Published in final edited form as:

Angew Chem Int Ed Engl. 2008 ; 47(37): 7018–7021. doi:10.1002/anie.200802088.

An Engineered Aryl Azide Ligase for Site-Specific Mapping of Protein-Protein Interactions via Photocrosslinking**

Hemanta Baruah, Sujiet Puthenveetil, Yoon-Aa Choi, Samit Shah, and Alice Y. Ting*

Department of Chemistry, Room 18-496, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 (USA)

Abstract

We re-engineered the small-molecule binding site of *E. coli* lipoic acid ligase (LplA) to accept a fluorinated aryl azide probe in place of lipoic acid. The mutated ligase can covalently attach the aryl azide to recombinant proteins fused to a 17-amino acid recognition sequence for LplA (called "LAP"). Labeling is highly specific, modifying LAP fusion proteins to the exclusion of all endogenous mammalian proteins. In cell lysate, we labeled FKBP with aryl azide and demonstrated rapamycin-dependent photocrosslinking to FRB.

Keywords

lipoic acid ligase; photocrosslinking; protein engineering; protein-protein interactions; site-specific protein; labeling

Protein-protein interactions (PPIs) underlie all signaling networks in cells, and thus technologies for identifying and characterizing them are extremely important. One powerful approach to PPI detection consists of photocrosslinking.[1,2] A photocrosslinker probe such as an aryl azide (Figure 1a) or benzophenone is conjugated to a protein of interest, and irradiation with UV light generates a reactive species that covalently crosslinks to any nearby protein (Figure 1b). Purification of the crosslinked complexes and analysis by mass-spectrometry, Western blotting, or other techniques reveals the identities of the interaction partners of the protein of interest.

PPI detection by photocrosslinking has considerable advantages over other, more commonly-used PPI detection methods, such as yeast two-hybrid,[3] protein complementation assays (e.g., split-Green Fluorescent Protein,[4] split- β -lactamase,[5] and split-luciferase,[6,7] and FRET (Fluorescence Resonance Energy Transfer[8]). For instance, photocrosslinking can detect interactions with endogenous proteins, whereas other methods detect only interactions with recombinant reporter-fused proteins. The small size of photocrosslinking probes introduces minimal steric bulk, which reduces the rate of false negatives. Transient interactions can be captured. Finally, photocrosslinking is compatible with all cell types and all subcellular compartments, in contrast to yeast two-hybrid, for instance.[9] Yet photocrosslinking is rarely used in living cells for PPI detection, principally for the reason that chemical photocrosslinkers cannot be easily targeted to specific proteins of interest within the cellular context. The best

**This work was supported by the U.S. National Institutes of Health (R01 GM072670 and P20 GM072029) and the Sloan and Dreyfus Foundations. Samsung Corporation provided a Lee Kun Hee scholarship to Y-A. C. We thank Susan M. Lydon for her assistance with graphics. We thank Kathleen T. Xie and Jackie Chan for help with site-directed mutagenesis experiments and Marta Fernández-Suárez for her assistance and advice.

*Fax: (+1) 617-253-7929, E-mail: ating@mit.edu.

current method for photocrosslinker introduction is unnatural amino acid mutagenesis.[10–16] The non-position specific form of this technique[15,16] introduces high background, reduced signals, and potential interference with protein function. Position-specific unnatural amino acid mutagenesis[10–14] is highly powerful, but limited for mammalian cell applications by the prevalence of natural amber codons,[17] low suppression efficiencies,[18] and the generation of prematurely truncated protein product, which can sometimes produce dominant negative effects.[19] A simple, robust, high efficiency, and minimally-invasive method for photocrosslinker targeting in cells is required, to bring the advantages of this approach for PPI detection to cell biology.

As a first step towards this goal, we report here the design and characterization of an aryl azide photoaffinity probe ligase. To produce this ligase, we re-engineered the active site of the *E. coli* enzyme lipoic acid ligase (LplA)[20] to accept a fluorinated aryl azide probe and catalyze its covalent conjugation to a recognition peptide (Figure 1b). The ligation reaction is extremely sequence-specific; endogenous mammalian proteins are not recognized and modified. We demonstrate the utility of this method by performing site-specific aryl azide ligation to FK506 binding protein (FKBP) in mammalian cell lysate, and then showing rapamycin-dependent photocrosslinking to its interaction partner, FKBP-rapamycin binding domain (FRB), also within lysate.

