Published in final edited form as:

Chembiochem. 2012 April 16; 13(6): 888–894. doi:10.1002/cbic.201100764.

Site-specific protein modification using lipoic acid ligase and bis-aryl hydrazone formation

Dr. Justin D. Cohen, Peng Zou, and Prof. Alice Y. Ting

Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge MA, 02139 (USA), Fax: (+1) 617-253-7929

Alice Y. Ting: ating@mit.edu

Abstract

A screen of Trp37 mutants of *E. coli* lipoic acid ligase (LplA) produced enzymes capable of ligating an aryl-aldehyde or an aryl-hydrazine substrate to LplA's 13-amino acid acceptor peptide (LAP). Once site-specifically attached to recombinant proteins, aryl-aldehydes could be chemo-selectively derivatized with hydrazine-probe conjugates, and aryl-hydrazines could be derivatized in an analogous manner with aldehyde-probe conjugates. Such two-step labeling was demonstrated for AlexaFluor568 targeting to monovalent streptavidin *in vitro*, and to neurexin-1 β on the surface of living mammalian cells. To further highlight this technique, we also labeled low density lipoprotein receptor on the surface of live cells with fluorescent phycoerythrin protein to allow single molecule imaging and tracking over time.

Keywords

protein engineering; ligases; chemo-selective conjugation; fluorescence labeling; live cell imaging

Introduction

Chemoselective ligation chemistries are important tools for the detection, analysis, and perturbation of proteins in order to elucidate structure and function. For this reason, numerous bio-compatible versions of such chemistries have been developed, including the Staudinger ligation^[1], copper-catalyzed 3+2 azide-alkyne cycloaddition^[2], strain-promoted azide-alkyne cycloaddition^[3], and native chemical ligation^[4]. When these reactions are performed on proteins, the ability to conjugate in a site-specific manner can become crucial, because protein function is intimately linked to structure, and conjugation at the wrong site can reduce activity, block interactions with other proteins, or even cause structural destabilization. Therefore, these bio-compatible ligation reactions have been augmented through combination with site-specific protein labeling techniques. For example, our lab showed that the alkyl azide partner of the strain-promoted 3+2 cycloaddition reaction could be site-specifically introduced onto a peptide tag of recombinant cellular proteins through the action of *E. coli* lipoic acid ligase (LplA)^[5,6]. Thereafter, the protein-linked alkyl azide

Correspondence to: Alice Y. Ting, ating@mit.edu.

Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author.

could be chemoselectively derivatized with cyclooctyne-fluorophore conjugates. Similarly, unnatural amino acid mutagenesis has been used to introduce azides^[7] and ketones^[8] into proteins, for chemoselective derivatization by alkyne- and aminoxy-probe conjugates.

The goal of this work is to couple a powerful bio-compatible ligation reaction, aniline-catalyzed bis-aryl hydrazone formation^[9], with our LplA-based site-specific protein labeling methodology. Bis-aryl hydrazone formation between an aromatic aldehyde and 6-hydrazinopyridyl moiety is one of the fastest chemoselective ligation reactions, with a rate constant in the range of 10^2 – 10^3 M⁻¹ s⁻¹^[9], compared to 10^{-4} to 10^{-2} M⁻¹ s⁻¹ for the Staudinger ligation^[1] and 10^{-3} to 1 M⁻¹ s⁻¹ for strain-promoted azide-alkyne cycloaddition^[3;5]. The resulting covalent adduct is also highly stable at physiological pH, although, usefully, the linkage can be reversed when desired with the addition of aniline and hydroxylamine^[10]. The formation kinetics of the bis-aryl hydrazone bond are 20–900 fold higher than those of conventional hydrazone/oxime bonds between alkyl aldehydes/ketones and alkyl hydrazide or aminooxy probes^[9;11]. Studies have shown that 10 mM aniline treatment for 90 minutes (to catalyze oxime formation) is non-toxic to living mammalian cells^[12]. Despite all these attractive features, methods have not been described for the site-specific introduction of the bis-aryl hydrazone reaction partners into full-length proteins, either in vitro or within complex environments such as cell lysates or living cells.

