Lipid Modification of Proteins through Sortase-Catalyzed Transpeptidation

John M. Antos, Gwenn M. Miller, Gijsbert M. Grotenbreg, and Hidde L. Ploegh*

Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142

Received August 27, 2008; E-mail: ploegh@wi.mit.edu

Abstract: A general chemoenzymatic method for the site-specific attachment of lipids to protein substrates is described. Sortase A is used to append short lipid-modified oligoglycine peptides to the C terminus of protein substrates bearing a five amino acid sortase A recognition sequence (LPETG). We demonstrate the attachment of a range of hydrophobic modifications in excellent yield (60−90%), including a simple step for removing the sortase enzyme postreaction. Lipoproteins prepared using these procedures were subsequently shown to associate with mammalian cells in a lipid tail-dependent fashion and localized to the plasma membrane and endosomes.

Introduction

Methods for the construction of protein−lipid conjugates are valuable tools for studying the role of post-translational lipid modification in controlling protein function and localization. Lipid attachment also provides a means for manipulating the properties of peptides and proteins that normally do not possess these types of modifications. Highlighting the utility of synthetic lipoproteins, Bertozzi and co-workers recently reported the synthesis of GFP-containing GPI-anchor mimetics as a means for probing the role of individual sugar residues in this complex membrane anchor.[1,2] Numerous reports have also shown that the addition of lipids and fatty acids can endow proteins and peptides with cell-penetrating capabilities,[3−7] modulate immunogenicity,[8,9] and provide anchors for incorporation into liposomes.[10,11]

With respect to full-size protein substrates, the attachment of lipids in site-specific fashion remains a significant bioconjugation challenge. Only a relatively limited set of chemical and chemoenzymatic strategies are available for this purpose. These include the use of enzymes such as myristoyl[12,13] and prenyl[14−16] transferases, wherein the user is limited to the natural substrates of these enzymes or closely related analogues. In the case of chemical labeling, a common strategy is cysteine derivatization through alkylation[17,18] or disulfide formation,[19,20] an approach that requires site-directed mutagenesis to ensure that protein substrates contain only a single reactive cysteine residue. Among the available strategies for lipid attachment, expressed protein ligation (EPL)[1,2,21−25] has proven to be the most versatile. For example, this method has been used to prepare GFP bearing a range of lipid modifications[1,2,24] and to generate lipidated Ras[25] and Rab[22] proteins. However, EPL mandates the expression of protein substrates as fusions with large intein domains, a necessity that can reduce protein expression yields and may require careful selection of the intein

(1) Paulick, M. G.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20332–20337.
(2) Paulick, M. G.; Wise, A. R.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2007**, *129*, 11543–11550.
(3) Martin, S. E.; Peterson, B. R. *Bioconjugate Chem.* **2003**, *14*, 67–74.
(4) Boonyarattanakalin, S.; Athavankar, S.; Sun, Q.; Peterson, B. R. *J. Am. Chem. Soc.* **2006**, *128*, 4917–4917.
(5) Covic, L.; Gresser, A. L.; Talavera, J.; Swift, S.; Kuliopulos, A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 643–648.
(6) Sun, Q.; Cai, S.; Peterson, B. R. *J. Am. Chem. Soc.* **2008**, *130*, 10064–10065.
(7) Nelson, A. R.; Borland, L.; Allbritton, N. L.; Sims, C. E. *Biochemistry* **2007**, *46*, 14771–14781.
(8) Moyle, P. M.; Toth, I. *Curr. Med. Chem.* **2008**, *15*, 506–516.
(9) BenMohamed, L.; Wechsler, S. L.; Nesburn, A. B. *Lancet Infect. Dis.* **2002**, *2*, 425–431.
(10) Reulen, S. W.; Brusselaars, W. W.; Langereis, S.; Mulder, W. J.; Breurken, M.; Merkx, M. *Bioconjugate Chem.* **2007**, *18*, 590–596.
(11) Torchilin, V. P. *Nat. Rev. Drug Discovery* **2005**, *4*, 145–160.
(12) Ha, V. L.; Thomas, G. M.; Stauffer, S.; Randazzo, P. A. *Methods Enzymol.* **2005**, *404*, 164–174.
(13) Heal, W. P.; Wickramasinghe, S. R.; Bowyer, P. W.; Holder, A. A.; Smith, D. F.; Leatherbarrow, R. J.; Tate, E. W. *Chem. Commun.* **2008**, 480–482.
(14) Kalinin, A.; Thoma, N. H.; Iakovenko, A.; Heinemann, I.; Rostkova, E.; Constantinescu, A. T.; Alexandrov, K. *Protein Expression Purif.* **2001**, *22*, 84–91.
(15) Gauchet, C.; Labadie, G. R.; Poulter, C. D. *J. Am. Chem. Soc.* **2006**, *128*, 9274–9275.
(16) Duckworth, B. P.; Zhang, Z. Y.; Hosokawa, A.; Distefano, M. D. *ChemBioChem* **2007**, *8*, 98–105.
(17) Breydo, L.; Sun, Y.; Makarava, N.; Lee, C. I.; Novitskaia, V.; Bocharova, O.; Kao, J. P.; Baskakov, I. V. *Biochemistry* **2007**, *46*, 852–861.
(18) Bader, B.; Kuhn, K.; Owen, D. J.; Waldmann, H.; Wittinghofer, A.; Kuhlmann, J. *Nature* **2000**, *403*, 223–226.
(19) Eberl, H.; Tittmann, P.; Glockshuber, R. *J. Biol. Chem.* **2004**, *279*, 25058–25065.
(20) Hicks, M. R.; Gill, A. C.; Bath, I. K.; Rullay, A. K.; Sylvester, I. D.; Crout, D. H.; Pinheiro, T. J. *FEBS J.* **2006**, *273*, 1285–1299.
(21) Muir, T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249–289.
(22) Alexandrov, K.; Heinemann, I.; Durek, T.; Sidorovitch, V.; Goody, R. S.; Waldmann, H. *J. Am. Chem. Soc.* **2002**, *124*, 5648–5649.
(23) Pellois, J. P.; Muir, T. W. *Angew. Chem., Int. Ed.* **2005**, *44*, 5713–5717.
(24) Grogan, M. J.; Kaizuka, Y.; Conrad, R. M.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2005**, *127*, 14383–14387.
(25) Gottlieb, D.; Grunwald, C.; Nowak, C.; Kuhlmann, J.; Waldmann, H. *Chem. Commun.* **2006**, 260–262.

