Structure-Guided Engineering of a Pacific Blue Fluorophore Ligase for Specific Protein Imaging in Living Cells

Justin D. Cohen, Samuel Thompson, and Alice Y. Ting*

Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

* Supporting Information

ABSTRACT: Mutation of a gatekeeper residue, tryptophan 37, in E. coli lipoic acid ligase (LplA), expands substrate specificity such that unnatural probes much larger than lipoic acid can be recognized. This approach, however, has not been successful for anionic substrates. An example is the blue fluorophore Pacific Blue, which is isosteric to 7-hydroxycoumarin and yet not recognized by the latter’s ligase (W37VLplA) or any tryptophan 37 point mutant. Here we report the results of a structure-guided, two-residue screening matrix to discover an LplA double mutant, E20G/W37TLplA, that ligates Pacific Blue as efficiently as W37VLplA ligates 7-hydroxycoumarin. The utility of this Pacific Blue ligase for specific labeling of recombinant proteins inside living cells, on the cell surface, and inside acidic endosomes is demonstrated.

Probe incorporation mediated by enzymes (PRIME) is a method to tag recombinant proteins in living cells with chemical probes. The method utilizes mutants of E. coli lipoic acid ligase (LplA), whose natural function is to ligate lipoic acid onto acceptor proteins involved in oxidative metabolism.1 Instead of lipoic acid, LplA mutants catalyze the covalent attachment of unnatural chemical probes, such as 7-hydroxycoumarin,2 an aryl azide,3 or an alkyl azide,4 onto recombinant proteins fused to a 13-amino acid recognition sequence called LAP (LplA acceptor peptide).5 The advantages of PRIME in comparison to other protein labeling methods are the small tag size, compatibility with the interior of living cells, and high labeling specificity.6 In previous studies, uptake of the unnatural substrate by LplA was achieved by mutation of a “gatekeeper” residue, W37, at the end of the lipoic acid binding pocket (Figure 1B). Enlarging this pocket, for example by a W37 → V mutation, allows LplA to accept structures much larger than lipoic acid, such as the blue fluorophore 7-hydroxycoumarin (HC) (Figure 1A, top). In the course of our screening, however, we discovered several structures that are not accepted by W37 point mutants. One of the most interesting examples is Pacific Blue (PB),5 a fluorophore that differs from HC only in the two fluorine atoms at C6 and C8 of the coumarin ring (Figure 1A, bottom). Because of these two electron-withdrawing fluorines, PB has a reduced 7-hydroxyl pKₐ of 3.7, compared to 7.5 for HC,7 and is therefore fully anionic and fluorescent at physiological pH (7.4) as well as endosomal pH (5.5–6.5). In contrast, only ~50% of HC is in the anionic and fluorescent form at pH 7.4, and it is mostly protonated and hence nonfluorescent in acidic endosomes.8

We hypothesized that PB is not recognized by HC’s ligase, W37VLplA, and other W37 point mutants because its negative charge clashes with the mostly hydrophobic binding pocket of LplA.8 In addition, near the W37 gatekeeper residue at the end of the lipoic acid binding tunnel is a negatively charged side chain, E20, that may electrostatically repel PB (Figure 1B). E20 could play a steric role as well, since a previous alanine scan in the lipoate binding pocket identified E20A as one of two mutants (along with W37A) with any detectable ligation activity for an aryl azide probe.3

The goal of this work was to use PB as a model compound to explore strategies for engineering new LplA activity, such as recognition of anionic substrates, beyond point mutations at W37. A PB ligase is also a useful alternative to HC ligase for studying proteins in acidic cellular compartments, where HC fluorescence is very low. By performing in-vitro screens using a panel of E20 and W37 single and double mutants, we discovered that E20G/W37TLplA ligates PB with comparable kinetics to W37VLplA ligation of HC (Figure 1A).
In-Vitro Screening of LplA Mutants (Figure 2A). Ligation reactions were assembled as follows for Figure 2A: 2 μM purified LplA mutant, 150 μM synthetic LAP peptide (GFEIDKVVYDLD; synthesized by the Tufts Peptide Synthesis Core Facility), 5 mM ATP, 500 μM fluorophore probe, 5 mM magnesium acetate, and 25 mM Na2HPO4 pH 7.2 in a total volume of 25 μL. Reactions were incubated for 12 h at 30 °C.

LplA mutant/probe combinations giving high activity under these conditions were then reassayed with 10-fold lower probe (50 μM) for 2 h.

Product formation was analyzed by ultraperformance liquid chromatography (UPLC) on a Waters Acquity instrument using a reverse-phase BEH C18 column 1.7 μM (1.0 × 50 mm) with inline mass spectrometry. Chromatograms were recorded at 210 nm. A gradient of 30−70% (acetonitrile + 0.1% trifluoroacetic acid) in (water with 0.1% trifluoroacetic acid) over 0.78 min was used.

