

Biosynthesis of Glycosylphosphatidylinositol (GPI)-anchored Membrane Proteins in Intact Cells: Specific Amino Acid Requirements Adjacent to the Site of Cleavage and GPI Attachment

Krishna Kodukula, Louise D. Gerber, Rodolfo Amthauer, Larry Brink, and Sidney Udenfriend

Department of Neurosciences, Roche Institute of Molecular Biology, Roche Research Center, Nutley, New Jersey 07110

Abstract. Mutational studies were previously carried out at the ω site in intact cells (Micanovic, R., L. Gerber, J. Berger, K. Kodukula, and S. Udenfriend. 1990. *Proc. Natl. Acad. Sci. USA*. 87:157-161; Micanovic R., K. Kodukula, L. Gerber, and S. Udenfriend. 1990. *Proc. Natl. Acad. Sci. USA*: 87:7939-7943) and at the $\omega + 1$ and $\omega + 2$ sites in a cell-free system (Gerber, L., K. Kodukula, and S. Udenfriend. 1992. *J. Biol. Chem.* 267:12168-12173) of nascent proteins destined to be processed to a glycosylphosphatidylinositol (GPI)-anchored form. We have now mutated the $\omega + 1$ and $\omega + 2$ sites in placental alkaline phosphatase (PLAP) cDNA and transfected the wild-type and mutant cDNAs into COS 7 cells. Only glycine at

the $\omega + 2$ site yielded enzymatically active GPI membrane-anchored PLAP in amounts comparable to the wild type (alanine). Serine was less active and threonine and valine yielded very low but significant activity. By contrast the $\omega + 1$ site was promiscuous, with only proline being inactive. These and the previous studies indicate that the ω and $\omega + 2$ sites of a nascent protein are key determinants for recognition by COOH-terminal signal transamidase. Comparisons have been made to specific requirements for substitution at the -1 , -3 sites of amino terminal signal peptides for recognition by NH₂-terminal signal peptidase and the mechanisms of NH₂ and COOH-terminal signaling are compared.

PROTEINS that are anchored to the plasma membrane by a glycosyl-phosphatidylinositol (GPI)¹ moiety are derived from a nascent precursor that contains two signal peptides; one at the NH₂ terminus and the other at the COOH terminus. The NH₂-terminal signal peptide, after directing the nascent protein (prepro protein) into the ER, is cleaved by NH₂-signal peptidase to yield the pro protein. Secreted proteins that contain no other signal peptides, are processed by this well-studied mechanism (4). In the case of GPI-anchored proteins, the COOH-terminal signal peptide apparently directs the pro protein to a putative transamidase (8, 10, 15, 30). The latter, which is probably present in the membrane of the ER, catalyzes attachment of the GPI moiety to an internal residue of the pro protein near the COOH terminus (ω residue) with concomitant cleavage of the signal peptide. We proposed previously that the amino acid in the nascent protein which is present at the putative cleavage site of the COOH-terminal signal peptide, to which the GPI anchor is attached and becomes the new COOH terminus, be designated as the " ω " site (11). This simplifies comparisons between the NH₂-terminal and COOH-terminal signal peptides and also among different GPI-anchored proteins which vary in size. Examination of the 20

or so fully characterized nascent forms of GPI-anchored proteins reveals that, with few exceptions, the ω residue and the two residues adjacent and COOH-terminal to it, i.e., $\omega + 1$ and $\omega + 2$, are small amino acids (small amino acid domain) (10, 11). Mutational studies carried out at the ω site of the nascent forms of placental alkaline phosphatase (PLAP) in intact cells (18, 19) and of the engineered protein miniPLAP in a cell-free system (14) indicated that only glycine, alanine, serine, cysteine, aspartic acid, and asparagine are allowed. These same six amino acids are the only ones that have been reported to be present at the ω sites of the GPI-anchored proteins that have been characterized thus far (8, 10, 11). MiniPLAP [23.9 kD] is an engineered protein designed at the cDNA level from human PLAP [57.5 kD]. The engineered protein retains the NH₂ and COOH termini and all the antibody recognition sites of PLAP but is devoid of the active site, all the putative glycosylation sites and most of the cysteines. MiniPLAP was designed specifically for cell-free studies.

Recently we reported mutational studies at the $\omega + 1$ and $\omega + 2$ sites of the engineered protein, preprominiPLAP (11). Those studies, that were carried out in a cell-free system, indicate that there are even greater limitations at the $\omega + 2$ site than at the ω site. COOH-terminal processing occurred equally well with alanine or glycine at the $\omega + 2$ sites. However, of the other amino acids tested only the serine mutant

1. *Abbreviations used in this paper:* GPI, glycosylphosphatidylinositol; PLAP, placental alkaline phosphatase.

was partially processed. By contrast the $\omega + 1$ site was quite promiscuous, only the proline mutant being inactive. We have now carried out comparable mutational studies at the $\omega + 1$ and $\omega + 2$ sites of wild-type PLAP in transfected COS cells to determine whether the same site specificities apply to processing of a natural GPI-anchored protein in intact cells.

Materials and Methods

Mutagenesis and Recombinant Plasmids

The $\omega + 1$ and $\omega + 2$ mutants of pBC12BI/PLAP 513 Δ Kpn were made by treating the $\omega + 1$ and $\omega + 2$ mutants of pGEM4Z/miniPLAP 208 (11) with *Asp*718 and *Ecl*XI (Boehringer Mannheim Biochemicals, Indianapolis, IN) to produce a 632-bp fragment containing the mutation sites. The original pBC12BI/PLAP 513 Δ Kpn was also treated with *Asp*718 and *Ecl*XI to produce two fragments, 4,770 and 632 bp. The 632 bp fragment from each mutant and the 4,770 bp fragment from pBC12BI/PLAP 513 Δ Kpn were gel purified, eluted from the gel using Spin-X tubes (Costar, Cambridge, MA), ethanol precipitated, and redissolved in water. The solutions were quantified, mixed 3:1 (mol/mol) small fragment to large fragment, and ligated using T4 ligase (New England Biolabs, Beverly, MA). The resultant $\omega + 1$ and $\omega + 2$ mutants of pBC12BI/PLAP 513 Δ Kpn were then used to transform competent DH5 α cells (New England Biolabs). The plasmids were screened with appropriate restriction enzymes, and DNA sequences were verified by sequencing, by the dideoxy method, using the TAQuence Version 2.0 DNA Sequencing Kit (United States Biochemical, Cleveland, OH).

