Semisynthesis of Functional Glycosylphosphatidylinositol-Anchored Proteins

Renée F. Roller*, Ankita Malik*, Maria A. Carillo, Monika Garg, Antonella Rella, Marie-Kristin Rauf, Bernd Lepenies, Peter H. Seeberger, and Daniel Varón Silva*

**Abstract:** Glypiation is a common posttranslational modification of eukaryotic proteins involving the attachment of a glycosylphosphatidylinositol (GPI) glycolipid. GPIs contain a conserved phosphoglycan that is modified in a cell- and tissue-specific manner. GPI complexity suggests roles in biological processes and effects on the attached protein, but the difficulties to get homogeneous material have hindered studies. We disclose a one-pot intein-mediated ligation (OPL) to obtain GPI-anchored proteins. The strategy enables the glypiation of folded and denatured proteins with a natural linkage to the glycolipid. Using the strategy, glypiated eGFP, Thy1, and the Plasmodium berghei protein MSP1

ubiquitous in eukaryotes and protozoa.[9] GPIs consist of a phospholipid, a phosphoethanolamine unit, and a conserved pseudo-pentasaccharide glycan core that is modified in a cell- and tissue-specific manner.[2] GPI-anchored (“glypiated”) proteins (GPI-APs) are functionally diverse, including hydrolytic enzymes, adhesion molecules, complementary regulatory proteins, receptors, and protozoan proteins, as well as cytokines and prion proteins.[1,3] The challenge of expressing GPI-APs containing a single, defined glycolipid in eukaryotic cell lines and their purification is a major limitation for determining the effects of GPIs and their modifications on the structure and function of the anchored protein.

Protozoan parasites cause severe human diseases such as leishmaniasis,[4] Chagas disease,[5] and malaria.[6] The cell membranes of these parasites contain large amounts of GPI-APs that modulate the host immune response during infection.[7] These proteins are of different size and complexity and include important vaccine candidates such as the surface variant glycoproteins from Trypanosoma brucei, mucins and mucin-like proteins from Trypanosoma cruzi, the metalloprotease GP63 from Leishmania donovani, and the circumsporozite protein (CSP) on sporozoites and merozoite surface proteins (MSPs) from Plasmodium falciparum.[8] among others. GPI-APs cover the parasite surface during all developmental stages and contribute to the survival of the parasites in the vector as well as to binding and invasion during the human liver and blood stages.[8b] The Plasmodium GPIs are very heterogeneous in the lipid part and contain an additional mannose (Man-IV) on the glycan core as the main modification (Figure 1a).[13]

Peptides and proteins have been attached mainly to GPI analogues by coupling fully protected peptides and glycans,[12] through Staudinger reaction and expressed-protein ligation,[13] and by transpeptidation with Sortase A.[14] However, access to GPI-APs containing a natural GPI limits the evaluation of the activity of these complex glycoconjugates. Herein, we describe a strategy to obtain GPI-anchored proteins containing fully lipidated GPIs. The strategy uses the *Nostoc punctiforme* DNAe split intein (Npu*Npu*) to generate protein thioester intermediates of the protein of interest that undergo a chemoselective reaction with synthetic cysteine-containing GPI-glycolipids in a one-pot ligation (OPL).[15] We established conditions for the reaction using enhanced green fluorescence protein (eGFP) as a protein model and show the semisynthesis of two naturally GPI-anchored proteins: Thy1 (CD90) from thymocytes and MSP1 from *Plasmodium berghei* ANKA (PbA). Moreover, we show that glypiation does not affect the structure of...
MSP1\textsubscript{19} in solution but enhances the in vitro production of the pro-inflammatory cytokines IL-12 and TNF-\(\alpha\) by bone-marrow-derived dendritic cells.

Results and Discussion

Three fragments were necessary to obtain GPI-APs by OPL using the Npu\textsuperscript{N}/Npu\textsuperscript{C} split intein fragments: a neX-pressed fusion protein of the protein of interest (POI) and the Npu\textsuperscript{N} fragment (POI-Npu\textsuperscript{N}), and Thy1 (Thy1-Npu\textsuperscript{N}) were recombinantly expressed in E. coli either as soluble proteins or in inclusion bodies and were isolated by affinity chromatography using a Histag attached at the C-terminus of Npu\textsuperscript{N} (for details see Supporting Information). The C-terminal intein fragment coupled to the first four residues of the C-Extein from DNA polymerase III from Nostoc punctiforme (Npu\textsuperscript{C}AA) was synthesized by the Fmoc strategy on Rink amide resin using microwave-assisted solid-phase peptide synthesis. The process provided the Npu\textsuperscript{C} peptide in good amount and purity. However, the formation of aspartimide and a difficult separation from the product reduced the yield of the synthesis (Figure S1 in the Supporting Information).