Scheme 1. Site-Specific Lipid Attachment through Sortase-Mediated Transpeptidation
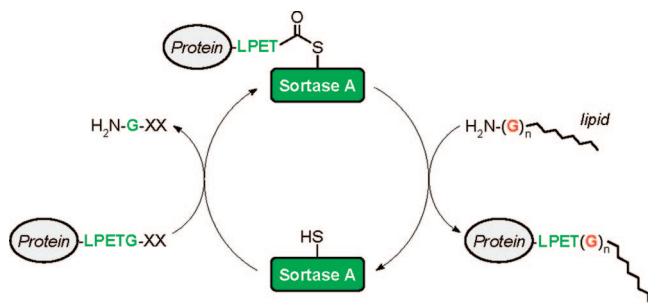
domain to prevent premature intein cleavage and a corresponding drop in the yield of ligation products.[26–29]

With the goal of complementing existing strategies for lipoprotein synthesis, we have developed a general method for the site-specific installation of lipids and fatty acids using sortase-catalyzed transpeptidation (Scheme 1). This chemoenzymatic approach has seen increasing use as a method for protein engineering.[30–38] Sortase A from *Staphylococcus aureus* recognizes a five amino acid sequence (LPXTG) near the C-terminus of its natural protein targets.[39–41] The LPXTG motif has been successfully transposed onto unstructured regions of other proteins to generate new sortase substrates. The active site cysteine of sortase cleaves between the threonine and glycine residues of the recognition site to produce a thioester intermediate which is then attacked by the amino terminus of an oligoglycine nucleophile, resulting in the formation of a new amide bond. The method is extremely versatile, stemming from the remarkable tolerance of the enzyme for substituents C-terminal to the oligoglycine unit. Synthetic nucleophiles, containing 1–5 glycine residues, have been decorated with a range of substituents including fluorophores,[33,34] photoaffinity probes,[34] peptide nucleic acids,[36] polymers,[35] solid supports,[32,35,42] or other polypeptides,[33,38] allowing the site-specific ligation of

these moieties to peptide and protein substrates. Here we describe the synthesis of lipid-modified oligoglycine nucleophiles compatible with the sortase-mediated transpeptidation reaction and demonstrate their high-yielding ligation to a protein substrate. Lipoproteins prepared using this procedure were shown to strongly associate with mammalian cells in a lipid tail-dependent fashion and were found to localize to the plasma membrane and endosomes.

Results and Discussion

Design and Synthesis of Lipid-Modified Triglycine Nucleophiles. Nucleophiles compatible with sortase-mediated transpeptidation have the simple structural requirement of a stretch of glycine residues with a free amino terminus. While successful transpeptidation has been reported with nucleophiles containing anywhere from one to five glycines, maximum reaction rates are obtained when two or more glycines are present.[32,38] It should be noted that primary alkylamines have been shown to participate in the transpeptidation reaction,[31,35] though in general these appear to be less efficient than oligoglycine derivatives. With these considerations in mind, we synthesized a panel of lipid-modified triglycine nucleophiles using solid-phase synthesis (Scheme 2).

Resin-bound precursor **1** was prepared on Rink amide resin using standard Fmoc chemistry. An orthogonally protected lysine residue was incorporated to allow selective liberation and subsequent acylation of the ε-amino group. Following removal of the 4-methyltrityl (Mtt) group, the lysine side chain was acylated with a range of hydrophobic carboxylic acids, including naturally occurring post-translational modifications such as myristic (**1**-C14) and palmitic (**1**-C16) acid, as well as modifications not found in eukaryotic cells (**1**-ad). The couplings of C18, C20, C22, and C24 acids were performed at 50 °C to maintain the solubility of all the reaction components. Following Fmoc deprotection and resin cleavage, the identity of nucleophiles derived from **1** was confirmed by LC–ESI-MS (Figure S1, Supporting Information). All peptides were obtained with good purity and did not require extensive purification prior to transpeptidation.