In each of the mutants studied, the substitutions were made specifically at $\omega + 1$ or $\omega + 2$ sites, and in each case the wild-type residues were maintained at adjacent positions. For example, when mutating alanine at $\omega + 1$, the wild-type amino acids at ω and $\omega + 2$ (aspartic acid and alanine, respectively) were preserved, thus the $\omega + 1$ mutants would have a sequence Asp-X-Ala and the $\omega + 2$ mutants have an Asp-Ala-X motif in the small amino acid domain. Similarly, the site mutations at the ω position had an X-Ala-Ala motif as reported earlier (18, 19).

Cell Culture and DNA Transfections

COS 7 cells were maintained in culture as described (12) and transfections were carried out in 35-mm well culture plates, by the DEAE-Dextran method (9). Approximately 300–350 ng of DNA per well was used for each transfection. Mock transfections were carried out substituting PBS for DNA in the transfection cocktail. Transfections were typically carried out in quadruplicate wells for each mutant.

Metabolic Labeling

Metabolic labeling of cells with [³⁵S]methionine was carried out 48–60 h after transfection. For steady state labeling, cells were starved of methionine by preincubating for 1 h at 37°C in 2.0 ml of methionine-free DME supplemented with 5% dialyzed FBS (Gibco Laboratories, Grand Island, NY). Fresh medium containing [³⁵S]methionine (~200 μ Ci per 35-mm well; 1.0 Ci = 37.0 GBq; Amersham Corp., Arlington Heights, IL) was then added and incubation was continued for 12–16 h.

Treatment of Cells with Phosphatidylinositol Specific Phospholipase C (PI-PLC)

PLAP, from either radiolabeled or unlabeled transfectants, was released with phosphatidylinositol specific phospholipase C (PI-PLC) purified from *Bacillus thuringiensis* (28). Each culture well was washed twice with 2.0 ml of PBS and twice with release buffer (25 mM Tris/HCl, pH 7.5, 10 mM glucose, 250 mM sucrose). Subsequently duplicate wells for each mutant were treated with 1.0 ml of release buffer containing protease inhibitors (antipain, aprotinin, bestatin, chymostatin, leupeptin, and pepstatin, each at 0.2 μ g/ml) and PI-PLC (125–150 U/well). One unit of PI-PLC activity is defined as that amount of enzyme that will hydrolyze 8×10^{-4} μ mol of phosphatidylinositol in 1 min at 37°C. PI-PLC was omitted from the above mixture while treating control wells for each mutant. Samples were incubated at 37°C for 2 h, supernatants were transferred to eppendorf tubes (Brinkman Instruments, Inc., Westbury, NY) and cell debris was removed by centrifugation (14,000 *g* for 10 min). 15–20- μ l aliquots were taken from

the supernatant to measure PLAP activity. The remaining cells on the culture plate were scraped into 2.0 ml of PBS, washed once with PBS, and resuspended in 100 μ l of homogenization buffer (50 mM Tris/HCl, pH 7.5, 1.0 μ M magnesium chloride, and 20 mM zinc sulfate). Homogenization was carried out using a hand held micro-homogenizer and aliquots of the homogenate (2.0–4.0 μ l) were taken for measuring PLAP activity.

Alkaline Phosphatase and Protein Assay

PI-PLC-treated supernatants or total cell homogenates were assayed for alkaline phosphatase activity using *p*-nitrophenyl phosphate as substrate (16). 1 mU of the enzyme produces an absorbance change of 0.04 per minute at 405 nm at room temperature (23°C). Traces of endogenous alkaline phosphatases in COS cells were inhibited by adding 10 mM (final concentration) L-homoarginine to the reaction mixture. The assay is therefore specific for PLAP. The cell homogenates were treated with 0.1% SDS and total protein concentrations were estimated by the Bradford method (14).

Immunoprecipitation and SDS-PAGE

An aliquot of each PI-PLC treated supernatant (900 μ l) was added to 100 μ l of 10% SDS–20% β -mercaptoethanol (vol/vol), and the mixture was first vortexed and then boiled for 5 min. A 200- μ l aliquot of the boiled mixture was diluted to 1.0 ml with radio immunoprecipitation buffer (3) and PLAP related proteins were precipitated by adding anti-PLAP polyclonal antibody (Accurate Chemical Corp., Westbury, NY) at 1:500 and incubating for 12–16 h at 4°C.

Whole cell lysates were prepared for immunoprecipitation in the following manner. Metabolically labeled cells were lysed directly on each plate with 400 μ l of 2% SDS–5% β -mercaptoethanol (vol/vol), transferred to eppendorf tubes and boiled for 5 min. Boiled lysate (200 μ l) was diluted to 1.0 ml with RIPA buffer (1% Triton X-100, 0.5% deoxycholate, 0.2% SDS, 100 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, pH 7.5) and treated with antibody, as described above. The immunoprecipitated samples were treated with protein A-Sepharose CL-4B (Sigma Chemical Co., St. Louis, MO) and incubated for 1 h on a rotator. Samples were then extracted into 2 \times SDS-PAGE sample buffer and electrophoresis was carried out on 7.5% polyacrylamide gels. After electrophoresis gels were fixed, treated with amplify (Amersham Corp.), dried, and exposed to X-OMAT-AR film (Eastman Kodak Co., Rochester, NY).

Immunofluorescent Labeling of Cells

Immunofluorescent labeling of transfected cells was carried out in 35-mm culture wells as described (9). At 72 h post-transfection, cells were fixed with 95% ethanol–5% acetic acid at –20°C for 10 min and then washed three times with cold PBS. To block non-specific binding, cells were incubated with normal goat serum for 16 h at 4°C and were then washed with PBS and incubated with anti-PLAP antibody (diluted 1:500) in BSA buffer (1% BSA, 0.3% gelatin, 25 mM sodium phosphate, pH 7.5, 150 mM sodium chloride). After this, the cells were washed with PBS to remove unbound antibody and finally incubated with fluorescein-conjugated goat anti-rabbit IgG (Organon Teknica, West Chester, PA) diluted 1:50 in triton buffer (0.3% Triton X-100 in PBS). Finally, the labeled cells were mounted under glycerol/PBS (9:1, vol/vol) containing 5% *n*-propyl gallate and examined under a fluorescence microscope. Photomicrography was carried out using Ektachrome 800 P film (Eastman Kodak Co.).

