannosyltransfer</td>
<td>hSmp3</td>
<td>Smp3p</td>
</tr>
<tr>
<td>10</td>
<td>Etn-P transfer to Man-2</td>
<td>hGpi7</td>
<td>Gpi7p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-F</td>
<td>(Gpi11p?)</td>
</tr>
<tr>
<td>11</td>
<td>GPI transamidase (GPIT)</td>
<td>PIG-K&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Gpi8p&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAA1</td>
<td>Gaa1p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-S</td>
<td>Gpi17p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-T</td>
<td>Gpi16p&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIG-U</td>
<td>Gab1p</td>
</tr>
<tr>
<td>p</td>
<td>Inositol deacylation</td>
<td>PGAP1</td>
<td>Bst1p</td>
</tr>
<tr>
<td>p</td>
<td>sn-2 deacylation</td>
<td>PERLD1/PGAP3</td>
<td>Per1p</td>
</tr>
<tr>
<td></td>
<td>sn-2 C&lt;sub&gt;26&lt;/sub&gt; acylation</td>
<td>Gup1p</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----------------------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>sn-2 acylation</td>
<td>PGAP2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Cwh43p?)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Steps 1-11 correspond to those shown in Fig. 2A; steps labeled y are shown in Fig. 2B; steps labeled p concern modification of the GPI structure after it is attached to protein.
2 Proteins with a single transmembrane span; all other proteins listed in this Table are predicted to be bitopic or polytopic membrane proteins.
Figure 1

Thy-1 - INVIRDKLVKC\text{CGGISSL}VQNTSWL\text{LLLL}LSL\text{SFLQATDFISLN}\text{COOH} \\
\omega - 10 \rightarrow \omega - 1 \quad \omega + 3 \rightarrow \omega + 9 \\
\text{generally polar} \quad \text{moderate polarity} \quad \text{hydrophobic}

Gas1p - SASSSSSSKKNAATNVKANLAQVVF\text{TSI}S\text{L}S\text{IAAGVG}\text{FALV}\text{COOH}
Figure 2

A

ER

synthesis and flipping of lipid donors

GDP-GDP

PtdSer

PSD

UDP-UDP

Acetate

Palmitoyl-CoA

CoA

Dol-P Dol-P Dol-P

rxn 6

DAG PE

DAG PE

Dol-P Dol-P Dol-P

Dol-P Dol-P Dol-P

Dol-P Dol-P Dol-P

Dol-P Dol-P Dol-P

P

N

H8 H7 H6 H5

GPI13 GPI1

GPI7?

GPI11

SMP3

GPI10

GPI18

GPI14

MCD4

CP2

B

notes for fill colors (RGB): (0,200,50) ; (255,150,0) ; (0,0,250) ; (0,0,250)
Figure 4

plasma membrane

$\beta_{1,6}$Glc