E. coli LplA is a 38 kD enzyme that catalyzes the ATP-dependent ligation of the cofactor lipoic acid (Figure 1a) onto a specific lysine sidechain of one of its three endogenous protein substrates in *E. coli*,[20,21] or a 17-amino acid unnatural peptide called LAP (“LplA Acceptor Peptide”[22]). We previously harnessed LplA for site-specific fluorophore tagging of LAP fusion proteins on the surface of mammalian cells[22] by capitalizing on (1) LplA’s extremely high sequence specificity, which prevents it from recognizing and modifying any endogenous mammalian proteins, and (2) the malleability of LplA’s active site, which allows it to accept and ligate alkyl azide analogs of lipoic acid to LAP. Azide-derivatized cell surface LAPs were selectively imaged with cyclooctyne-fluorophore conjugates.[22]

For site-specific photocrosslinker introduction, we synthesized two fluorinated aryl azide substrates for LplA (Figure 1a; syntheses shown in Supporting Figure 1). These structures represent a compromise between high photocrosslinking efficiency,[1,2] longer activation wavelength, and small probe size, to better fit within the LplA active site. Using an HPLC assay[22] with either E2p (a hybrid lipoyl domain derived from one of LplA’s natural protein substrates[21] or LAP-HP1 fusion[22] as protein substrate, we tested the ability of wild-type LplA to utilize probes **1** or **2**. No product conjugate was detected in the reactions with either probe (data not shown).

Because aryl azide probes **1** and **2** are both considerably larger than lipoic acid or the alkyl azide used in our previous study,[22] we decided to mutagenize the LplA active site to increase its volume. The co-crystal structure of *E. coli* LplA with lipoic acid has been reported,[23] but a number of unusual features about this structure led us to instead examine the co-crystal structure of *T. acidophilum* LplA with lipoyl-AMP[24] (Supporting Figure 2). The structure shows 9 amino acid sidechains (L18, D21, Y39, R72, Y73, T74, H81, L159, and H161) that lie within 7.5 Å of the dithiolane ring of lipoic acid. We aligned the *E. coli* and *T. acidophilum* LplA protein sequences, and individually mutated each of the corresponding amino acids in the *E. coli* enzyme (L17, E20, W37, R70, S71, S72, H79, F147, and H149) to alanine, in order to decrease their size. HPLC testing of this initial panel of 9 LplA mutants, using LAP-HP1 as protein substrate, showed that W37A, E20A, and S71A could accept probe **1** to a small degree (Supporting Figure 3). Under identical reaction conditions, W37A gave more product than did E20A and S71A (Supporting Figure 3). No product was detected for probe **2** with any of the mutants, despite the fact that it is smaller, showing that other factors

besides size play a role in substrate recognition. Based on these hits, we then produced a second panel of LplA mutants, shown in Figure 2a. We tested each of these using the same HPLC assay, with LAP-HP1, and found that W37V LplA exhibited the highest probe **1** ligation activity, followed by W37S (Figure 2a). Under identical conditions (saturating probe **1**, ATP, and LAP-HP1, 30 min), W37V gave 20-fold more product than did W37A, the best enzyme from the original mutant panel (Figure 2a). It is interesting that W37V LplA is superior to W37A LplA; we surmise that W37V better complements the shape and size of probe **1**, whereas W37A may over-enlarge the LplA active site. We selected W37V as our best mutant and used this clone for all subsequent experiments.

We characterized the W37V-catalyzed ligation of aryl azide **1** in terms of kinetics and specificity. First, we repeated the HPLC assay, with negative controls with ATP omitted, or W37V replaced by wild-type LplA. As expected, product was only observed with ATP and W37V mutant (Figure 2b). We collected the product peak from the HPLC and characterized it by mass-spectrometry. Supporting Figure 4 shows that the mass corresponds to the calculated mass for one molecule of aryl azide **1** plus one LAP-HP1 protein, minus water. Next, we measured the kinetics of ligation. Using HPLC as readout, we determined a k_{cat} of $0.31 \pm 0.04 \text{ s}^{-1}$, which surprisingly is even faster than the k_{cat} for the natural ligation reaction (lipoylation of E2p, 0.25 s^{-1})[22] (Supporting Figure 5). Separate measurements showed that the K_{m} was $80 \mu\text{M}$ or less (data not shown), which is considerably greater than the K_{m} of wild-type LplA for lipoic acid ($1.7 \mu\text{M}$ [20] or $4.5 \mu\text{M}$ [23]).