Results and Discussion

E. coli LplA catalyzes highly sequence-specific lipoic acid conjugation to a 13-amino acid recognition sequence, called LAP (LplA Acceptor Peptide)^[13]. We have previously shown that mutation of the lipoic acid binding pocket can confer the ability to ligate a range of unnatural substrate structures, including 7-hydroxycoumarin^[14], an aryl azide photocrosslinker^[15], and trans-cyclooctene^[16]. To test if mutants of LplA could accept aryl aldehyde and aryl hydrazine substrates, we synthesized the two structures shown in Figure 1A, in addition to analogs with one less methylene. These four substrates were screened against wild-type LplA and the seven mutants shown in Figure 1B. We have previously found that the W37 position, which is located at the end of the lipoic acid binding tunnel, acts as a “gatekeeper” residue whose mutation allows LplA to accept substrates whose size and shape differ greatly from lipoic acid. We tested a small panel of W37 mutants which have previously shown activity for unnatural probe ligation^[14;16;17]. No activity was detected with any of the LplA mutants with the shorter aldehyde and hydrazine substrates (data not shown). However, the longer aryl aldehyde (“Ald”) shown in Figure 1A was recognized and ligated to the LAP peptide by several of the W37 mutants, with ^{W37I}LplA having the highest activity (Figure 1B). Using 1 μM ^{W37I}LplA, 500 μM Ald probe, and 150 μM LAP peptide, the reaction proceeded to 62 % completion in 5 minutes (Figure 1B).

We found that the aryl hydrazine (“Hyd”) probe was also ligated by many of the LplA mutants, but not as efficiently as the aryl aldehyde (“Ald”). Interestingly, the relative activity of the W37 mutants for the Hyd probe was similar to that with the Ald probe, with ^{W37I}LplA again having the highest activity. However, the overall activity with the Hyd probe was lower than that for the Ald probe, reacting to 50% completion using ^{W37I}LplA over 60 min. We determined the k_{cat} values for ^{W37I}LplA-catalyzed attachment of the Ald

and Hyd probes to LAP peptide. Ald ligation had a k_{cat} of $0.33 \pm .01 \text{ s}^{-1}$ while Hyd ligation had a k_{cat} of $0.021 \pm .003 \text{ s}^{-1}$ (Supporting Methods). Both ligations required ATP and could not be catalysed by wild-type LplA (Figure 1C). Identities of product peaks were confirmed by mass spectrometry (Figure S1).

In vitro protein labeling with LplA and bis-aryl hydrazone formation

We proceeded to test whether our LplA-mediated protein tagging method could be used for specific modification of proteins in vitro. We first turned our attention to streptavidin, a protein used ubiquitously in biotechnology due to its extremely high affinity and specificity for the small-molecule biotin. The ability to form site-specific conjugates of streptavidin to reporters such as fluorophores, enzymes (e.g., horse radish peroxidase, alkaline phosphatase) and phycoerthyrin could be extremely beneficial for enhancing activity and hence performance in applications ranging from ELISA and western blotting to live cell imaging.

We prepared streptavidin protein displaying a single LAP tag by utilizing our previously described monovalent streptavidin technology^[18]. Monovalent streptavidin is prepared by refolding one equivalent of wild-type streptavidin (“alive”, A) with three equivalents of “dead” (non-biotin-binding, D) streptavidin (Figure 2A). The resulting mixture of heterotetramers is then purified by gradient nickel affinity chromatography to isolate the species with exactly one wild-type subunit and three dead subunits, i.e., a single biotin binding site in the context of a tetrameric protein. We genetically fused the 13-amino acid LAP tag to the N-terminus of the wild-type subunit. Therefore, the resulting purified monovalent streptavidin (mSA) had a single LAP tag on the functional biotin-binding subunit of the tetrameric protein.

Labeling with ^{37}I LplA was performed with either Ald or Hyd substrate for 1 hr. After labeling, the crude mixtures were combined with either AlexaFluor568-hydrazide (AF568-Hyd) or fluorescein-aldehyde to selectively derivatize Ald or Hyd, respectively. Reactions were performed in the presence of 20 mM aniline catalyst at pH 5.0, overnight at room temperature. Figure 2B shows specific conjugation of AF568-Hyd to Ald-functionalized mSA-LAP, and fluorescein-aldehyde to Hyd-functionalized mSA-LAP. Importantly, negative controls with ATP omitted from the first step, or wild-type LplA used in place of ^{37}I LplA, showed no labeling.