Two essential residues (Asn35 and Cys36) for the transfer and splicing in the natural intein were exchanged for Ala in this peptide, thereby de-function-lalizing the autocatalytic potential of the Npu\textsuperscript{C} fragment for the corresponding transfer of the protein fused to the Npu\textsuperscript{N} fragment.

The synthesis of GPI\textsubscript{1} and GPI\textsubscript{2}, which contain a cysteine residue, was completed using established methods\textsuperscript{[13b]} A convergent assembly of the glycan and two phosphorylations were used to install the lipid and a phosphoethanolamine...
coupled to cysteine. The synthesis of the GPI of *Plasmodium* proteins (GPI3, Figure 1b) started with the preparation of the pseudo-hexasaccharide 5 using two protected building blocks and a [4+2] glycosylation strategy (Scheme 1 and the Supporting Information).[16,17] Pseudo-hexasaccharide 5 was deallylated with palladium chloride in acetic acid.[18] A phosphorylation of myo-inositol with H-phosphonate 7[16] and subsequent oxidation with iodine delivered the corresponding glycolipid 8. Cleavage of the TIPS ether using scandium triflate and water to release the alcohol 9, subsequent phosphorylation with the H-phosphonate 10[13b] and careful oxidation with iodine and water gave the protected product 11. Final removal of all protecting groups from 11 by treatment with hydrazine acetate, palladium-catalyzed hydrogenolysis, and treatment with trifluoroacetic acid and with mercury trifluoroacetate delivered GPI3 for the glypiation of proteins (Scheme 1).[19]

The one-pot ligation (OPL) for C-terminal protein modification involved three steps. The first step is the self-assembly of the split-intein fragments POI-NpuN and NpuAA to give an active intein domain that in the second step induces an N- to S-acyl shift to form the thioester POI-NpuN intermediate (Figure 2a). Due to the lack of cysteine in NpuAA, the POI-S-NpuN intermediate cannot undergo an intramolecular transthioesterification with the C-extein, but it reacts in the third step with a thiol to form the desired ligation product by one of two pathways (Figure 2a). The first pathway (Path A) involves a transthioesterification of the intermediate POI-S-NpuN with a thiol additive, such as methyl 3-mercaptopropionate (MMP), that is added to the reaction to form a protein thioester (POI-SR) in situ. This thioester undergoes a second transthioesterification with a cysteine-containing molecule (GPI1, GPI2, and GPI3 or biotin-4) and a subsequent irreversible S- to N-acyl shift furnishes the desired product. In the second pathway, the product is formed through direct transthioesterification of POI-S-NpuN with the cysteine-containing molecule and subsequent rearrangement to the ligation product (Figure 2a). The OPL is technically an expressed-protein ligation reaction, or native chemical ligation,[20] that involves in situ formation of the protein thioesters mediated by an active split intein domain.

To establish the best conditions for the C-terminal modification of proteins, we investigated the OPL between eGFP-NpuN and biotin-4 using methyl 3-mercaptopropionate (MMP), 4-mercaptophenylacetic acid (MPAA), or sodium 2-mercaptopethansulfonate (MESNa) as a thiol reagent.[21] SDS-PAGE and western-blot analysis of the reaction showed formation of the eGFP-biotin product after one hour and increased formation of the product after 24 h when using MMP or MESNa (Figure 2b and Figure S6). A kinetic study of this OPL monitored by RP-HPLC showed a fast association of the intein fragments, the formation of intermediates to generate the active protein thioester, and the appearance of the ligation product in small quantities already after one minute. To achieve high protein conversion and high product yields, incubation over several hours or days was required (Figure 2c). These results were in good agreement with a fast association of the intein and formation of the eGFP-SR thioester and a slower ligation to the desired product. They showed also that a capture of the intein intermediate from the N- to S-acyl shift with cysteine-containing compound would deliver the product without the need for additional thiol reagents. However, since GPs are difficult to synthesize and are available only in limited amounts, we performed the OPL using the cysteine-GPIs as limiting reagents.