Further in-Vitro Screening of Top Five LplA Double Mutants (Figure 2B,C). Reactions for the top five LplA double mutants were assembled as above, but with 500 μM probe and a reaction time of 45 min. Reactions were quenched with EDTA to a final concentration of 100 mM. Product formation was analyzed on a Varian Prostar HPLC using a reverse-phase C18 Microsorb-MV 100 column (250 × 4.6 mm). Chromatograms were recorded at 210 nm. We used a 10 min gradient of 30−60% acetonitrile in water with 0.1% trifluoroacetic acid under 1 mL/min flow rate. Percent conversions were calculated by dividing the product peak area by the sum of (product + starting material) peak areas.

Michaelis–Menten Kinetic Assay. The Michaelis–Menten curve shown in Figure S4 was generated as previously described. Reaction conditions were as follows: 2 μM E20G/W37TLplA, 600 μM synthetic LAP peptide, 2 mM magnesium acetate, and 25 mM Na2HPO4 pH 7.2.

Mammalian Cell Culture and Imaging. HEK and HeLa 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 pretreated with 50 μg/mL fibronectin (Millipore) to increase their adherence.

Cells were imaged in Dulbecco’s Phosphate Buffered Saline (DPBS) at room temperature. The images in Figures 3 and 4 were collected on a Zeiss AxioObserver.Z1 microscope with a 40× oil-immersion objective and 2.5× Optovar, equipped with a Yokogawa spinning disk confocal head containing a Quad-band notch dichroic mirror (405/488/568/647 nm). Pacific Blue/coumarin (405 nm laser excitation, 445/40 emission filter), YFP (491 nm laser excitation, 528/38 emission filter), Alexa Fluor 568 (561 nm laser excitation, 617/73 emission filter), and DIC images were collected using Slidebook software (Intelligent Imaging Innovations). Images were acquired for 100 ms to 1 s using a Cascade II:512 camera. Fluorescence images in each experiment were normalized to the same intensity range.

Cell Surface Labeling. HEK cells were transfected with 200 ng of LAP4.2-LDLR-pcDNA4 and 100 ng of H2B-YFP cotransfection marker plasmid, per 0.95 cm2 at ∼70% confluency, using Lipofectamine 2000 (Invitrogen). Fifteen hours after transfection, the growth media was removed, and the cells were washed three times with DPBS. The cells were...
were performed under both forcing conditions (12 h, 500 μM of each probe. N.D. indicates not detected. The best LplA mutants gave 50% conversion to PB ligation product in a 12 h reaction, >50% conversion in 2 h reaction. (B) Quantitative product yields for the top five PB ligases in (A), after 45 min reaction with 500 μM ligase, 1 mM ATP, and 5 mM Mg(OAc)2 in DPBS and three at E20) and their 12 crossed double mutants, shown in Figure 2A. Screening was performed using 500 μM Pacific Blue or hydroxycoumarin probe, 2 μM ligase, 1 mM ATP, and 5 mM Mg(OAc)2, in DPBS at room temperature for 40 min. Cells were then washed three times with DPBS and either imaged immediately or incubated at 37 °C for an additional 30 min to allow receptor internalization prior to imaging. 

Intracellular Protein Labeling. HEK cells were transfected at ~70% confluency with 200 ng of LAP-YFP-NLS-pcDNA3 and 50 ng of FLAG-E20G/W37TLplA-pcDNA3 per 0.95 cm² using Lipofectamine 2000 (Invitrogen). Fifteen hours after transfection, the growth media was removed, and the cells were washed three times with serum-free MEM. The cells were labeled by applying 20 μM PB3-AM1 in serum-free MEM at 37 °C for 20 min. The cells were then washed three times with fresh MEM. Excess probe was removed by changing the media several times over 40 min.

To visualize LplA expression levels, cells were fixed with 3.7% formaldehyde in PBS pH 7.4 for 10 min, followed by methanol at −20 °C for 5 min. Fixed cells were washed with DPBS and then blocked overnight with blocking buffer (3% BSA in DPBS with 0.1% Tween-20). Anti-FLAG M2 antibody (Sigma) was added at a 1:300 dilution in blocking buffer for 1 h at room temperature. Cells were then washed three times with DPBS before treatment with a 1:300 dilution of goat anti-mouse antibody conjugated to Alexa Fluor 568 (Invitrogen) in blocking buffer for 1 h at room temperature. Cells were washed three times with DPBS prior to imaging.