To test if these site-specific mSA-LAP-fluorophore conjugates were active, we used them to perform labeling and imaging of biotinylated cell surface proteins. In Figure S2, HEK cells were transfected with plasmids for acceptor peptide (AP)-tagged low density lipoprotein receptor (LDLR) and endoplasmic reticulum (ER)-targeted biotin ligase. Previous work has shown that such conditions result in site-specific biotinylation of the AP tag in the ER lumen by biotin ligase^[19]. These cells were then treated with the mSA-LAP-AlexaFluor568 conjugate prepared as in Figure 2. Specific fluorescence labeling is seen in transfected cells expressing AP-LDLR and the nuclear yellow fluorescent protein (YFP) transfection marker. Labeling is not seen when the AP tag is mutated, excess biotin is added to quench mSA, or cells are not transfected. Hence Figure S2 demonstrates that the mSA-fluorophore conjugate prepared by LplA and bis-aryl hydrazone formation is functional for live cell labeling and imaging.

To illustrate generality, we performed similar labeling of two other proteins. One is alkaline phosphatase, an enzyme frequently attached to antibodies and streptavidin and used to generate a chromogenic signal in ELISA assays. We prepared a LAP fusion to the N-terminus of alkaline phosphatase, labeled with LplA and Ald, and then derivatized with fluorescein-Hyd. Figure S3 shows that this labeling was effective and dependent on ATP. The second protein we labeled was E2p, a 9 kDa domain of pyruvate dehydrogenase, one of LplA's natural protein substrates in *E. coli*.^[20] Figure S3 shows successful conjugation of fluorescein-Ald to Hyd-labeled E2p protein, as well as the reverse scheme.

A major benefit of the LplA protein labeling strategy is the exceptional sequence specificity of LplA. Hence, we explored the ability of our two-step labeling protocol to specifically conjugate fluorophores to LAP in complex mixtures containing thousands of competing proteins. Figure S4 shows a labeling experiment with a LAP-YFP fusion in mammalian cell lysate. AlexaFluor568 and fluorescein are conjugated to LAP-YFP only, and not any endogenous mammalian proteins, using LplA and bis-aryl hydrazone formation. Negative controls with LAP-YFP omitted or wild-type LplA in place of ^{W37I}LplA show no labeling.

Cell surface protein labeling with LplA and bis-aryl hydrazone formation

We next tested our labeling protocol in the context of the living mammalian cell surface. This context tests both the specificity of our labeling scheme, and its biocompatibility. In Figure 3, we co-transfected HEK cells with expression plasmids for LAP4.2-neurexin-1 β and a nuclear YFP transfection marker. Neurexin-1 β is a single transmembrane protein with an extracellular N terminus that functions as a neuronal adhesion protein. LAP4.2^[13] is a less hydrophobic variant of LAP that frequently gives improved surface targeting compared to LAP fusions (sequence given in Methods). Labeling was performed with ^{W37I}LplA, ATP, and 100 μ M Ald for 45 min at 37°C. Reagents were washed away, and then 100 μ M AF568-Hyd was added together with 10 mM aniline at 4°C for 30 min. After washing, cells were immediately imaged. Figure 3 shows that cell surface labeling was specific to transfected cells expressing LAP4.2-neurexin-1 β . Negative controls using wild-type LplA, ATP omitted, or a LAP containing an alanine mutation showed no labeling.

Cell surface protein labeling with phycoerythrin and single molecule imaging

Single molecule imaging is a powerful way to study protein trafficking in cells without losing information through ensemble averaging. Single molecule imaging in the cellular context requires fluorophores that are exceptionally bright and photostable. Quantum dots have excellent photophysical properties but commercial versions are very large and multivalent^[19]. Small organic dyes such as the AlexaFluors and cyanine dyes are much dimmer and require intense illumination to achieve reasonable signal-to-noise ratios at the single molecule level. Under these conditions, photobleaching occurs rapidly and prevents single molecule tracking for longer than a few minutes or even seconds^[21].

For biotechnological applications requiring extreme fluorophore brightness, such as fluorescence activated cell sorting (FACS), phycoerythrin has been used as a much brighter alternative to organic dyes and a smaller and less expensive alternative to QDs. R-phycoerythrin (PE) is a 240 kD protein with a disk shape (11 nm $\times$ 6 nm), containing 34

embedded phycobilin-type chromophores. It is usually obtained by purification from red algae^[22]. With an extinction coefficient (ϵ) of $2.0 \times 10^6 \text{ M}^{-1}\text{cm}^{-1}$ at 566 nm, and quantum yield (QY) of 0.85, it is >25 times brighter than AlexaFluor 568 ($\epsilon = 91,300 \text{ M}^{-1}\text{cm}^{-1}$ at 568 nm; QY = 0.69), which emits at the same wavelength.