Glypiation of eGFP via the OPL strategy was investigated with GPI1, which contains a single lipid chain, using the conditions for the ligation with biotin-4. OPL with this glycolipid was extended to several days to increase conversion. The glypliated eGFP-GPI1 product was detected by western blot using an anti-GPI antibody. LC-MS analysis confirmed the formation of the product and the applicability of the strategy to glypliated eGFP (Figures S7 and S8). We next evaluated the OPL for glypiation of eGFP with GPI2, which bears a natural lipid with two alkyl chains shows low solubility in the ligation buffer. β-Octylglucopyranoside (22 mM) was added as a surfactant to the buffer to increase the solubility of GPI2 and the OPL was accomplished using the same reaction time and concentrations used with GPI1. SDS-PAGE analysis showed a slower conversion of the eGFP-SR thioester and lower yield formation of the eGFP-GPI2 product compared to the OPLs with biotin-4 and GPI1 (Figure 2d). This slower conversion may be a result of the low solubility of GPI2 and possible formation of micelles even in the presence of the surfactant.

Next, we performed the glypiation of the hydrophobic Thy1 (CD90) protein by OPL. The solution structure of GPI-
anchored glycoprotein Thy1 is affected by glypiation.\[22\] The GPI of this protein was the first fully elucidated GPI and has as sialic acid containing trisaccharide glycan side chain on Man-1.\[23\] In addition to the GPI anchor, Thy1 has three N-glycosylation sites and is naturally a mixture of more than 120 glycoforms. The biological function of Thy1 is not completely understood; however, it has been associated with T-cell proliferation and different types of cancer.\[24\] The influence of glycosylation and the GPI anchor on Thy1 activity remains unknown.\[22\]

Expression of the Thy1-Npu\( ^N \) fusion protein in E. coli yielded the protein in inclusion bodies requiring high concentrations of urea for solubilization (8 m) and purification (> 2 m). OPL with biotin-\( ^4 \) using the same conditions used for the ligations with eGFP, but including 2 m urea for keeping the Thy1-Npu\( ^N \) fusion protein in solution, allowed the intein fragments to fold, thereby inducing the formation of the activated intein domain and generation of the corresponding protein thioester. OPL with Thy1 proceeded similarly to that with eGFP; the Thy1-SR thioester formed fast, the ligation product Thy1-biotin appeared after a few minutes, and high conversion was achieved after several days (Figure S9).

Thy1 was ligated to GPI1 using the conditions established with biotin-\( ^4 \). After OPL for seven days, the Thy1-GPI1 product was purified. The unreacted fusion protein Thy1-Npu\( ^N \) and cleaved Npu\( ^N \) carrying a C-terminal HisTag were retained on the HisTrap column and were eluted with imidazole (Figure 2c). The product and the Npu\( ^N \)AA intein fragment lacking the tag were not retained in the HisTrap column and were separated either by dialysis or using SEC on a Superdex 30 column. The purification was monitored by SDS-PAGE (Figure 2c) and the product was analyzed by western blotting and LC–MS (Figure S10). The hydrophobic Thy1-GPI1 ligation product was obtained, but MS characterization proved challenging as various complex mixtures of adducts and fragmentation products were observed.

Encouraged by the results obtained coupling eGFP and Thy1, OPL glypiation of the C-terminal fragment of MSP1 (MSP1\( _{19} \)), one of the most abundant and highly conserved GPI-APs in Plasmodium sp. was undertaken.\[25\] MSP1 is a cysteine-rich protein stabilized by multiple disulfide bonds and is expressed as an approximately 200 kDa protein precursor. The protein is linked to the GPI via cysteine and it is an essential protein for the invasion and survival of the parasite in the host.\[26\] MSP1 undergoes two proteolytic cleavages leaving the C-terminal fragment (MSP1\( _{19} \)) anchored to the membrane via the GPI.\[27\] MSP1\( _{19} \) remains on the membrane during the red blood cell infection and is internalized by endocytosis during the ring-stage into the vacuolar system, where it localizes at the food vacuole.\[28\] a unique organelle in Plasmodium for degradation of hemoglobin and target for multiple antimalarial drugs.\[29\]

OPL between MSP1\( _{19} \)-Npu\( ^N \) and biotin-\( ^4 \) using MMBA as thiol reagent resulted in fast formation of the MSP1\( _{19} \)-biotin product as determined by western blotting and LC–MS. SDS-PAGE monitoring of the process showed almost complete consumption of the fusion protein after four days (Figure S11). MSP1\( _{19} \)-Npu\( ^N \) was ligated next to the synthetic GPI3 using OPL using the conditions established to obtain eGFP-GPI3 (Figure 2a and the Supporting Information).\[29\] The ligation progressed in a similar manner to the ligation with biotin-\( ^4 \) and delivered the product MSP1\( _{19} \)-GPI3 with high conversion of the fusion protein after five days (Figure S12).