We were concerned that re-modeling of the LplA active site may change the peptide specificity of the enzyme. Thus we characterized the specificity of W37V-catalyzed aryl azide ligation in three ways. First, we prepared total lysate from HEK cells expressing a LAP fusion to cyan fluorescent protein (LAP-CFP[22]). We labeled the lysate under forcing conditions with aryl azide, ATP, and W37V LplA, then we derivatized the aryl azide with phosphine FLAG using the Staudinger ligation.[25] Western blotting with anti-FLAG antibody showed that even though LAP-CFP expression level is so low that it cannot be detected above endogenous proteins in the Coomassie-stained gel, only LAP-CFP receives the aryl azide, and endogenous proteins in the lysate are not modified by W37V LplA (Figure 2c). Negative controls showed that labeling is specific for LAP and dependent on W37V (Figure 2c).

For our second specificity test, we performed aryl azide **1** ligation on three non-LAP proteins: FRB-AP, FKBP-AP and BCCP-87.[26] We characterized the products by mass spectrometry (data not shown). None were found to be modified by aryl azide, despite exposure to reaction conditions that gave complete conversion of LAP-HP1 to the aryl azide conjugate.

Third, we tested peptide specificity using a live cell imaging assay. A useful feature of the aryl azide probe is that it not only mediates photocrosslinking, but can also serve as a functional group handle for introduction of fluorescent probes. Though it is more common to derivatize *alkyl* azides with alkynes via 3+2 cycloaddition,[22,27,28] or phosphines via Staudinger ligation,[29,30] we and other groups have observed facile reaction of *aryl* azides with both cyclooctynes[31] and phosphines.[29,30,32] To make use of this reactivity, we expressed a LAP-CFP-TM cell surface construct[22] in HeLa cells, and performed extracellular labeling with W37V LplA, aryl azide, and ATP. After washing away excess aryl azide, we derivatized the LAP-ligated azide with cyclooctyne-Cy3 conjugate.[22] Imaging showed that only transfected (CFP-positive) cells became labeled with Cy3, while neighboring untransfected cells did not receive Cy3 (Figure 3). Negative controls with aryl azide omitted from the labeling reaction, or LAP-CFP-TM replaced by an alanine point mutant (in the LAP sequence) gave no fluorophore labeling (Figure 3). Thus, our three assays collectively demonstrate that aryl azide ligation has the same high sequence-specificity as wild-type LplA catalyzed lipoylation.

To demonstrate the utility of the new aryl azide ligase, we attempted to detect a known PPI within a complex mixture. FKBP is known to interact with FRB in the presence but not absence of the small molecule rapamycin.[33,34] We prepared a FKBP-LAP construct and expressed it in HEK cells. In separate HEK cells, we expressed a CFP-FRB fusion. Total lysates were prepared from both samples, and they were mixed. We labeled the combined lysates with aryl azide **1**, ATP, and W37V LplA, before photocrosslinking with 300–360 nm light in the presence or absence of rapamycin. To analyze the samples, we ran them on SDS-PAGE gels, and detected CFP-FRB by in-gel CFP fluorescence (Figure 4). Only in the presence of rapamycin, with UV light applied, and LAP tag intact, did we observe a higher molecular weight band corresponding to the covalent FKBP-FRB heterodimer. Negative controls with either construct omitted gave no heterodimer band.

In summary, we have engineered an aryl azide ligase that brings us closer to the dream of routine PPI detection inside cells by photocrosslinking. Our ligase, generated by active site engineering of *E. coli* LplA, catalyzes the extremely sequence-specific covalent ligation of a fluorinated aryl azide probe onto LAP fusion proteins, expressed on live cell surfaces, or in total lysate. Our immediate goal is to extend this ligation reaction to the cytosol of living cells.

In addition to photocrosslinking, the aryl azide ligation reaction can also be used for fluorophore tagging, as previously demonstrated with an alkyl azide analog of lipoic acid. [22] The aryl azide undergoes facile reaction with cyclooctyne fluorophore conjugates on the surface of living cells.