PE has rarely been explored as a reagent for single molecule imaging. Previously, Irvine et al. used PE for single timepoint imaging of single peptide molecules binding to major histocompatibility complex (MHC) on the surface of antigen presenting cells^[23]. We wished to explore the use of our LplA method to target PE to specific cell surface proteins, and to image them at the single molecule level. Since PE can only be practically added to cells at low micromolar concentrations, it is essential that it be targeted using a method with an extremely high second order rate constant. For instance, calculations shows that the yield would be <1% using a targeting method with a rate constant of $\sim 0.1 \text{ M}^{-1} \text{ s}^{-1}$, such as azide-azadibenzocyclooctyne cycloaddition^[5], after 1 hour of labeling. With its extremely fast kinetics and cell compatibility, the bis-aryl hydrazone conjugation is therefore ideal for this application.

We prepared HEK cells expressing LAP4.2-neurexin-1 β as in Figure 3, and labeled them with the Hyd probe using ^{37}I -LplA. After labeling, cells were washed and treated with 20 mM aniline and PE modified with 4-formylbenzamide (PE-Ald). After 45 min the cells were washed and imaged. As shown in Figure 4A, clear labeling was observed in transfected cells. No labeling was seen in negative controls using wild-type LplA, with ATP omitted, or with an alanine mutation in LAP.

To perform single molecule imaging with PE, we prepared COS7 cells expressing LAP4.2-LDLR on their surfaces. LDLR is a constitutively internalized receptor that promotes the plasma clearance of LDL particles via clathrin-mediated endocytosis. A single-molecule imaging platform for LDLR based on our hydrazine-labeling technique could potentially provide insight into the mechanisms of LDLR targeting to clathrin-coated pits. We labeled the LDLR using our Hyd probe, followed by treatment with 20 mM aniline and PE-Ald. As shown in Figure 4B, individual labeled LDLR molecules appeared as single diffraction-limited spots on the cell surface, imaged by total internal reflection fluorescence (TIRF) microscopy. To confirm that the labeled spots were indeed single receptors and not aggregates, we compared the intensity distribution of > 2900 spots on cells to individual PE molecules randomly distributed on glass slides. Similar distributions are seen in Figure S5. The labeled receptors are also dynamic, as shown in time-lapse imaging experiments captured at a rate of 1 fps over a period of 60 s (Supporting Movie S1). The brightness of PE molecules offers high signal-to-background ratios unmatched by organic fluorophores, and photobleaching is reduced because of the lower laser intensity required for illumination.

Conclusion

In summary, LplA provides a general method for targeting small molecule probes with extremely high specificity to proteins in vitro, in lysate, and in living cells. Bis-aryl hydrazone formation is an extremely fast and biocompatible ligation reaction. By combining these two technologies in this study, we have developed a method to prepare protein-small

molecule and protein-protein conjugates with high specificity and great facility. We demonstrated the methodology on monovalent streptavidin, alkaline phosphatase, YFP, LDL receptor and neuexin-1 β , preparing conjugates to AlexaFluor568, fluorescein, and the extremely bright fluorescent protein phycoerythrin.

Presently, several methods exist to incorporate the reaction partners for conventional hydrazone/oxime formation, such as alkyl aldehydes via formylglycine generating enzyme (FGE)^[24;25] or ketones via incorporation of the unnatural amino acid *p*-acetylphenylalanine^[8]. In comparison to these methods, our LplA-based labeling takes advantage of the enhanced kinetics and stability of bis-aryl hydrazone formation, and we show that the same LplA mutant can target both the aryl aldehyde reaction partner *and* the hydrazinopyridine reaction partner.

Although we have demonstrated specific labeling on the surface of live cells, we note that expansion of this methodology for the labeling of intracellular proteins is likely to be complicated by the presence of endogenous aldehydes in the cell's interior.

This study expands the panel of probes that can be ligated by LplA mutants for specific labelling of proteins. In comparison to lipoic acid ligation by wild-type LplA (k_{cat} 0.22 s⁻¹), and 7-hydroxycoumarin ligation by ^{W37V}LplA (k_{cat} 0.019 s⁻¹), the measured k_{cat} for Ald ligation ($0.33 \pm .01$ s⁻¹) is extremely rapid and among the best for an unnatural probe / LplA mutant pair^[14]. The hydrophobic nature of the substrate recognition may partially explain the ten-fold greater activity of Ald versus Hyd, as the polar nature of the hydrazine may interfere with binding.