Results

Expression of $\omega + 1$ and $\omega + 2$ Site-directed Mutants of PLAP and Processing to GPI-anchored Forms

The mutants that were investigated are shown in Fig. 1. Each was transiently expressed in COS cells and monitored by [³⁵S]methionine labeling followed by immunoprecipitation with the polyclonal antibody to PLAP. As shown in Fig. 2 A, all the $\omega + 1$ mutants yielded PLAP-related proteins with the same molecular masses as wild-type PLAP and the extent of expression was comparable, within experimental error, for all mutants. Each of the $\omega + 1$ mutant forms (with the exception of proline) yielded a broadened band in the

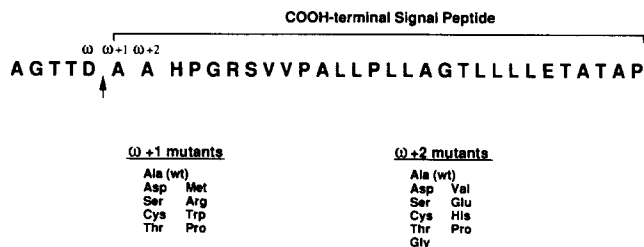


Figure 1. The COOH terminus of PLAP indicating the signal peptide that is cleaved during GPI addition. The arrow indicates the site of cleavage and ω represents the terminal amino acid residue of the mature protein to which the GPI moiety is attached.

range of 65 to 70 kD. Based on earlier studies (19), this represents a mixture of the proPLAP form (ca. 65 kD) and the mature, GPI-anchored form of PLAP (ca. 68 kD). The resolved, faster running band (ca. 60 kD) represents the nascent protein, preproPLAP. Wild-type PLAP, like all GPI-anchored proteins, can be released from the cell surface by PI-PLC (10, 15). Accordingly, cells transfected with each of the $\omega + 1$ mutant cDNAs and labeled with [35 S]methionine were treated with PI-PLC. Aliquots of the PI-PLC-treated supernatants were immunoprecipitated with antibody and the labeled PLAP-related proteins were resolved on SDS-PAGE. Fig. 2 *B* shows protein released into the medium. PLAP was not released from control cells that were not treated with PI-PLC (data not shown). Except for proline, PI-PLC treatment released mature forms of PLAP from all $\omega + 1$ mutants. The amount released from the tryptophan mutant was small but significant. The PI-PLC-released proteins appear as single broad bands at 68 kD. Broadening is most likely due to glycosylation during processing.

Expression of $\omega + 2$ mutants is shown in Fig. 3 *A*. As with $\omega + 1$ mutants all mutant forms were expressed. It should be noted, however, that, in contrast to $\omega + 1$ mutants (Fig. 2 *A*) not all of the $\omega + 2$ bands were broadened. The explanation for this becomes apparent when one examines Fig. 3 *B* which shows the release of the GPI-anchored forms of PLAP from the cell surface by PI-PLC. Only four of the mutants yielded PI-PLC releasable PLAP that migrated at 68 kD;

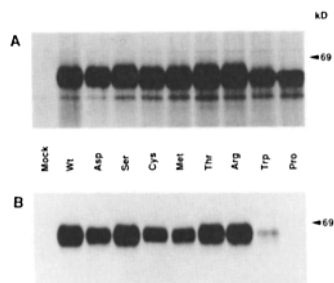


Figure 2. Expression of $\omega + 1$ mutants of PLAP in COS cells. Transfected COS cells were metabolically labeled with [35 S]methionine and proteins from whole cell lysates or PI-PLC treated supernatants were immunoprecipitated with anti-PLAP antiserum and subjected to SDS-PAGE. (A) Total PLAP from whole cell

lysates. Wild-type PLAP and the various mutants yield broad bands indicating that they are extensively glycosylated and represent a mixture of two forms of the protein—the NH₂ terminally processed pro form and the mature GPI-linked form. The mutants that are not processed at the COOH terminus exhibit a compact band of the pro form at ~65 kD. The minor product migrating fastest on the gel (~60 kD) is nascent preproPLAP. (B) The wild-type and mutant forms of PLAP that are released by PI-PLC from the cells into the medium are shown in this panel. They represent the GPI-linked forms of PLAP (~67 kD).

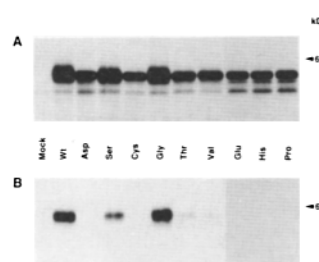


Figure 3. Expression of $\omega + 2$ mutants of PLAP in COS cells. Cells were labeled with [35 S]methionine following which cell lysates and PI-PLC treated supernatants were immunoprecipitated as described under methods and the legend for Fig. 2. (A) Total PLAP from whole cell lysates. (B)

Wild-type and $\omega + 2$ mutant forms of PLAP that were released by PI-PLC from the cell surface into the medium.

alanine \cong glycine > serine >> threonine and valine. It was the wild-type, serine and glycine mutants which exhibited broadened peaks in Fig. 3 *A*, indicating that each represented a mixture of proPLAP (65 kD) and GPI-anchored PLAP. Those $\omega + 2$ mutants that showed little or no release by PI-PLC (Fig. 3 *B*) gave sharp bands at ca. 65 kD in Fig. 3 *A*, indicating that they represented largely the proforms of PLAP and had not been processed to the GPI-anchored form.

Expression of Enzymatically Active PLAP

Processing of the COOH terminus is a prerequisite for PLAP to be expressed on the cell surface (18, 19). Transfected cells, both $\omega + 1$ and $\omega + 2$ mutants, were homogenized and total PLAP enzyme activity was assayed. As shown in Table I *A*, all $\omega + 1$ mutants, except proline and tryptophan, exhibited

Table I. PLAP Enzyme Activity in COS Cells Transfected with Various Mutants

Mutant	(A) Total cellular enzyme activity		(B) Enzyme activity released by PI-PLC	
	$\omega + 1$	$\omega + 2$	$\omega + 1$	$\omega + 2$
	(milliunits/ $\sim 10^6$ cells)			
Ala*	103	84	123	91
Arg	38	—	48	—
Asp	24	9	66	0
Cys	18	0	38	5
Glu	—	2	—	0
Gly	—	63	—	78
His	—	5	—	0
Met	17	—	28	—
Pro	1	4	0	0
Ser	27	16	114	31
Thr	22	10	39	11
Trp	6	—	6	—
Val	—	10	—	8

For total enzyme activity each value represents the average of duplicate analyses obtained in one experiment. With respect to PIPLC activity released into the medium at least three different experiments were carried out with each mutant. Because of differences in transfection efficiency, values from only one of the experiments are shown. Although the absolute values obtained varied from transfection to transfection, the relative value for each mutant with respect to the wild type was fairly constant. Values for mock transfections averaged about 1.5 milliunit and were subtracted as background.