OPL was scaled up for eGFP-Npu\( ^N \) and MSP1\( _{19} \)-Npu\( ^N \) with biotin-\( ^4 \), GPI1, and GPI3 to obtain eGFP-biotin, eGFP-GPI1, MSP1\( _{19} \)-biotin, and MSP1\( _{19} \)-GPI3 to investigate the effects of glypiation on the structure and biological activity of the proteins. These ligations were carried out under optimized conditions using 10 \( \mu \)M MSP1\( _{19} \)-Npu\( ^N \), 30 \( \mu \)M Npu\( ^N \)AA, 20 mM TCEP solution, and 70 mM MMBa in Tris buffer at pH 7.2 containing octylglucoside (22 mM) for ligation with GPI3. The MSP1\( _{19} \)-biotin and MSP1\( _{19} \)-GPI3 products were isolated using

![Image](image-url)

**Figure 3.** Characterization of GPI-anchored proteins. a) RP-HPLC chromatogram of purified MSP1\( _{19} \)-GPI3. b) Deconvoluted ESI-MS spectrum of MSP1\( _{19} \)-GPI3. c) SDS-PAGE after His-Trap purification of the preparative synthesis of MSP1\( _{19} \)-Biotin and MSP1\( _{19} \)-GPI3. d) CD spectra of MSP1\( _{19} \)-Biotin and MSP1\( _{19} \)-GPI3. e, f) Production of TNF-\( \alpha \) (e) and IL-12 (f) by BMDCs stimulated with the MSP1\( _{19} \)-Biotin and MSP1\( _{19} \)-GPI3. CM = Culture medium.
HisTrap and SEC columns and were analyzed by SDS-PAGE, MALDI MS, and LC–MS (Figure 3).

Possible structural changes on the proteins during OPL or by the glypiation were analyzed by circular dichroism (CD). Previous CD studies showed different profiles with and without glypiation that were attributed to a structural change in the protein.[22,31] The CD spectra of eGFP-Biotin, eGFP-GPI1, and eGFP-OH exhibit a minimum at 210 nm that corresponds with the β-sheet folds in the β-barrel structure reported for this protein (Figures S15). Similarly, the CD spectra of MSP1<sub>α</sub>-biotin and MSP1<sub>α</sub>-GPI3 showed a profile corresponding to a random coil structure that is in good agreement with previously reported CD spectra for unmodified MSP1<sub>α</sub> containing two epidermal growth factor (EGF) like domains (Figure 3d). CD analysis indicates that OPL and glypiation did not affect the structure of eGFP and MSP1<sub>α</sub> in solution, which is a requirement for the immunogenicity of this protein.[34]

Isolated GPIs and GPI-anchored proteins from parasites induce a strong proinflammatory response and the production of cytokines by macrophages.[35] This type of response is characteristic in patients with severe malaria that show enhanced levels of IL-6, TNF-α, and IL-12 correlating with parasitemia.[36] Thus, a proinflammatory response is considered important for protection against malaria and for clearance of the parasites, but may also contribute to malaria-associated pathology such as cerebral malaria.[37]

The proinflammatory activity of MSP1<sub>α</sub>-biotin and MSP1<sub>α</sub>-GPI3 was tested in vitro and compared to bacterial lipopolysaccharide (LPS), a strong proinflammatory stimulator. Bone-marrow-derived dendritic cells (BMDCs) from mice were cultured in the presence of Iseove’s modified Dulbecco’s medium (CM, negative control), 1 μg/mL lipopolysaccharide (LPS, positive control), or each of the modified MSP1<sub>α</sub> proteins at 0.01, 0.1, and 1 μM concentrations (Figures 3e and f). ELISA determination of tumor necrosis factor alpha (TNF-α) and interleukin 12 (IL-12) indicated cytokine production for cells stimulated with MSP1<sub>α</sub>-biotin and MSP1<sub>α</sub>-GPI3. The cytokine production was concentration-dependent with highest cytokine secretion for MSP1<sub>α</sub>-GPI3 at 1 μM, and showed a tendency for higher TNF-α and IL-12 levels for the glypiated protein. These findings agree with previous reports describing this property for isolated GPIs from protozoa.[7c,38] Future studies will focus on the effect of glypiation for MSP1<sub>α</sub> activity in vivo and its potential as a malaria vaccine candidate.