We envision the use of this method for preparation of improved conjugates of streptavidin and antibodies to reporters, particularly enzyme reporters such as peroxidase and alkaline phosphatase, where non-specific chemical conjugation could block their active sites and reduce activity. Such reagents could lead to improved sensitivity and reproducibility for ELISAs, western blots, and immunofluorescence staining. Finally, we note that our method showcases the use of phycoerythrin for single molecule imaging of specific proteins in the context of live cells. We believe this should be generalizable and provide an alternative to small organic dyes (due to increased brightness) and QDs (due to smaller size and lower cost).

Experimental Section

Plasmids

For expression of His₆-tagged LplA in *E. coli*, we used the LplA-pYJF16 plasmid.^[14] The cloning of LAP-streptavidin-pET21a for bacterial expression is described in the Supporting Methods. For expression of LAP fusion proteins in mammalian cells, we used LAP4.2-neurexin-1 β -pNICE^[14] and LAP4.2-LDLR-pcDNA4^[27]. Mammalian expression plasmids for BirA-ER^[28], AP-LDLR^[29], and H2B-YFP^[19] have been described previously.

***In vitro* screening for Ald and Hyd ligation activity (Figure 1)**

Ligation reactions were assembled as follows: 1 μ M of purified LpIA mutant^[14], 150 μ M synthetic LAP peptide (GFEIDKVVYDLDA), 5 mM ATP, 500 μ M of either Ald or Hyd probe, 5 mM magnesium acetate, and 25 mM Na₂HPO₄ pH 7.2 in a total volume of 20 μ L. Reactions were incubated for 5 to 60 min at 30 °C and then quenched with EDTA to a final concentration of 45 mM. Samples were diluted to a total volume of 80 μ L in conjugation buffer (10 mM Na₂HPO₄, 3.2 mM KH₂PO₄, 2.7 mM KCl, 140 mM NaCl, pH 5.0) and analyzed on a Varian Prostar HPLC using a reverse-phase C18 Microsorb-MV 100 column (250 $\times$ 4.6 mm). Chromatograms were recorded at 210 nm. For analysis of the aldehyde ligation reaction we used a 10-minute gradient of 30–60% acetonitrile in water with 0.1% trifluoroacetic acid under 1 mL/minute flow rate. For analysis of the hydrazine ligation reaction a gradient of 25–60% over 14 minutes with the same solvents was used. Percent conversions were calculated by dividing the product peak area by the sum of (product + starting material) peak areas. Reactions were performed in triplicate (Ald) or duplicate (Hyd) and the average values are shown. Reactions in Figure 1C were performed using the conditions above with a 70 minute reaction time for Ald and 120 minute reaction time for Hyd.

LAP-monovalent streptavidin expression and purification

Monovalent streptavidin containing a single LAP tag fused to the N-terminus of the “alive” subunit was expressed and purified as previously described^[28]. Briefly, the alive (LAP-tagged, His₆-tagged) and dead (untagged) subunits of streptavidin were expressed separately in *E. coli*. The inclusion bodies were solubilized and the alive and dead proteins were combined in a 3:1 ratio. After refolding to obtain a statistical mixture, monovalent streptavidin containing exactly one alive subunit and three dead subunits was purified using gradient nickel affinity chromatography. Monovalency was confirmed using a DNA gel shift assay. LAP-mSA was mixed with 250 bp biotinylated DNA at a 1:1 and 10:1 molar ratio and run on a 1.5% agarose gel. A band corresponding to binding of a single biotinylated DNA was observed. In comparison, wild-type streptavidin under the same conditions binds between 1 to 4 biotinylated DNA molecules.

***In vitro* labelling of LAP fusion proteins (Figure 2)**

Reactions were assembled using 2 μ M LAP-mSA, 500 nM ^{W37I}LpIA, 5 mM ATP, 100 μ M of either Ald or Hyd, 5 mM magnesium acetate, and 25 mM Na₂HPO₄ pH 7.2 in a total volume of 20 μ L. Reactions were incubated at room temperature for 1 hr. Each reaction was then diluted to a volume of 500 μ L of PBS and the buffer adjusted to pH 5 using HCl. Thereafter, the solution was concentrated to ~30 μ L using an ultrafiltration concentrator with a MWCO of 5 kDa (Vivaspin 500, GE Healthcare). This was repeated twice in order to fully exchange the buffer and eliminate excess probe. Conjugation was then performed by adding 20 mM aniline and 200 μ M of either AlexaFluor568-hydrazide (Invitrogen) or fluorescein-aldehyde (4FB-PEG3-fluorescein, Solulink). Reactions were incubated overnight and analyzed on a 10% SDS-PAGE gel. In gel fluorescence imaging was performed using a Fujifilm FLA-9000.