It should be noted that more measurable enzyme activity was released into the medium by PIPLC than was present in the cells before treatment with the lipase. This is so because in homogenates of cells the GPI-anchored PLAP is attached to membrane fragments and has a lower specific activity than the PLAP that is solubilized by lipase treatment. In addition, because cell lysates are slightly turbid, measurement of enzyme activity is not as precise as it is in supernatants after PI-PLC release.

* Indicates the wild type amino acid present at the $\omega + 1$ or $\omega + 2$ positions.

~20 to 30% of the activity of wild-type PLAP. More significantly, treatment with PI-PLC did not release PLAP enzyme activity from cells transfected with the proline $\omega + 1$ mutant (Table I B). Some enzyme was released from cells transfected with the tryptophan mutant; considerable amounts of PLAP activity were released from all other mutants. These findings are in agreement with the data in Fig. 2 B. Total PLAP enzyme activity in $\omega + 2$ transfected cells is also shown in Table I A. Only the glycine mutant exhibited activity comparable to the wild-type. Except for serine, all the others yielded 10% or less of the activity of wild-type PLAP. Again, monitoring enzyme activity released by PI-PLC is most revealing. Only the glycine and serine $\omega + 2$ mutants yielded appreciable activity while most other mutants were either totally inactive or exhibited 10% or less of the wild-type activity. These findings are again in accord with the metabolic labeling experiments shown in Fig. 3 B.

Immunocytochemical Localization of Mutant Forms of PLAP

Mature GPI-anchored PLAP is present on the outer surface of the plasma membrane. To determine what forms of PLAP were expressed by COS cells after transfection with mutant cDNAs, the cells were examined by immunofluorescent microscopy with and without permeabilization. Only those forms of PLAP that are on the cell surface should be visible before permeabilization, while permeabilized cells should also permit visualization of intracellular forms of PLAP (prepro and proPLAP) (29). Representative mutants were investigated. As shown in Fig. 4, PLAP from the methionine and threonine mutants at the $\omega + 1$ site (residue 485), as well as alanine were present predominantly on the cell surface similar to the wild-type. However, only after permeabilization was the $\omega + 1$ proline mutant seen within the cells indicating that it was expressed but not processed to the membrane anchored form. Of the $\omega + 2$ mutants, only glycine exhibited surface fluorescence approaching that of the wild-type (alanine) (Fig. 5). The valine mutant yielded some fluorescence and the aspartic acid mutant essentially none. After permeabilization all $\omega + 2$ mutants yielded considerable fluorescence indicating, once more, that processing rather than expression was the limiting factor in producing PLAP that reached the cell surface. These immunocytochemical observations agree with and support the biochemical finding presented above.

Discussion

A comparison between our earlier findings with miniPLAP mutants in cell-free preparations (11) and our present findings in COS cells transfected with mutant cDNAs of PLAP show fairly good agreement in Table II. Both studies demonstrate a relative lack of specificity at the $\omega + 1$ site and rather stringent requirements at the $\omega + 2$ site. It is of interest that in the cell-free studies with miniPLAP only glycine could substitute effectively at the $\omega + 2$ site for wild-type alanine. By contrast, in the current studies with PLAP mutants in intact cells, serine was ~30 to 50% as effective as alanine or glycine and the threonine and valine mutants demonstrated low but measurable activity. There could be many reasons why $\omega + 2$ substitution is more limited in the cell-free system than it is in intact cells. An obvious one is

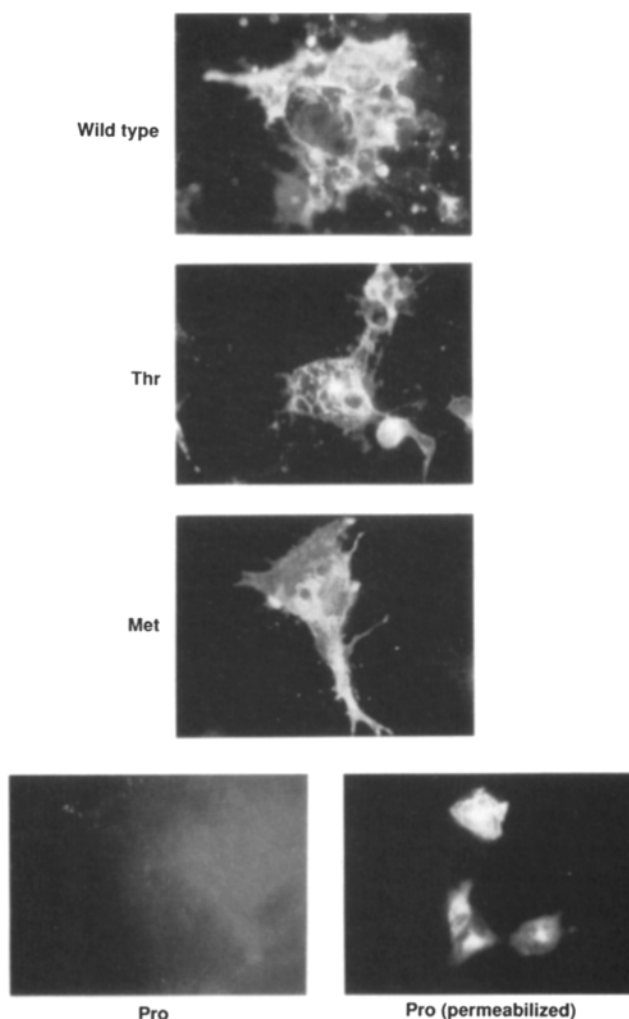


Figure 4. Detection of PLAP encoded by wild-type and selected $\omega + 1$ mutant cDNAs in transfected COS cells. Transfected cells were fixed and nonspecific binding sites blocked with normal goat serum. Intact and permeabilized cells were treated with anti-PLAP antibodies and bound antibody was visualized by staining with fluorescein-conjugated rabbit anti-rabbit immunoglobulin antibodies.

that although PLAP and miniPLAP have the same NH_2 - and COOH -terminal sequences, they are different proteins. Mature PLAP contains 484 residues and is *N*-glycosylated whereas mature miniPLAP contains 179 residues and is not glycosylated (14). Conceivably, secondary structure and the extent of glycosylation of a protein during the transamidation reaction with the GPI moiety might be factors in selection of $\omega + 2$ sites. Ideally one should use the same protein to compare cellular and cell-free processing. Unfortunately, native GPI-anchored proteins are frequently large and also *N*-glycosylated. Accordingly, they do not lend themselves readily to monitoring small changes in molecular mass that occur during NH_2 - and COOH -terminal processing in a cell-free system (19). Conversely, miniPLAP, which was designed specifically for cell-free studies is not processed as well in intact cells (14), perhaps because of its lack of *N*-glycosylation sites. Nevertheless, agreement between the two studies is fairly good as is agreement between experimentally determined values and those observed in characterized