The difficulty in accessing GPI-anchored proteins has limited investigations into the biological role of glypiation and its effect on the structure and activity of proteins. The herein reported OPL is an efficient method to glypitate the C-terminus of proteins and provides access to naturally glypiated proteins with biological activity. This OPL method will help us to evaluate the effect of the GPI structure on protein folding, the behaviour of GPI-APs in membranes, and the activity of glypiated proteins in vitro and in vivo.

**Conclusion**

We have established a one-pot ligation strategy for the glypiation of proteins using the DnaE split intein from *Nostoc punctiforme*. An active protein thioester is formed in situ and reacts with a synthetic cysteine-containing GPI. The method tolerates folded and denatured proteins, glycolipids with poor solubility in buffers, and additives for protein solubilization. We completed the semisynthesis of glypiated eGFP, Thy1, and *Plasmodium berghei* ANKA MSP1<sub>α</sub> proteins and showed that the process does not affect protein folding. Structural analysis of the glypiated proteins by CD revealed no effect of glypiation on the overall structure of the eGFP and MSP1<sub>α</sub> proteins in solution. Glypiation enhanced the proinflammatory activity of the MSP1<sub>α</sub> protein in vitro, thus suggesting potential application of glypiated proteins in vaccine development.

**Acknowledgements**

We thank DFG and the priority program 1623, the Max Planck Society and the RIKEN-Max Planck Joint Center for Systems Chemical Biology for financial support. We thank Prof. Dr. Henning Mootz for the donation of the eGFP-Npu<sub>N</sub> clone and for his helpful suggestions.

**Conflict of interest**

The authors declare no conflict of interest.

**Keywords:** GPI anchor · glycoproteins · glypiation · protein modifications · protein semisynthesis

---

[1] M. A. Ferguson, J. Cell. Sci. 1999, 112, 2799–2809.
[2] Y. H. Tsai, X. Y. Liu, P. H. Seeberger, Angew. Chem. Int. Ed. 2012, 51, 11438–11456; Angew. Chem. 2012, 124, 11604–11623.
[3] a) M. A. J. Ferguson, A. F. Williams, Annu. Rev. Biochem. 1988, 57, 285–320; b) H. Ikezawa, Biol. Pharm. Bull. 2002, 25, 409–417.
[4] S. Burza, S. L. Croft, M. Boelaert, Lancet 2018, 392, 951–970.
[5] C. Bern, N. Engl. J. Med. 2015, 373, 453–466.
[6] a) F. Debierre-Grockiego, R. T. Schwarz, Glycobiochemistry 2010, 20, 801–811; b) E. A. Ashley, A. Pye Phyo, C. J. Woodrow, Lancet 2018, 391, 1608–1621.
[7] a) M. A. S. Campos, I. C. Almeida, O. Takeuchi, S. Akira, E. P. Valente, D. O. Procópio, L. R. Travassos, J. A. Smith, D. T. Golenbock, R. T. Gazzinelli, J. Immunol. 2001, 167, 416–423; b) S. D. Tachado, R. Mazhari-Tabrizi, L. Schofield, Parasite Immunol. 1999, 21, 609–617; c) M. M. Camargo, I. C. Almeida, M. E. Pereira, M. A. Ferguson, L. R. Travassos, R. T. Gazzinelli, J. Immunol. 1997, 158, 5890–5901.
[8] a) P. R. Gilson, B. S. Crabb, Int. J. Parasitol. 2009, 39, 91–96; b) P. R. Sanders, P. R. Gilson, G. T. Cantin, D. C. Greenbaum, T. Nebl, D. J. Carucci, M. J. McConville, L. Schofield, A. N. Hodder, J. R. Yates, B. S. Crabb, J. Biol. Chem. 2005, 280, 40169–40176.
[9] A. Guha-Niyogi, D. R. Sullivan, S. J. Turco, Glycobiochemistry 2001, 11, 45R–59R.