Mammalian cell culture

HEK and COS-7 cells were cultured in growth media consisting of Minimum Essential Medium (MEM, Cellgro) supplemented with 10% fetal bovine serum (FBS, PAA Laboratories). Cells were maintained at 37 °C under 5% CO₂. For imaging, HEK cells were grown on glass coverslips pre-treated with 50 µg/mL fibronectin (Millipore) to increase their adherence. COS-7 cells were grown in LabTek II chambered coverglass system 8-well plates.

Microscopy

Cells were imaged in Dulbecco's Phosphate Buffered Saline (DPBS) at room temperature. The confocal images in Figures 3 and 4A were collected on a Zeiss AxioObserver.Z1 microscope with a 40x oil-immersion objective and 2.5x Optovar. The images were collected in confocal mode using a Yokogawa spinning disk confocal head with a Quad-band notch dichroic mirror (405/488/568/647 nm). YFP (491 nm laser, 528/38 emission filter), AlexaFluor568/ Phycoerythrin (561 nm laser, 617/73 emission filter), and Normarski-type DIC images were collected using a Cascade II:512 camera and Slidebook software (Intelligent Imaging Innovations). Fluorescence images in each experiment were normalized to the same intensity range.

TIRF images (Figure 4) were acquired on the same microscope using a TIRF slider. YFP (491 nm laser excitation, 525/30 emission filter, 502 nm dichroic mirror), Alexa Fluor 568 / Phycoerythrin (561 nm laser excitation, 605/30 emission filter, 585 nm dichroic mirror) and Normarski-type DIC images were collected at 100x magnification using Slidebook software (Intelligent Imaging Innovations). Digital images (16 bit) were obtained with a cooled EMCCD camera (QuantEM:512SC, Photometrics) with exposure times between 50 ms and 200 ms.

Cell surface labelling (Figure 3 and 4)

For some constructs in this work (neurexin-1 β and LDLR), an alternative peptide sequence called LAP4.2^[13] was used (GFEIDKLVHDFPA), in order to boost cell surface expression levels. For Figure 3, HEK cells were transfected with 200 ng LAP4.2-neurexin-1 β and 200 ng H2B-YFP co-transfection marker plasmid, per 0.95 cm² cells at ~70% confluency, using Lipofectamine 2000 (Invitrogen). 15 hours after transfection, the growth media was removed, and the cells were washed three times with DPBS with 0.5% casein. Casein was added to DPBS for all washing and labeling steps as a blocking agent and was required to reduce non-specific sticking of the probes. The cells were then labeled by applying 100 µM Ald probe, 1 µM ^{W37I}LpIA, 1 mM ATP, and 5 mM Mg(OAc)₂ in DPBS with 0.5% casein at 37°C for 45 minutes. Cells were then washed three times with DPBS with 0.5% casein and treated with 10 mM aniline and 100 µM AlexaFluor568-Hydrazide at 4°C for 30 min. Cells were washed an additional three times and imaged live. The cell surface labeling in Figure 4A was performed in the same fashion with the following changes: labeling was done using Hyd probe for 45 min at room temperature, and the fluorophore conjugation was done using 3 µM PE-Ald (4FB-R PE, Solulink) for 45 min at 4°C. Labeling for Figure 4B was performed as described for Figure 4A except COS-7 cells were transfected with 200 ng LAP4.2-LDLR and 100 ng H2b-YFP co-transfection marker, only 20 µM Hyd probe was

used in the initial labeling, and 0.3 μ M PE-Ald with 20 mM aniline for 45 min was used for the fluorophore conjugation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding was provided by the NIH (R01 GM086214) and the American Chemical Society. Prof. Ian R. Henderson donated the pQUANTagen(kx) plasmid, and Daniel Liu provided plasmids and helpful advice.