In addition to purely aliphatic substituents, we also prepared a cholesterol-modified triglycine nucleophile (**2**-chol, Scheme 2). Peptide **2** was prepared on Rink amide resin and then cleaved from the solid support, leaving the N-terminal Fmoc group intact. A tryptophan residue was included in this peptide as a UV-active chromophore to aid in purification by HPLC. Peptide **2** was then reacted with cholesterol NHS-carbonate **3**, followed by the addition of piperidine to remove the Fmoc group. The crude material was then purified by RP-HPLC to yield **2**-chol in 22% yield.

Sortase-Mediated Ligation of Triglycine Nucleophiles to a Protein Substrate. With a series of lipidated nucleophiles in hand, we proceeded to explore their coupling to a protein substrate using sortase. A model protein consisting of eGFP equipped with a C-terminal LPETG sequence followed by a His$_6$ affinity handle was expressed and purified from *Escherichia coli*. Recombinant sortase A containing an N-terminal His$_6$ tag was produced as described.[40] At the outset of these studies, we decided to incorporate a procedure for removing the sortase enzyme after the reaction, similar to a purification scheme proposed by Parthasarathy et al.[35] The configuration of His$_6$ tags on both sortase and the eGFP substrate provided a simple means for doing so, where Ni-NTA resin was added following transpeptidation to retrieve the sortase enzyme as well as any

(26) Luka, Z.; Wagner, C. *Protein Expression Purif.* **2003**, *28*, 280–286.
(27) Zhao, W.; Zhang, Y.; Cui, C.; Li, Q.; Wang, J. *Protein Sci.* **2008**, *17*, 736–747.
(28) Cui, C.; Zhao, W.; Chen, J.; Wang, J.; Li, Q. *Protein Expression Purif.* **2006**, *50*, 74–81.
(29) Tan, L. P.; Lue, R. Y.; Chen, G. Y.; Yao, S. Q. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 6067–6070.
(30) Pritz, S.; Kraetke, O.; Klose, A.; Klose, J.; Rothemund, S.; Fechner, K.; Bienert, M.; Beyermann, M. *Angew. Chem., Int. Ed.* **2008**, *47*, 3642–3645.
(31) Samantaray, S.; Marathe, U.; Dasgupta, S.; Nandicoori, V. K.; Roy, R. P. *J. Am. Chem. Soc.* **2008**, *130*, 2132–2133.
(32) Chan, L.; Cross, H. F.; She, J. K.; Cavalli, G.; Martins, H. F. P.; Neylon, C. *PLoS ONE* **2007**, *2*, e1164.
(33) Tanaka, T.; Yamamoto, T.; Tsukiji, S.; Nagamune, T. *ChemBioChem* **2008**, *9*, 802–807.
(34) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707–708.
(35) Parthasarathy, R.; Subramanian, S.; Boder, E. T. *Bioconjugate Chem.* **2007**, *18*, 469–476.
(36) Pritz, S.; Wolf, Y.; Kraetke, O.; Klose, J.; Bienert, M.; Beyermann, M. *J. Org. Chem.* **2007**, *72*, 3909–3912.
(37) Mao, H. *Protein Expression Purif.* **2004**, *37*, 253–263.
(38) Mao, H.; Hart, S. A.; Schink, A.; Pollok, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 2670–2671.
(39) Kruger, R. G.; Otvos, B.; Frankel, B. A.; Bentley, M.; Dostal, P.; McCafferty, D. G. *Biochemistry* **2004**, *43*, 1541–1551.
(40) Ton-That, H.; Liu, G.; Mazmanian, S. K.; Faull, K. F.; Schneewind, O. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 12424–12429.
(41) Ton-That, H.; Mazmanian, S. K.; Faull, K. F.; Schneewind, O. *J. Biol. Chem.* **2000**, *275*, 9876–9881.
(42) Clow, F.; Fraser, J. D.; Proft, T. *Biotechnol. Lett.* **2008**, *30*, 1603–1607.