This study demonstrates the malleability of the LplA active site. Remarkably, the small-molecule binding site can be made to accommodate a variety of probe structures, without sacrificing the extremely high peptide/protein specificity that makes this enzyme such a useful tool for site-specific labeling applications. We were surprised to find that a single amino acid mutation (W37V) opened up the LplA active site to a structure quite dissimilar to lipoic acid. Not all enlargements of the lipoate binding pocket had equivalent effects, however; W37A was a much poorer catalyst for aryl azide **1** than W37V. We were also not successful in finding any catalysts for aryl azide **2**, despite its smaller size. These lessons will guide us in our continued efforts to engineer LplA for biotechnological applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Tanaka Y, Bond MR, Kohler JJ. Mol. Biosyst 2008;4:473–480. [PubMed: 18493640]
2. Brunner J. Annu. Rev. Biochem 1993;62:483–514. [PubMed: 8352595]
3. Parrish JR, Gulyas KD, Finley RL Jr. Curr. Opin. Biotechnol 2006;17:387–393. [PubMed: 16806892]
4. Hu CD, Kerppola TK. Nat. Biotechnol 2003;21:539–545. [PubMed: 12692560]
5. Galarneau A, Primeau M, Trudeau LE, Michnick SW. Nat. Biotechnol 2002;20:619–622. [PubMed: 12042868]
6. Luker KE, Smith MC, Luker GD, Gammon ST, Piwnica-Worms H, Piwnica-Worms D. Proc. Natl. Acad. Sci. U. S. A 2004;101:12288–12293. [PubMed: 15284440]
7. Stefan E, Aquin S, Berger N, Landry CR, Nyfeler B, Bouvier M, Michnick SW. Proc. Natl. Acad. Sci. U. S. A 2007;104:16916–16921. [PubMed: 17942691]
8. Truong K, Ikura M. Curr. Opin. Struct. Biol 2001;11:573–578. [PubMed: 11785758]
9. van Crielinge W, Beyaert R. Biol. Proced. Online 1999;2:1–38. [PubMed: 12734586]
10. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG. J. Am. Chem. Soc 2002;124:9026–9027. [PubMed: 12148987]

11. Chin JW, Martin AB, King DS, Wang L, Schultz PG. *Proc. Natl. Acad. Sci. U. S. A* 2002;99:11020–11024. [PubMed: 12154230]
12. Tippmann EM, Liu W, Summerer D, Mack AV, Schultz PG. *Chembiochem* 2007;8:2210–2214. [PubMed: 18000916]
13. Hino N, Okazaki Y, Kobayashi T, Hayashi A, Sakamoto K, Yokoyama S. *Nat. Methods* 2005;2:201–206. [PubMed: 15782189]
14. Ye S, Kohrer C, Huber T, Kazmi M, Sachdev P, Yan EC, Bhagat A, RajBhandary UL, Sakmar TP. *J. Biol. Chem* 2008;283:1525–1533. [PubMed: 17993461]
15. Suchanek M, Radzikowska A, Thiele C. *Nat. Methods* 2005;2:261–267. [PubMed: 15782218]
16. Kirshenbaum K, Carrico IS, Tirrell DA. *Chembiochem* 2002;3:235–237. [PubMed: 11921403]
17. Liu W, Brock A, Chen S, Chen S, Schultz PG. *Nat. Methods* 2007;4:239–244. [PubMed: 17322890]
18. Wang L, Xie J, Schultz PG. *Annu. Rev. Biophys. Biomol. Struct* 2006;35:225–249. [PubMed: 16689635]
19. Lu T, Ting AY, Mainland J, Jan LY, Schultz PG, Yang J. *Nat. Neuroscience* 2001;4:239–246.
20. Green DE, Morris TW, Green J, Cronan JE Jr, Guest JR. *Biochem. J* 1995;309:853–862. [PubMed: 7639702]
21. Ali ST, Guest JR. *Biochem. J* 1990;271:139–145. [PubMed: 2121129]
22. Fernández-Suárez M, Baruah H, Martínez-Hernández L, Xie KT, Baskin JM, Bertozzi CR, Ting AY. *Nat. Biotechnol* 2007;25:1483–1487. [PubMed: 18059260]
23. Fujiwara K, Toma S, Okamura-Ikeda K, Motokawa Y, Nakagawa A, Taniguchi H. *J. Biol. Chem* 2005;280:33645–33651. [PubMed: 16043486]
24. Kim DJ, Kim KH, Lee HH, Lee SJ, Ha JY, Yoon HJ, Suh SW. *J. Biol. Chem* 2005;280:38081–38089. [PubMed: 16141198]
25. Vocadlo DJ, Hang HC, Kim EJ, Hanover JA, Bertozzi CR. *Proc. Natl. Acad. Sci. U. S. A* 2003;100:9116–9121. [PubMed: 12874386]
26. Chapman-Smith A, Turner DL, Cronan JE Jr, Morris TW, Wallace JC. *Biochem. J* 1994;302(Pt 3): 881–887. [PubMed: 7945216]
27. Agard NJ, Prescher JA, Bertozzi CR. *J. Am. Chem. Soc* 2004;126:15046–15047. [PubMed: 15547999]
28. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. *ACS Chem. Biol* 2006;1:644–648. [PubMed: 17175580]
29. Lin FL, Hoyt HM, van Halbeek H, Bergman RG, Bertozzi CR. *J. Am. Chem. Soc* 2005;127:2686–2695. [PubMed: 15725026]
30. Luchansky SJ, Goon S, Bertozzi CR. *Chembiochem* 2004;5:371–374. [PubMed: 14997530]
31. Wittig G, Krebs A. *Chem. Ber* 1961;94:3260–3275.
32. Restituyo JA, Comstock LR, Petersen SG, Stringfellow T, Rajski SR. *Org. Lett* 2003;5:4357–4360. [PubMed: 14601999]
33. Banaszynski LA, Liu CW, Wandless TJ. *J. Am. Chem. Soc* 2005;127:4715–4721. [PubMed: 15796538]
34. Choi J, Chen J, Schreiber SL, Clardy J. *Science* 1996;273:239–242. [PubMed: 8662507]