References

1. van Berkel SS, van Eldijk MB, van Hest JCM. *Angew Chem Int Ed Engl.* 2011; 50:8806–8827. [PubMed: 21887733]
2. Hong V, Steinmetz NF, Manchester M, Finn MG. *Bioconjug Chem.* 2010; 21:1912–1916. [PubMed: 20886827]
3. Jewett JC, Sletten EM, Bertozzi CR. *J Am Chem Soc.* 2010; 132:3688. [PubMed: 20187640]
4. Kent SBH. *Chem Soc Rev.* 2009; 38:338–51. [PubMed: 19169452]
5. Yao JZ, Uttamapinant C, Poloukhine AA, Baskin JM, Codelli JA, Sletten EM, Bertozzi CR, Popik VV, Ting AY. *J Am Chem Soc.* 2012
6. Fernandez-Suarez M, Baruah H, Martinez-Hernandez L, Xie KT, Baskin JM, Bertozzi CR, Ting AY. *Nat Biotechnol.* 2007; 25:1483–1487. [PubMed: 18059260]
7. de Graaf AJ, Kooijman M, Hennink WE, Mastrobattista E. *Bioconjug Chem.* 2009; 20:1281–1295. [PubMed: 19191567]
8. Hutchins BM, Kazane SA, Staflin K, Forsyth JS, Felding-Habermann B, Smider VV, Schultz PG. *Chem& Biol.* 2011; 18:299–303. [PubMed: 21439474]
9. Dirksen A, Dawson PE. *Bioconjug Chem.* 2008; 19:2543–2548. [PubMed: 19053314]
10. Dirksen A, Yegneswaran S, Dawson PE. *Angew Chem Int Ed Engl.* 2010; 49:2023–2027. [PubMed: 20183838]
11. Dirksen A, Dirksen S, Hackeng TM, Dawson PE. *J Am Chem Soc.* 2006; 128:15602–15603. [PubMed: 17147365]
12. Zeng Y, Ramya TNC, Dirksen A, Dawson PE, Paulson JC. *Nat Methods.* 2009; 6:207–209. [PubMed: 19234450]
13. Puthenveetil S, Liu DS, White KA, Thompson S, Ting AY. *J Am Chem Soc.* 2009; 131:16430–16438. [PubMed: 19863063]
14. Uttamapinant C, White KA, Baruah H, Thompson S, Fernandez-Suarez M, Puthenveetil S, Ting AY. *Proc Natl Acad Sci U S A.* 2010; 107:10914–10919. [PubMed: 20534555]
15. Baruah H, Puthenveetil S, Choi YA, Shah S, Ting AY. *Angew Chem Int Ed Engl.* 2008; 47:7018–7021. [PubMed: 18677791]
16. Liu DS, Tangpeerachaikul A, Selvaraj R, Taylor MT, Fox JM, Ting AY. *J Am Chem Soc.* 2011
17. Jin X, Uttamapinant C, Ting AY. *ChemBioChem.* 2011; 12:65–70. [PubMed: 21154801]
18. Howarth M, Chinnapen DJF, Gerrow K, Dorrestein PC, Grandy MR, Kelleher NL, El Husseini A, Ting AY. *Nat Methods.* 2006; 3:267–273. [PubMed: 16554831]
19. Howarth M, Liu WH, Puthenveetil S, Zheng Y, Marshall LF, Schmidt MM, Wittrup KD, Bawendi MG, Ting AY. *Nat Methods.* 2008; 5:397–399. [PubMed: 18425138]
20. Green DE, Morris TW, Green J, Cronan JE, Guest JR. *Biochem J.* 1995; 309:853–862. [PubMed: 7639702]
21. Altman RB, Terry DS, Zhou Z, Zheng Q, Geggier P, Kolster RA, Zhao Y, Javitch JA, Warren JD, Blanchard SC. *Nat Methods.* 2012; 9:68–71. [PubMed: 22081126]

22. Chang, Wr; Jiang, T.; Wan, Zl; Zhang, Jp; Yang, Zx; Liang, Dc. J Mol Biol. 1996; 262:721–722. [PubMed: 8876649]
23. Irvine DJ, Purbhoo MA, Krogsgaard M, Davis MM. Nature. 2002; 419:845–849. [PubMed: 12397360]
24. Wu P, Shui WQ, Carlson BL, Hu N, Rabuka D, Lee J, Bertozzi CR. Proc Natl Acad Sci U S A. 2009; 106:3000–3005. [PubMed: 19202059]
25. Carrico IS, Carlson BL, Bertozzi CR. Nat Chem Biol. 2007; 3:321–322. [PubMed: 17450134]
26. Blanden AR, Mukherjee K, Dilek O, Loew M, Bane SL. Bioconjug Chem. 2011; 22:1954–1961. [PubMed: 21932849]
27. Cohen JD, Thompson S, Ting AY. Biochemistry. 2011; 50:8221–8225. [PubMed: 21859157]
28. Howarth M, Ting AY. Nat Protoc. 2008; 3:534–545. [PubMed: 18323822]
29. Zou P, Ting AY. ACSChem Biol. 2011; 6:308–313.