For labeling of vimentin-LAP (Figure 4B), HeLa cells were transfected with 250 ng of vimentin-LAP-Clontech, 50 ng of FLAG-E20G/W37TLplA-pcDNA3, and 100 ng of H2B-YFP transfection marker per 0.95 cm² using Lipofectamine 2000. Labeling was performed as above, with an extended 60 min washout period to remove excess probe. Cells were then imaged live in DPBS.

We note that, compared to intracellular labeling with hydroxycoumarin, labeling with PB3 generally requires longer washout times, up to 60 min in some cases. Shorter wash times result in higher PB background in all cells.

## RESULTS

### Screening for a Pacific Blue Ligase.

On the basis of the LplA crystal structure (Figure 1B),8 we decided to focus our engineering efforts on the W37 and E20 positions. We started with a preliminary screen of 19 W37 point mutants and 14 E20 point mutants, against four probe structures (Figure S1). These four structures, shown in Figure 2A, are two Pacific Blue probes and two Pacific Blue probes with shorter (n = 3) and longer (n = 4) linkers (PB3 and PB4) and two analogous 7-hydroxycoumarin probes (HC3 and HC4). Some Pacific Blue (PB) ligation product was detected after a 12 h reaction with W37T, V, I, and A LplA mutants (Figure S1), so we decided to introduce these mutations into our next screen. Note that the activity of the best point mutant, W37T

LplA, which gave ~50% conversion to PB ligation product after 12 h, is too slow for practical utility. For E20, none of the tested point mutants gave product with any of the four probes after 12 h. Nevertheless, in our next screen, we included E20 mutations to the smaller, neutral side chains Gly, Ala, and Ser.2,9

Our next library consisted of 7 single mutants (four at W37 and three at E20) and their 12 crossed double mutants, shown in Figure 2A. Screening was performed using 500 μM probe in an overnight reaction. Any ligase/probe combination with high activity under these conditions was reassayed using 50 μM

---

Figure 2. Screening of LplA mutants for Pacific Blue ligation activity. (A) Relative product conversions measured for 19 LplA single and double mutants with two hydroxycoumarin (HC) probes and two Pacific Blue (PB) probes. HC3 and PB3 have n = 3 linkers, and HC4 and PB4 have n = 4 linkers. To generate these grids, ligation reactions were performed under both forcing conditions (12 h, 500 μM probe) and milder conditions (2 h, 50 μM probe) and analyzed by ultraperformance liquid chromatography, as described in the Experimental Procedures. Sample traces are shown in Figure S1. The activity grid was generated with the following tiers: no activity, <25% conversion in a 12 h reaction, 25–50% conversion in a 12 h reaction, <25% conversion in 2 h reaction, 25–50% conversion in 2 h reaction, >50% conversion in 2 h reaction. (B) Quantitative product yields for the top five PB ligases in (A), after 45 min reaction with 500 μM of each probe. N.D. indicates not detected. The best LplA mutants for PB3, HC3, and HC4 are highlighted. Errors are reported as standard errors of the mean. (C) HPLC trace showing formation of LAP-PB3 conjugate, catalyzed by our best PB ligase, E20G/W37TLplA. The identity of the LAP-PB3 peak was confirmed by mass spectrometry, shown in Figure S3. Traces below show negative control reactions with ATP omitted (red) or E20G/W37TLplA replaced by wild-type LplA (black).
probe in a 2 h reaction. As before, the E20 single mutants had no detectable activity (Figure 2A). The W37 single mutants were minimally active with both PB probes, although high activity was seen with HC3 and HC4. The best single mutant/probe pair was W37VLplA with HC4.

The LplA double mutants, however, had interesting patterns of activity with PB. Although none of the mutants ligated PB4 efficiently, PB3 was ligated well by five double mutants (Figure 2A; re-evaluated quantitatively in Figure 2B). The best two have the W37T mutation, suggesting that not only size reduction but also polarity increase at this position is beneficial for PB recognition. We noticed that the W37A mutation performed poorly in the context of all double mutants for all four probes, perhaps because it destabilizes the binding pocket. The best E20 mutation to pair with W37T was Gly, perhaps because it generates the most space and conformational freedom. Together, our observations suggest that W37 and E20 mutations work synergistically to allow PB uptake: W37 mutations enlarge the binding pocket, while E20 mutations remove repulsive electrostatic interactions (Figure 1C).

We proceeded to fully characterize our best PB ligase to emerge from this screen, E20G/W37TLplA. First, HPLC analysis of the ligation reaction was repeated (Figure 2C), alongside negative controls omitting ATP or replacing PB ligase with wild-type LplA. Second, the kinetic constants for PB3 ligation to LAP were measured by HPLC (Figure S4). Both $k_{\text{cat}}$ ($0.014 \pm 0.001 \text{s}^{-1}$) and $K_M$ ($17.5 \pm 4.3 \mu M$) values are comparable to those previously determined for HC4 ligation catalyzed by W37VLplA ($k_{\text{cat}}$ $0.019 \pm 0.004 \text{s}^{-1}$ and $K_M$ $56 \pm 20 \mu M$). Finally, we tested the sequence specificity of PB3 ligation by labeling a LAP fusion protein within mammalian cell lysate. Figure S5 shows that only LAP is labeled by PB ligase and not any endogenous mammalian proteins.