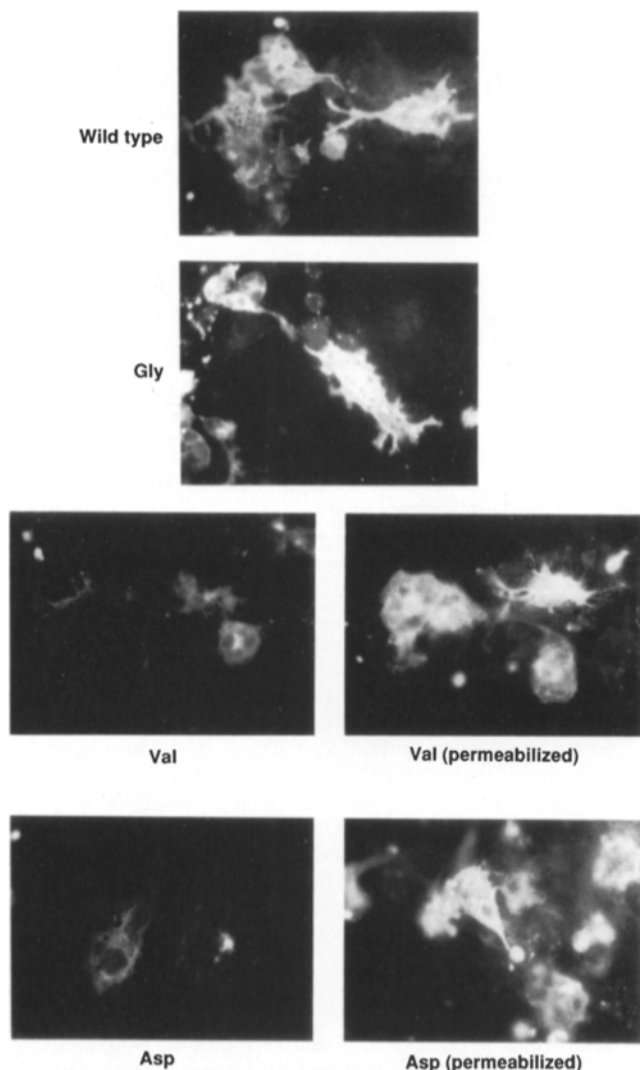


Figure 5. Detection of PLAP encoded by wild-type and selected $\omega + 2$ mutant cDNAs in transfected COS cells. Details as in legend to Fig. 4.

GPI proteins (Table II). While not all possible mutants at the $\omega + 1$ and $\omega + 2$ sites were produced at least one representative of each class of amino acids was investigated. Tryptophan is representative of the aromatic amino acids phenylalanine and tyrosine; methionine and valine also represent the long chain amino acids leucine and valine; arginine represents lysine as well; asparagine is representative of glutamine; and aspartic acid is representative of glutamic acid.

Selection of the ω , $\omega + 1$, and $\omega + 2$ sites for mutational studies was based on observations on the twenty or so proteins that have been fully characterized, i.e., have been both cloned and have had their ω sites determined (8, 11). Although we have not carried out mutational studies at sites preceding the ω site (i.e., $\omega - 1$ or $\omega - 2$) these are highly variable in all the fully characterized GPI-anchored proteins. Based on deletion mutations preceding the ω site in DAF, Caras et al. (7) concluded that "residues immediately NH₂-terminal to this region do not appear to play an essential role" as determinants in COOH-terminal processing. Earlier mutational studies at the ω site revealed that only one of six

Table II. Experimentally Determined Hierarchy of Amino Acid Substituents at the $\omega + 1$, and $\omega + 2$ positions

Mutant	ω	$\omega + 1$	$\omega + 2$
Ala	0.4+	1.0 (1.0)+*	1.0 (1.0)+*
Arg	ND	0.5 (0.7)	ND
Asn	0.8+	+	ND
Asp	0.4+*	0.4 (0.6)+	0.1 (0)
Cys	0.2+	0.3 (1.2)	0 (0)
Gln	0	ND	ND
Glu	0	(0.4)+	0 (0)
Gly	0.4+	+	0.7 (0.6)+
His	ND	ND	0 (0)
Leu	0.1	ND	ND
Lys	0	ND	ND
Met	0	0.3 (0.7)	ND
Pro	0	0 (0)	0 (0)
Ser	1.0+	0.6 (1.2)+	0.3 (0)+
Thr	0	0.3 (0.7)+	0.1 (0)+
Trp	0	0.1 (0)	ND (0)
Tyr	0	ND	ND
Val	0.1	ND	0.1 (0)

Values for the ω position have been calculated from an earlier report (18). Each value for $\omega + 1$ and $\omega + 2$ represents the average of all experiments carried out on that mutant not only the ones shown in Table I. For each site, the mutant yielding highest activity was arbitrarily assigned a value of 1.0; all other values are relative to the most active mutant. The calculation of probabilities for $\omega + 1$ and $\omega + 2$ substitution discussed in the text uses only the data obtained in the current experiments in cells transfected with PLAP cDNA and the various mutant cDNAs. Numbers in parenthesis under $\omega + 1$ and $\omega + 2$ are values previously reported for in vitro experiments with miniPLAP (11) and are shown here only to compare with the data obtained with PLAP in intact cells. A + indicates that this amino acid is present at the corresponding site of a characterized GPI protein.

* Denotes the wild type amino acid present at the respective position in PLAP.

small amino acids (Table II) could serve effectively (14, 18, 19). Direct observations of the $\omega + 1$ sites in the fully characterized GPI-anchored proteins (11) show some preference for small amino acids. However, this selectivity has not been borne out in site mutational studies in which arginine, methionine, and threonine proved quite effective. In fact, glutamic acid and asparagine are present at $\omega + 1$ sites in characterized proteins (human Thy 1 and VSG MIT 221, respectively) (11). Interestingly, mutational studies at the $\omega + 2$ site indicate an even greater selectivity than has been observed at that site among the characterized GPI-anchored proteins. Thus, ω and $\omega + 2$ appear to serve as the amino acid determinants for cleavage and GPI attachment.