Scheme 2. Synthesis of Lipid-Modified Triglycine Nucleophiles
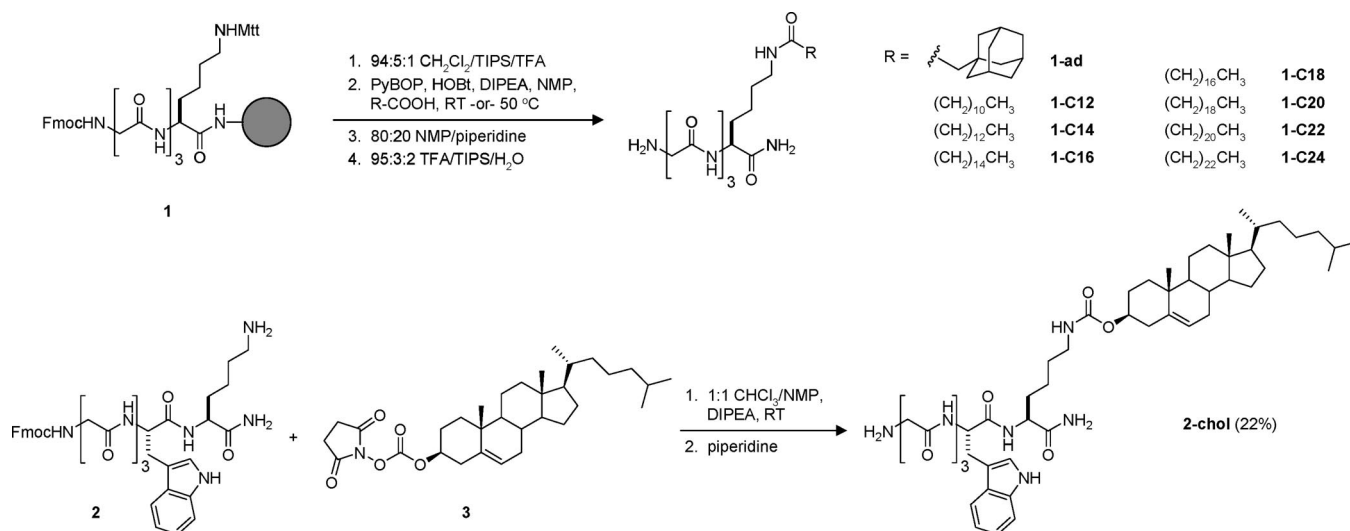
unreacted eGFP-LPETG-His$_6$. In this scenario, the transpeptidation product should no longer be His$_6$-tagged because residues C-terminal to the LPETG motif are lost in the course of transpeptidation, and thus, this material will not be bound by Ni-NTA. As shown in Figure 1a, exposing eGFP-LPETG-His$_6$ to 150 μM sortase and 2 mM triglycine (GGG) in Tris buffer (pH 7.5) containing 10 mM CaCl$_2$ resulted in the appearance of a lower molecular weight species just below the band corresponding to sortase (lane 2). Excess sortase was used to drive the reaction to completion in a reasonable time frame. Near-quantitative conversion of the input eGFP-LPETG-His$_6$ protein (lane 1) was observed over the course of a 3 h incubation at 37 °C. A slurry of Ni-NTA resin containing 1 M NaCl and 40 mM imidazole was then added, and the mixture was incubated for 2 h. The addition of imidazole and NaCl at the specified concentrations was critical to preventing nonspecific binding of the transpeptidation products. After filtration to remove Ni-NTA, the filtrate was analyzed by SDS−PAGE. A single polypeptide was observed with no evidence for the presence of residual sortase (lane 3). Characterization by ESI-MS confirmed that the protein observed in lane 3 was indeed the desired ligation product (Figure S2, Supporting Information). Also, judging from the intensity of the band in lane 3 as compared to the corresponding eGFP bands in lanes 1 and 2, we were confident that very little material was lost over the course of the two-step transpeptidation/sortase-depletion protocol.

Initial attempts to ligate one of the lipid-modified nucleophiles (**1**-C22) using the conditions described above yielded only small amounts of the putative transpeptidation product as well as an unidentified higher molecular weight species[43] (Figure 1a, lane 4). Not surprisingly, the solubility of **1**-C22 in aqueous buffer was extremely poor, and we suspected that this may limit the amount of nucleophile available in solution. To improve solubility, we examined the effect of adding mild detergents to the reaction mixture. As shown in Figure 1a (lane 8), the addition of 1% (w/v) *n*-dodecyl maltoside afforded a significant

increase in the level of transpeptidation product relative to that of the other detergents tested. This material was further
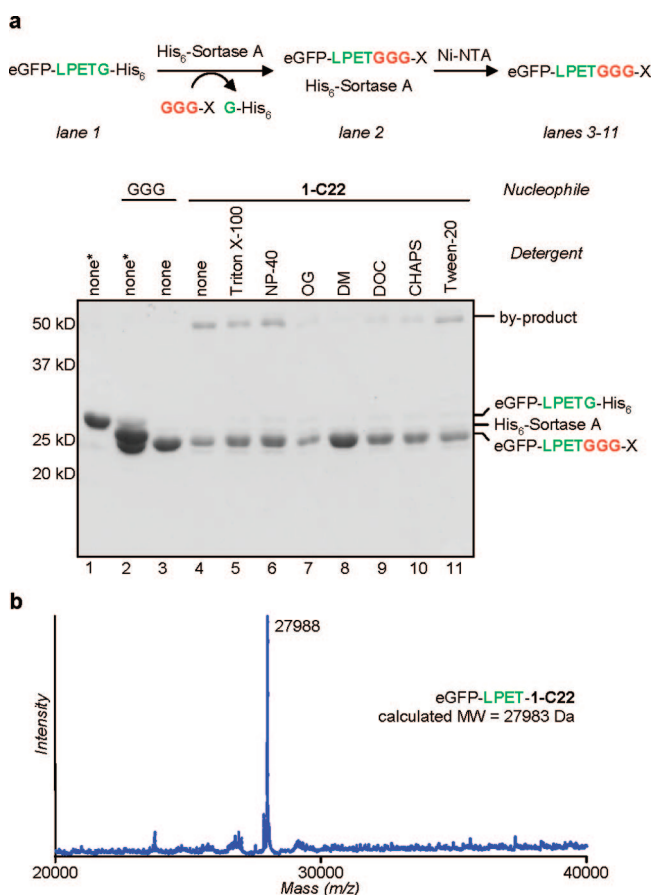