Figure 1. Photoaffinity probes and photocrosslinking scheme. (a) Structures of aryl azide probes **1** and **2**, and lipoic acid. (b) Scheme illustrating site-specific aryl azide **1** ligation to LAP fusion protein, followed by photocrosslinking to interacting protein. LplA is lipoic acid ligase, and LAP is the LplA acceptor peptide.

Figure 2.

Engineering the aryl azide ligase and characterization of specificity. (a) Second-generation LplA mutants screened for aryl azide **1** ligation activity. Wild-type LplA and 6 mutants were compared under identical reaction conditions (800 μ M aryl azide **1**, 400 nM enzyme, 30 °C for 30 minutes), and product (aryl azide-LAP-HP1 conjugate) was detected by HPLC. The % conversion to product obtained with W37V LplA was normalized to 100% and other conversion values are reported relative to this. All measurements were performed in triplicate. (b) HPLC traces showing W37V LplA-catalyzed ligation of aryl azide **1** to LAP-HP1 protein. Negative controls with ATP omitted or W37V replaced by wild-type LplA are shown. The two starred peaks were purified and analyzed by mass-spectrometry in Supporting Figure 4. (c) Aryl azide ligation performed on mammalian cell lysate, to test labeling specificity. Lysate was generated from HEK cells expressing LAP-CFP (37 kD). Lysate was labeled *in vitro* with LplA and aryl azide **1**, then incubated with phosphine-FLAG to derivatize the aryl azide. Blotting with anti-FLAG antibody is shown to the right of the Coomassie-stained gel. Negative

controls are shown with an alanine mutation in LAP (lane 2), wild-type LplA instead of W37V (lane 4), and untransfected cells (lane 3).

Figure 3.

Specific fluorophore labeling on live cells via aryl azide **1** ligation. HeLa expressing LAP-CFP-TM were labeled with W37V LplA and aryl azide **1** for 40 minutes, followed by cyclooctyne-Cy3 conjugate[22] for 20 minutes, to selectively derivatize the aryl azide. Imaging shows pink Cy3 staining on the membranes of transfected CFP-positive cells. Negative controls are shown with omission of aryl azide (middle row) and LAP-CFP-TM replaced by an alanine mutant (bottom row). CFP and Cy3 images are merged. DIC is differential interference contrast.

Figure 4.

Aryl azide-mediated photocrosslinking of FKBP and FRB in mammalian cell lysate. Lysate containing FKBP-LAP and CFP-FRB was labeled with W37V LpIA and aryl azide **1**. Photocrosslinking was then performed with 300–350 nm light for 4 minutes, in the presence (lane 3) or absence (lane 4) of rapamycin. The samples were run on SDS-PAGE and crosslinked heterodimer was detected by CFP imaging (right gel). Negative controls are shown with no UV irradiation (lane 1), an alanine mutation in FKBP-LAP (lane 2), FKBP-LAP not expressed (lane 5), or CFP-FRB not expressed (lane 6). Coomassie visualization of the same samples (left gel) shows equal loading in all lanes. Note that proteins run differently in the CFP gel, because they were not denatured, in order to preserve CFP fluorescence.