One should consider the possibility that a mutation at any of the sites, ω , $\omega + 1$, or $\omega + 2$ that leads to a GPI-linked protein, might shift the reading frame so as to convert an upstream site (i.e., $\omega - 3$ or $\omega - 4$) to a new GPI site. Obviously, frameshifts of this kind would lead to cleavage at a site preceding the normal GPI attachment site, which in PLAP is Asp⁴⁸⁴. Such frameshifts towards the amino terminus can be ruled out in these studies because the mutants were found to react with an antibody directed to the COOH-terminal sequence of mature PLAP, CDLAPPAGTTD⁴⁸⁴. We have shown that when Asp⁴⁸⁴ is removed the epitope is lost (unpublished results). Since $\omega + 1$ and $\omega + 2$ mutants reacted with this antibody, cleavage and GPI addition could not occur at a residue preceding the ω site (data not shown).

The similarity between COOH-terminal signal transamidase and NH₂-terminal signal peptidase with respect to amino acid substituents at their cleavage sites was first

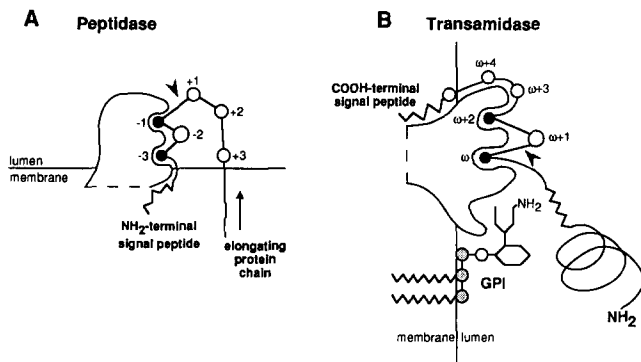


Figure 6. Binding of NH_2 - and COOH -terminal signal peptides with their respective cleavage enzymes. (A) Representation of the luminal portion of NH_2 -terminal signal peptidase with a signal peptide bound at the -1 and -3 sites is adapted from the model proposed by von Heijne (31). The arrowhead indicates the site of cleavage between the -1 site of the signal peptide and the $+1$ site. The latter represents the NH_2 terminus of the mature protein. (B) Representation of the luminal portion of the putative COOH -terminal transamidase with a signal peptide bound at the ω and $\omega + 2$ sites. The arrowhead indicates the site of cleavage between the $\omega + 1$ site of the signal peptide and the ω site. The latter represents the COOH terminus of the mature protein, which concomitant with cleavage, combines with the NH_2 group of the GPI moiety. The transamidase is shown in the lumen in a vertical position to emphasize similarities between the binding sites of the two cleavage enzymes.

pointed out by Ferguson and Williams (10) and experimentally verified by us, initially using miniPLAP in a cell-free system (11) and currently with PLAP in intact cells. In both signal peptides only the first and third residues at the cleavage site have a requirement for small amino acids. In the case of NH_2 -terminal signal peptidase this is referred to as the $-1, -3$ rule for predicting the site of cleavage of a cDNA deduced sequence in a protein containing a typical hydrophobic NH_2 -terminal signal peptide (31). According to von Heijne (31, 32) the significance of the $-1, -3$ requirements are that these two amino acids represent sites of interaction with the NH_2 -terminal signal peptidase. The mechanism proposed by von Heijne for interaction of the NH_2 -terminal signal peptide with the peptidase is shown in Fig. 6 A (31, 32). The $-1, -3$ rule, which is generally accepted among molecular biologists as a useful means of predicting the amino terminus of a mature protein based on its cDNA deduced sequence, has an acknowledged reliability of 75 to 80% (31, 32). Our studies with miniPLAP in a cell-free processing system suggested a comparable mechanism whereby the ω and $\omega + 2$ substituents represent key recognition sites for interaction with the putative COOH -terminal signal transamidase (Fig. 6 B). The present studies confirm such a mechanism so that we may consider the ω and $\omega + 2$ substituents as fairly reliable predictors of the COOH -terminal processing site in a nascent GPI-anchored protein.

The cDNA deduced COOH -terminal regions of eight characterized GPI-anchored proteins are shown in Fig. 7. Generally, COOH -terminal signal peptides contain 15–30 residues. The region between $\omega + 1$ and $\sim \omega + 8$ is generally, but not always, rich in charged amino acids and/or proline. A similar stretch of amino acids rich in charged amino acids and proline precedes the cleavage sites of NH_2 -

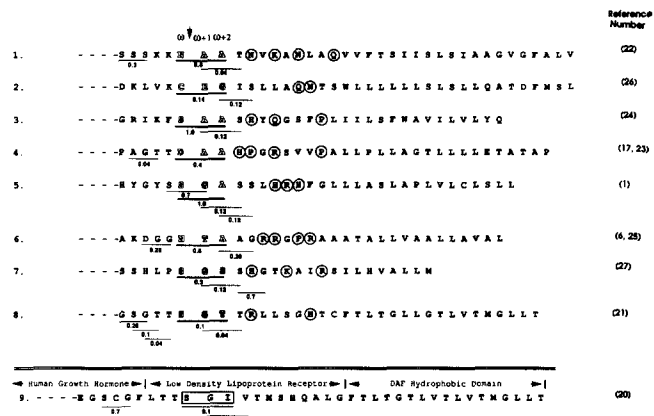


Figure 7. cDNA-deduced COOH -terminal regions of several GPI-anchored membrane proteins. The arrow indicates sites of cleavage. The highlighted amino acids represent ω , $\omega + 1$, and $\omega + 2$ sites, respectively, and circled amino acids are within putative hinge regions. The underlined sequences in each protein bracket residues with some degree of probability as representing the actual cleavage site based on hierarchical values in Table II. The numbers under each line represent the relative probability of each putative ω , $\omega + 2$ site. The experimentally determined ω , $\omega + 2$ sequence in each protein is underlined in bolder type. The proteins are as follows: 1, GAS 1 (*Saccharomyces cerevisiae*); 2, Thy-1 (human); 3, 5'-nucleotidase; 4, alkaline phosphatase (human placenta); 5, renal dipeptidase; 6, promastigote surface protease GP63 (*L. major*); 7, VSG YNat 1.1 (*Trypanosoma congolense*); and 8, decay accelerating factor. Protein 9 represents the deduced COOH -terminal amino acid sequence of a chimeric cDNA obtained by combining portions of three unrelated cDNAs. The ω site (boxed) is based on the site of cleavage proposed by the authors but not experimentally verified. The underlined sequences represent possible GPI attachment sites based on the ω , $\omega + 2$ rule.