Figure 1. (a) Optimization of a two-step protocol for lipid ligation followed by removal of the sortase enzyme. Conditions: 77 μM eGFP-LPETG-His$_6$, 150 μM sortase A, 2 mM nucleophile, 1% (w/v) detergent, 50 mM Tris, pH 7.5, 150 mM NaCl, 10 mM CaCl$_2$, 3 h at 37 °C. After transpeptidation, Ni-NTA resin was added as a slurry in 1 M NaCl and 40 mM imidazole and incubated for 2 h at room temperature. Abbreviations: OG = *n*-octyl glucoside, DM = *n*-dodecyl maltoside, DOC = deoxycholate. The asterisks indicate samples not treated with Ni-NTA resin. (b) ESI-MS spectrum of eGFP-LPET−**1**-C22 ligation product.

(43) The high molecular weight byproduct may correspond to the acyl enyme intermediate formed between sortase A and the eGFP substrate. This assignment implies that the affinity of the His$_6$ tag on the N-terminus of sortase is impaired in the acyl enzyme intermediate because this species is not removed during depletion of His$_6$-tagged proteins using Ni-NTA resin. Further characterization of this byproduct was not performed.
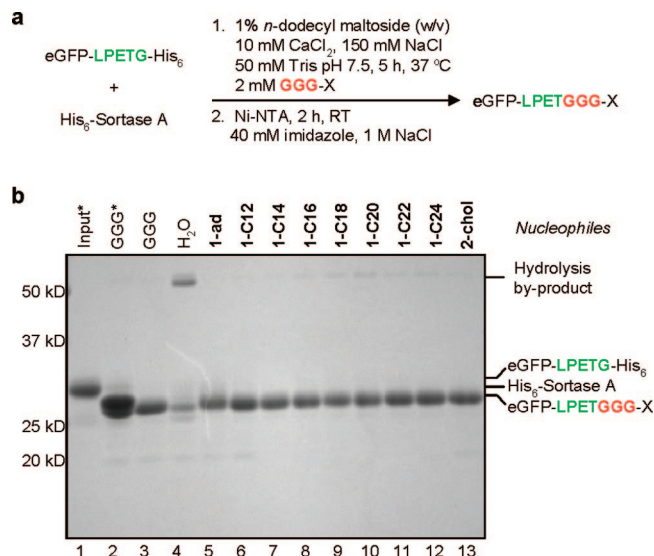
Figure 2. Lipid attachment through sortase-catalyzed transpeptidation: (a) optimized procedure for lipid coupling followed by removal of His_6-tagged proteins with Ni-NTA resin, (b) SDS−PAGE analysis of lipid-modified eGFP following transpeptidation and depletion of His_6-tagged proteins ("Input" = eGFP-LPETG-His_6 in the absence of sortase and nucleophile; the asterisks indicate samples incubated with Ni-NTA resin and 300 mM imidazole to block binding of His_6-tagged proteins).

Table 1. Mass Spectral Characterization and Yield (%) of eGFP−Lipid Conjugates

eGFP conjugate	calcd mass (Da)	obsd mass (Da)	yield (%)
eGFP−GGG	27534	27531	93
eGFP−**1**-ad	27837	27835	87
eGFP−**1**-C12	27843	27842	87
eGFP−**1**-C14	27871	27870	72
eGFP−**1**-C16	27899	27898	64
eGFP−**1**-C18	27927	27926	57
eGFP−**1**-C20	27955	27954	66
eGFP−**1**-C22	27983	27982	83
eGFP−**1**-C24	28011	28010	84
eGFP−**2**-chol	28260	28257	87

characterized by ESI-MS and found to consist exclusively of the desired lipid-modified eGFP derivative (Figure 1b).

Using 1% (w/v) *n*-dodecyl maltoside to maintain nucleophile solubility, we were able to successfully ligate the entire panel of lipid-modified triglycine nucleophiles to our eGFP model substrate (Figure 2). In all cases, a single major protein that migrated at a lower apparent molecular weight than unreacted eGFP-LPETG-His_6 (lane 1) was observed in the filtrate of reactions using lipidated nucleophiles (lanes 5−13). As expected, these samples also showed complete removal of the sortase enzyme. Coupling reactions using only triglycine (lanes 2 and 3) served as positive controls for transpeptidation. A reaction lacking an oligoglycine nucleophile resulted in slow hydrolysis of the acyl enzyme intermediate and the formation of a higher molecular weight byproduct[43] (lane 4). As an additional negative control, we confirmed that the formation of transpeptidation products was dependent on the presence of sortase (Figure S3, Supporting Information). The identity of all eGFP conjugates was ultimately established by mass spectrometry (Table 1 and Figure S2, Supporting Information).