**Cell Surface Labeling with Pacific Blue Ligase.** To test our PB ligase on living cells, we first performed labeling of a cell surface protein. The neuronal adhesion protein neurexin-1 with LAP4.2 (a variant of LAP5 whose sequence is given in the Experimental Procedures) fused to its extracellular N-terminus was expressed in human embryonic kidney (HEK) cells. Labeling was performed by adding purified PB ligase, PB3 probe, and ATP to the cellular media for 30 min. Figure S6 shows a ring of PB fluorescence around cells expressing LAP4.2-neurexin, as indicated by the presence of the cotransfection marker, whereas untransfected neighboring cells are not labeled. Negative controls performed with wild-type LplA, ATP omitted, or an alanine mutation in LAP resulted in no visible labeling (Figure S6).

A potential advantage of PB ligase over HC ligase is for visualization of proteins in acidic organelles, where HC fluorescence is low due to its $pK_a$ of 7.5. To test this experimentally, we used PB ligase or HC ligase to label LAP4.2-LDL receptor on the surface of HEK cells. After labeling, cells were incubated for 30 min at 37 °C to allow internalization of fluorescently tagged receptors. Figure 3 shows that PB-tagged LAP4.2-LDL receptor is clearly visible.
visible within internalized puncta, whereas HC-tagged LAP4.2-LDL receptor is not. Separate experiments showed that many of the PB-labeled internal puncta overlap with FM4-64, an endosomal marker (data not shown).

Intracellular Protein Labeling with Pacific Blue Ligase. We tested PB ligase for labeling of intracellular proteins in living mammalian cells. To deliver PB3 across the cell membrane, we first protected the carboxylic acid and 7-hydroxyl groups of PB3 with acetoxyethyl (AM) groups to give PB3-AM₁ (structure shown in Supporting Information). Endogenous intracellular esterases remove the AM groups to give PB3 inside the cell. HEK cells were cotransfected with plasmids for PB ligase and LAP-YFP-NLS (NLS is a nuclear localization signal; YFP is yellow fluorescent protein). To perform labeling, PB3-AM₁ was incubated with cells for 20 min, and then the media was replaced 3 times over 40 min to allow cells to pump out excess, unconjugated probe. The cells were then fixed, and anti-FLAG immunostaining was performed to visualize enzyme expression. As expected for specific labeling, PB fluorescence overlapped well with the YFP fluorescence of LAP-YFP-NLS (Figure 4). PB is not seen in neighboring untransfected cells. PB labeling is also absent when wild-type LpLA is used in place of PB ligase, or the PB-YFP-NLS contains a Lys → Ala mutation in the LAP sequence. To illustrate generality, we also performed PB labeling in live cells of vimentin-LAP, an intermediate filament protein (Figure 4B).

**DISCUSSION**

In this study, we identified an LpLA double mutant capable of recognizing and ligating a charged probe, Pacific Blue. Unlike previous studies where simple enlargement of the binding pocket via a point mutation at W37 was sufficient to allow recognition of large hydrophobic probes, the synergistic effect of mutating both the E20 and W37 positions was required for recognition of Pacific Blue. Guided by the LpLA crystal structure, we were able to create a small and focused library of single and double LpLA mutants to screen for the desired PB ligation activity. No single mutation had significant activity, but the augmentation of the most active W37 single mutants by E20 mutations resulted in a kinetically efficient PB ligase. We anticipate that these insights into the substrate binding pocket of LpLA will prove useful in future engineering efforts. The engineered PB ligase has k₄₉ and k₅₉ values similar to those of our previously reported 7-hydroxycoumarin ligase. PB ligase also retained sequence specificity for LAP over all endogenous mammalian proteins and could therefore be used for specific protein labeling inside and on the surface of living mammalian cells.

With this report, PRIME labeling can now be performed with any of three coumarin probes: Pacific Blue, 7-hydroxycoumarin, or 7-aminocoumarin (AC). This decision of which coumarin to use is dependent on the specific application. HC is the brightest of the three probes, followed by PB and then AC due to its decreased extinction coefficient. However, as demonstrated here, PB and AC have the added benefit of pH insensitivity, whereas the pKₐ of HC makes it unsuitable for imaging in acidic organelles such as endosomes.

**REFERENCES**