terminal signal peptides and is referred to as the hinge region (13). In the case of GPI-anchored proteins, a putative hinge region (circled amino acids, Fig. 7) is followed by a stretch of 8 to 20 highly hydrophobic amino acids. Occasionally, the hydrophobic region contains or terminates in a charged amino acid or proline. Determining the most probable site for GPI anchoring in a cDNA deduced protein should take all the above factors into account: the size of the hydrophobic sequence, the hinge region, as well as the probabilities for the various amino acids to reside at the ω , $\omega + 1$, and $\omega + 2$ sites. It appears that, except for proline and tryptophan, almost all substitutions at the $\omega + 1$ site are effective. We can therefore limit considerations to the ω and $\omega + 2$ sites and rule out putative ω sites that are followed by a proline or tryptophan. Using the hierarchical values determined for these two sites (Table II), the probability of a specific amino acid being the ω site for GPI anchoring in a given nascent protein would then be arrived at by multiplying the individual probabilities of ω and $\omega + 2$ constituents. Thus, if serine at an ω site has a probability of 1 and alanine at $\omega + 2$ also has a probability of 1 the resultant probability would be $1 \times 1 = 1.00$ (most probable). Cysteine at an ω site has a probability of 0.2 and with an alanine at a putative $\omega + 2$ site the overall probability would be 0.2 (less probable). However, if cysteine were at the ω site and the $\omega + 2$ site were serine (0.3) then the overall probability for each motif would be 0.06 (even less probable). Thus, in Fig. 7 an ω , $\omega + 2$ rule would

have clearly predicted the site of cleavage/attachment in the first six proteins. Based only on the ω , $\omega + 2$ rule, protein 7 has a site of higher probability (S-H-G) immediately following the experimentally determined site (S-G-S). It should be noted, however, that the S-H-G site is within a typical hinge area (circled amino acids) and is also very close to the COOH terminus of the signal peptide thereby making the S-G-S site the more likely. Protein eight contains three possible sites yet it is one with a lower probability that is actually the site of cleavage and GPI attachment (21). Thus, proteins 7 and 8 indicate that other factors, including secondary structure, or location within a hinge region must also play significant roles as determinants. Nevertheless, application of the ω , $\omega + 2$ rule, based on data in Table II, to the 20 fully characterized GPI-anchored proteins (reference 11, Table I) indicate that it is a fairly good predictor of the site of cleavage and GPI addition. In three proteins (reference 11, Table I; proteins 9, 17, and 19), the observed ω site was the only one with any degree of probability. In twelve other instances (reference 11, Table I; proteins 1, 3, 5-8, 10, 11, 14-16, and 18) more than one site was possible but the verified site had by far the highest probability. In two of the proteins (reference 11, Table I; proteins 12 and 13), the verified ω site was one of two with about equal probabilities. However, in three proteins (reference 11, Table I; proteins 2, 4, and 20) the actual cleavage site had a lower probability than one at a nearby site(s). In two of these cases the more probable site was towards the NH₂ terminus and in one case, toward the COOH terminus. It should be noted, however, that in all the characterized proteins the verified ω site was one with some degree of probability. Thus, an ω , $\omega + 2$ rule by itself would have predicted the correct cleavage site in 75% of the cases and would have limited selection to two possible sites in another 10%. If one rules out probable sites within a hinge region, predictability is even greater.

This brings up the recent report by Moran and Caras (20) on an interesting series of chimeric proteins in which elements from non-anchored proteins are fused to generate a functional signal(s) for GPI membrane anchor attachment. Protein 9 in Fig. 7 shows one of their constructs. What is interesting and significant is that, core hydrophobic portions from HN₂-terminal signal peptides from secreted proteins such as prolactin and human growth hormone could be substituted successfully for the COOH-terminal signal peptide of DAF. These findings further emphasize the similarities between NH₂- and COOH-terminal processing. However, Moran and Caras (20) went on to conclude that unlike the NH₂-terminal processing signal of two alternate small amino acids (-1, -3), COOH-terminal processing and GPI addition requires only two consecutive small amino acids (i.e., ω , $\omega + 1$). Such a conclusion is not borne out by published reports of other GPI anchored proteins (Fig. 7 and references 8, 10, 30) or by our experimental findings. It must be pointed out, however, that the serine claimed to be the site of cleavage in most of the chimeric proteins produced by Moran and Caras (20), was only inferred, and not experimentally verified. Nevertheless, the sequence of S-G-I in one of the constructs does have some probability (as noted in Fig. 7). If one assumes that isoleucine at $\omega + 2$ (which we did not investigate) is as probable as valine or methionine (0.1 then the overall probability would be 0.1. It should be noted however, that, seven residues towards the NH₂ termi-

nus there is a far more likely ω , $\omega + 2$ site, S-C-G, with a probability of 0.7. All the other constructs contain several probable sites in addition to the ones listed above. We fail to see how a study with chimeric proteins in which the ω sites were not experimentally determined and where no site mutations were made can claim that only two consecutive small amino acids can serve as a signal for GPI anchoring.

Studies with purified GPI transamidase and synthetic peptide analogs of COOH-terminal signal peptides will be necessary to more firmly establish the ω , $\omega + 2$ requirement for GPI anchoring. Until then application of an ω , $\omega + 2$ rule can be of value in helping to determine the site of GPI attachment in a cDNA deduced sequence of a protein. By narrowing the number of possible sites to one or two it limits the experiments required to establish the GPI site experimentally (17, 23, 24). Information from the ω , $\omega + 2$ rule can also be used to develop site-directed antibodies to help monitor studies on the biosynthesis of a GPI-anchored protein (2).

We thank Sarala Kodukula for valuable technical assistance, Michelle Smeyne and Dr. Richard Smeyne for their help with the fluorescence microscope, Drs. Frank Margolis and Radmila Micanovic for valuable discussions, and Enid Alston for expert secretarial assistance.

Received for publication 21 August 1992 and in revised form 9 October 1992.