The eGFP chromophore provided a convenient method for quantifying the yield of the eGFP conjugates. By comparing the absorbance at 488 nm of the conjugates relative to the input eGFP-LPETG-His_6 protein, we estimated that the lipid-modified
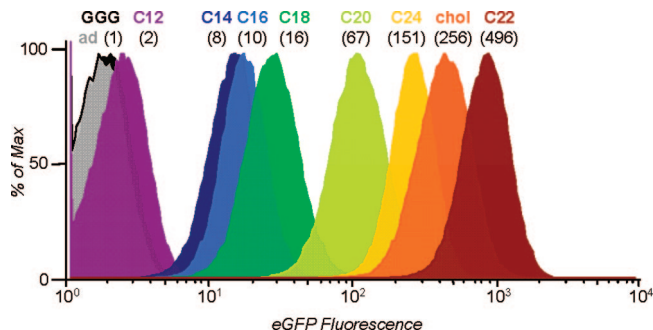
Figure 3. Association of lipid-modified eGFP with HeLa cells. Cells were incubated with lipid-modified eGFP (2.5 µg/mL) for 1 h in serum-free medium and then analyzed by flow cytometry. Numbers in parentheses represent the fold enhancement in mean cellular fluorescence relative to that of eGFP−GGG. Histograms for eGFP−GGG (black) and eGFP−**1**-ad (gray) are overlapping.

protein conjugates were all obtained in excellent (∼60−90%) yield (Table 1 and Figure S4, Supporting Information). A reproducible drop in reaction yield for nucleophiles containing fatty acids of intermediate length (**1**-C16, **1**-C18, **1**-C20) was observed. Whether this reflects a drop in the efficiency of the actual transpeptidation reaction or limited solubility of the resulting lipoprotein conjugates, which could contribute to losses during the sortase depletion step, is presently unclear. It should be noted, however, that the purity of these eGFP conjugates, as determined by both SDS−PAGE (Figure 2) and ESI-MS (Figure S2, Supporting Information), was not compromised.

Interaction of Lipid-Modified eGFP with Mammalian Cells. The presence of a hydrophobic tail on eGFP should render it capable of spontaneous association with biological membranes.[1,2,24,44] As an initial investigation into the effect of the identity of the lipid tail on the ability of these lipoproteins to associate with cells, we incubated HeLa cells with lipid-modified eGFP (2.5 µg/mL) in serum-free medium for 1 h at 37 °C. After washing, the cells were analyzed by flow cytometry (Figure 3). In general, we observed a steady increase in cellular fluorescence with increasing length of the lipid tail. In the case of eGFP−**1**-C22, we recorded a 500-fold enhancement in mean cellular fluorescence relative to that of eGFP−GGG. Interestingly, eGFP−**1**-C24 consistently yielded lower association than eGFP−**1**-C22. For all samples cell viability was >94% as determined by exclusion of propidium iodide, indicative of a lack of cytotoxicity of the lipidated eGFP derivatives. Also, because the eGFP preparations contained low levels of free lipidated nucleophile (∼1 µM in culture medium), we verified that this contaminant did not contribute to the observed increase in cellular fluorescence. eGFP−GGG in combination with 1 µM **1**-C22, **1**-C24, or **2**-chol produced only modest increases in fluorescence, <5% of the magnitude observed for the corresponding covalently modified eGFP conjugates (Figure S5, Supporting Information).

A preliminary evaluation of the subcellular localization of eGFP−**1**-C22 and eGFP−**2**-chol was performed using spinning disk confocal microscopy. HeLa cells and U373 cells were incubated with 2.5 µg/mL lipoprotein in serum-free medium, followed by washing and live cell imaging. For both cell types we observed staining of the cell surface as well as the appearance of internal fluorescent structures (Figure 4). For all cells, filamentous membrane extensions, presumably filopodia,