Figure 1.

Screening for aldehyde and hydrazine ligases. **(A)** Scheme for using lipoic acid ligase (LplA) to conjugate an aldehyde or hydrazine functional group to LAP fusion proteins (LAP is the 13-amino acid LplA Acceptor Peptide, with sequence GFEIDKLVWYDLDA^[13]). After ligation, the aldehyde or hydrazine moiety can be chemoselectively derivatized with fluorophores, photocrosslinkers, or even proteins (represented by red and green circles). The linkage is a hydrazone (highlighted in yellow) that is stabilized by conjugation to the two neighboring aromatic rings. **(B)** Wild-type (WT) LplA and seven W37 point mutants were screened against the aldehyde (Ald) and hydrazine (Hyd) substrates shown in (A), each at 500 μ M. Percent conversion of LAP peptide to the LAP-product conjugate is reported in the right columns. Data are represented as mean $\pm$ standard deviation. N.D., not detected. Representative traces are shown in Figure S1. **(C)** HPLC traces showing complete ligation of Ald (left) and Hyd (right) to LAP catalyzed by ^{W37I}LplA using extended reaction times of 70 and 120 min respectively. Negative controls are shown with ATP omitted (red trace) and ^{W37I}LplA replaced by wild-type LplA (black trace). Starred peaks were analyzed by mass spectrometry in Figure S1.

Figure 2.

Site-specific fluorophore conjugation to monovalent streptavidin (mSA). (A) Overview of site-specific labeling of mSA. mSA consists of three “dead” subunits and one “alive” subunit that has a single N-terminal LAP fusion. This allows site-specific conjugation of mSA to a single fluorophore (red sphere). (B) mSA with a single LAP tag was labeled with ^{W37I}LpIA and 100 μM of either Ald substrate (lanes 1–3) or Hyd substrate (lanes 4–6). The conjugation reactions were analysed by SDS-PAGE and in-gel fluorescence in which lanes 1–3 were derivatized with AlexaFluor568-Hyd, while lanes 4–6 were derivatized with fluorescein-Ald. Negative controls were performed with ATP omitted from the ligation reaction (lanes 2 and 5), or ^{W37I}LpIA replaced by wild-type LpIA (lanes 3 and 6). Coomassie staining of the same samples is shown at right. mSA-AlexaFluor568 conjugated as in lane 1 was used for cell surface protein labeling in Figure S2.

Figure 3.

Fluorophore labeling of cell surface proteins. HEK cells expressing LAP4.2-neurexin-1 β were labeled with ^{37}I LplA and Ald substrate for 45 min, washed, then treated with AlexaFluor568-Hyd in the presence of 10 mM aniline for 30 min, before imaging (top row). Nuclear YFP is a co-transfection marker. Negative controls were performed with wild-type LplA, ATP omitted, and an alanine mutation in LAP. Scale bars, 10 μm .

Figure 4.

Protein labeling with phycoerythrin on the surface of living mammalian cells. **(A)** HEK cells expressing LAP4.2-neurexin-1 β were labeled with ^{37}I LpIA and Hyd substrate for 45 min, washed, then treated with 3 μM phycoerythrin-Ald in the presence of 10 mM aniline for 45 min, before confocal imaging. Negative controls are shown with wild-type LpIA, ATP omitted, and an alanine mutation in LAP. Nuclear YFP was a co-transfection marker. **(B)** COS7 cells expressing LAP4.2-LDL receptor were labeled as in (A), except that a lower concentration of phycoerythrin-Ald (0.3 μM) and 20 mM aniline was used and the cells were imaged in TIRF mode. Nuclear YFP is a transfection marker. Images of a transfected cell (top) and untransfected cell (bottom) are shown. **(C)** Enlarged view of the boxed area in (B) showing individual phycoerythrin particles conjugated to LAP4.2-neurexin-1 β on the COS7 cell surface. Scale bars, 10 μm . A time lapse movie showing the labeled cell surface LDL receptor over a period of 1 min is shown in supporting movie 1.