References

- Adachi, H., T. Katayama, C. Inuzuka, S. Oikawa, M. Tsujimoto, and H. Nakazato. 1990. Identification of membrane anchoring site of human renal dipeptidase and construction and expression of cDNA for its secretory form. *J. Biol. Chem.* 265:15341-15345.
- Bailey, C. A., A. Howard, R. Micanovic, J. Berger, E. Heimer, A. Felix, L. Gerber, L. Brink, and S. Udenfriend. 1988. Site-directed antibodies for probing the structure and biogenesis of phosphatidylinositol glycan-linked membrane proteins: Application to placental alkaline phosphatase. *Anal. Biochem.* 170:532-541.
- Bailey, C. A., L. Gerber, A. D. Howard, and S. Udenfriend. 1989. Processing at the carboxyl terminus of nascent placental alkaline phosphatase in a cell-free system: Evidence for specific cleavage of a signal peptide. *Proc. Natl. Acad. Sci. USA.* 86:22-26.
- Blobel, G. 1983. Control of intracellular protein traffic. *Methods Enzymol.* 96:663-682.
- Bradford, M. M. 1976. Rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248-254.
- Button, L. L., and W. R. Macmaster. 1988. Molecular cloning of the major surface antigen of Leishmania. *J. Exp. Med.* 167:724-729.
- Caras, I. W., G. N. Weddel, and S. R. Williams. 1989. Analysis of the signal for attachment of a glycosylphospholipid membrane anchor. *J. Cell. Biol.* 108:1387-1396.
- Cross, G. A. M. 1990. Glycolipid anchoring of plasma membrane proteins. *Annu. Rev. Cell Biol.* 6:1-39.
- Cullen, B. R. 1987. Use of eukaryotic expression technology in the functional analysis of cloned genes. *Methods Enzymol.* 152:684-704.
- Ferguson, M. A. J., and A. F. Williams. 1988. Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. *Annu. Rev. Biochem.* 57:285-320.
- Gerber, L., K. Kodukula, and S. Udenfriend. 1992. Phosphatidylinositol-glycan (PI-G) anchored membrane proteins: amino acid requirements adjacent to the site of cleavage and PI-G attachment in the COOH-terminal signal peptide. *J. Biol. Chem.* 267:12168-12173.
- Gluzman, Y. 1981. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell.* 23:175-182.
- Haeuptle, M.-T., N. Flint, N. M. Gough, and B. Dobberstein. 1989. A tripartite structure of the signals that determine protein insertion into the endoplasmic reticulum membrane. *J. Cell Biol.* 108:1227-1236.
- Kodukula, K., R. Micanovic, L. Gerber, M. Tamburrini, L. Brink and S. Udenfriend. 1991. Biosynthesis of phosphatidylinositol glycan-anchored membrane proteins. Design of a simple protein substrate to characterize the enzyme that cleaves the COOH-terminal signal peptide. *J. Biol. Chem.* 266:4464-4470.
- Low, M. G. 1989. The glycosyl-phosphatidylinositol anchor of membrane proteins. *Biochim. Biophys. Acta.* 988:427-454.
- McComb, R. B., and G. N. J. Bowers. 1972. Study of optimum buffer con-

- ditions for measuring alkaline phosphatase activity in human serum. *Clin. Chem.* 18:97-104.
17. Micanovic, R., C. A. Bailey, L. Brink, L. Gerber, Y.-C. E. Pan, J. D. Hulmes, and S. Udenfriend. 1988. Aspartic acid-484 of nascent placental alkaline phosphatase condenses with a phosphatidylinositol glycan to become the carboxyl terminus of the mature enzyme. *Proc. Natl. Acad. Sci. USA.* 85:1398-1402.
 18. Micanovic, R., L. Gerber, J. Berger, K. Kodukula, and S. Udenfriend. 1990. Selectivity of the cleavage/attachment site of phosphatidylinositol-glycan anchored membrane proteins determined by site-specific mutagenesis at Asp-484 of placental alkaline phosphatase. *Proc. Natl. Acad. Sci. USA.* 87:157-161.
 19. Micanovic, R., K. Kodukula, L. Gerber, and S. Udenfriend. 1990. Selectivity at the cleavage/attachment site of phosphatidylinositol-glycan anchored membrane proteins is enzymatically determined. *Proc. Natl. Acad. Sci. USA.* 87:7939-7943.
 20. Moran, P., and I. W. Caras. 1991. Fusion of sequence elements from non-anchored proteins to generate a fully functional signal for glycoposphatidylinositol membrane anchor attachment. *J. Cell Biol.* 115:1595-1600.
 21. Moran, P., H. Raab, W. J. Kohr, and I. W. Caras. 1991. Glycophospholipid membrane anchor attachment. Molecular analysis of the cleavage/attachment site. *J. Biol. Chem.* 266:1250-1257.
 22. Nuoffer, C., P. Jenö, A. Conzelmann, and H. Reizmann. 1991. Determinants of glycophospholipid anchoring of the *Saccharomyces cerevisiae* GAS1 protein to the plasma membrane. *Mol. Cell Biol.* 11:27-37.
 23. Ogata, S., Y. Hayashi, N. Takami, and Y. Ikehara. 1988. Chemical characterization of the membrane-anchoring domain of human placental alkaline phosphatase. *J. Biol. Chem.* 263:10489-10494.
 24. Ogata, S., Y. Hayashi, Y. Misumi, and Y. Ikehara. 1990. Membrane-anchoring domain of rat liver 5'-nucleotidase: identification of the COOH-terminal Serine-523 covalently attached with a glycolipid. *Biochemistry.* 29:7923-7927.
 25. Schneider, P., M. J. Ferguson, M. J. McConville, A. Mehlert, S. W. Homans, and C. Bordier. 1990. Structure of the glycosyl-phosphatidylinositol membrane anchor of the *Leishmania major* promastigote surface protease. *J. Biol. Chem.* 265:16955-16964.
 26. Seki, T., N. Spurr, F. Obata, S. Goyert, P. Goodfellow, and J. Silver. 1985. The human Thy-1 gene: structural chromosomal location. *Proc. Natl. Acad. Sci. USA.* 82:6657-6661.
 27. Strickler, J. E., D. A. Binder, J. J. L'Italian, G. T. Shimamoto, S. W. Wait, L. J. Dalheim, J. Novotny, J. A. Radding, W. H. Konigsberg, M. Y. K. Armstrong, F. F. Richards, and T. M. Lalor. 1987. *Trypanosoma congolense*: structure and molecular organization of the surface glycoproteins of two early bloodstream variants. *Biochemistry.* 26:796-805.
 28. Taguchi, R., Y. Asahi, and H. Ikezawa. 1980. Purification and properties of phosphatidylinositol-specific phospholipase C of *Bacillus Thuringiensis*. *Biochim. Biophys. Acta.* 619:48-57.
 29. Tokumitsu, S. I., K. Kohnoe, M. Takeya, and T. Takeuchi. 1979. Localization of alkaline phosphatase isozymes in human cancer cell-lines *in vitro*. *Acta Histochem. Cytochem.* 12:631-635.
 30. Udenfriend, S., R. Micanovic, and K. Kodukula. 1991. Structural requirements of a nascent protein for processing to a PI-G anchored form: studies in intact cells and cell-free systems. *Cell Biol. Intl. Rep.* 15:739-759.
 31. von Heijne, G. 1983. Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* 133:17-21.
 32. von Heijne, G. 1990. The signal peptide. *J. Membr. Biol.* 115:195-201.