---

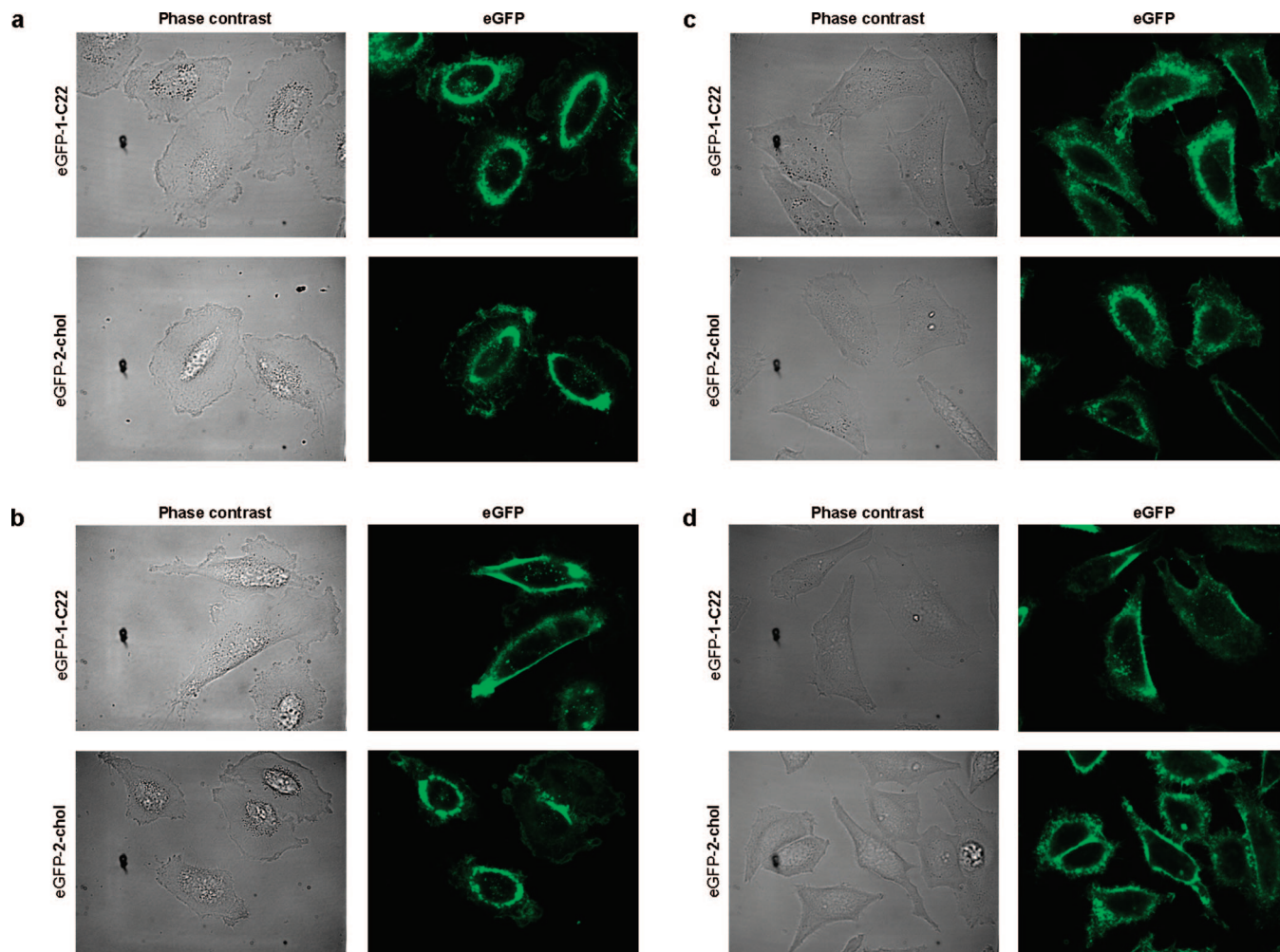(44) Zacharias, D. A.; Violin, J. D.; Newton, A. C.; Tsien, R. Y. *Science* **2002**, *296*, 913−916.

Figure 4. Subcellular distribution of eGFP−**1**-C22 and eGFP−**2**-chol in U373 and HeLa cells. Cells were incubated with lipid-modified eGFP (2.5 μg/mL) at 37 °C in serum-free medium and then imaged live using spinning disk confocal microscopy. Key: (a) U373 cells, 1 h of incubation, (b) U373 cells, 5 h of incubation, (c) HeLa cells, 1.25 h of incubation, (d) HeLa cells, 5 h of incubation.

were evident in the fluorescent micrographs, indicating the presence of the eGFP lipoproteins on the plasma membrane. In U373 cells, distinct vesicular structures were observed at early (1 h) and later (5 h) time points, with a more pronounced presence of vesicles observed for eGFP−**1**-C22 than for eGFP−**2**-chol. In general, less internalization was observed for HeLa cells, though punctate internal staining could be observed with eGFP−**1**-C22 after a 5 h incubation in these cells. In the case of U373, we were also able to verify that eGFP−**1**-C22 had access to early/recycling endosomal compartments as determined by partial colocalization with transferrin (Figure 5).[6,45] Colocalization between eGFP−**2**-chol and transferrin was not observed (Figure 5) for this cell line. Taken together, these results demonstrate that synthetic lipoproteins prepared using sortase-mediated transpeptidation can interact strongly with the plasma membrane and access endosomal compartments. The range of staining observed for our eGFP conjugates suggests that interactions with cells could be fine-tuned, and studies are under way to characterize the differences in the behavior of protein conjugates bearing different lipid anchors, specifically looking for differences in subcellular distribution and routes of endocytosis.

Conclusion

In summary, we have developed a general strategy using sortase-mediated transpeptidation as a means to install hydro-

phobic lipid modifications onto protein substrates in site-specific fashion. The ease of use of this method stands out as a significant asset. Protein substrates require only a five amino acid extension (LPETG) to serve as sortase substrates, a modest insertion that is not expected to impede the function of most proteins and should also have minimal impact on the expression yield of these polypeptides. Recombinant proteins containing the LPETG motif are also stable and unreactive until activated by the sortase enzyme. This affords a distinct advantage over the expressed protein ligation approach where premature cleavage of intein fusion proteins has been documented, leading to a reduction in the yield of final ligation products. The requisite lipid-modified nucleophiles compatible with sortase-mediated transpeptidation also have a very straightforward structural requirement of a short stretch of glycine residues and are prepared by standard solid-phase synthesis. Therefore, both natural and non-natural hydrophobic modifications can be easily incorporated using this system.

While the requirement for the sortase enzyme itself could be perceived as a disadvantage of the transpeptidation strategy, it should be noted that we routinely express batches with yields of >50 mg/L, and stock solutions of the enzyme can be stored for several months with no apparent degradation in enzymatic

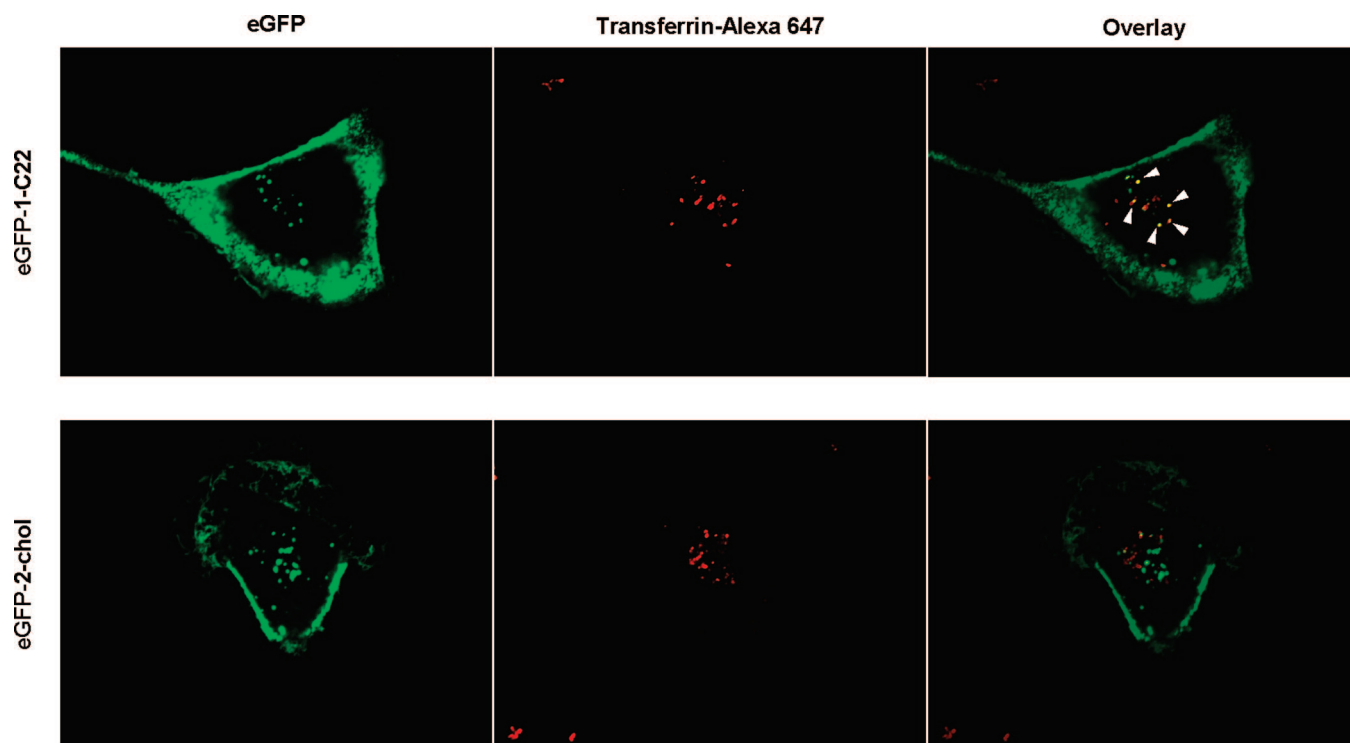(45) Sheff, D.; Pelletier, L.; O'Connell, C. B.; Warren, G.; Mellman, I. *J. Cell Biol.* **2002**, *156*, 797–804.

Figure 5. eGFP−**1**-C22 is able to access early endosomal compartments as determined by partial colocalization with transferrin−Alexa 647. U373 cells were incubated with eGFP−**1**-C22 (2.5 μg/mL) for 5 h in serum-free medium followed by the addition of transferrin−Alexa 647 (100 μg/mL) during the final 15 min. White arrows indicate vesicles with positive staining for both eGFP-1-C22 and transferrin-Alexa 647. Significant colocalization between eGFP−**2**-chol and transferrin−Alexa 647 was not observed.

activity. Moreover, in this work we provided a convenient method for removing sortase following transpeptidation. The ready availability of sortase and ease of removal following transpeptidation are important features as the enzyme is relatively inefficient and excess sortase is typically employed to drive the labeling reactions to high levels of conversion in a reasonable period of time.

In this initial work, we used eGFP as our model protein substrate due to its natural fluorescence and verified that eGFP proteins lipidated through sortase-mediated transpeptidation strongly associate with mammalian cells. We envision further applications of this chemical modification strategy to include anchorage of proteins to cellular membranes that do not normally occur on the cell surface to create artificial cell surface receptors. Additionally, because the transpeptidation strategy is able to attach extremely hydrophobic modifications, it could provide a new method for interfacing proteins with materials that possess poor water solubility, such as carbon nanotubes.

Supporting Information Available: Full experimental details and characterization data for triglycine nucleophiles and protein conjugates. These materials are available free of charge via the Internet at http://pubs.acs.org.

JA806779E