Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Supporting Information

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, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Supporting Figure S1: Syntheses of aryl azide probes **1** and **2**.

Supporting Figure S2: LplA active sites from *T. acidophilum* and *E. coli.*

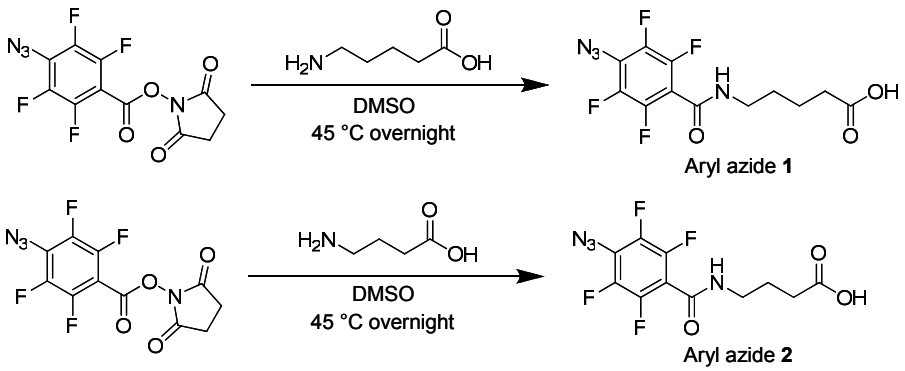
Supporting Figure S3: Comparison of aryl azide **1** ligation efficiencies of first-generation LplA mutants

Supporting Figure S4: Mass spectrum of the product of aryl azide **1** ligation to LAP-HP1

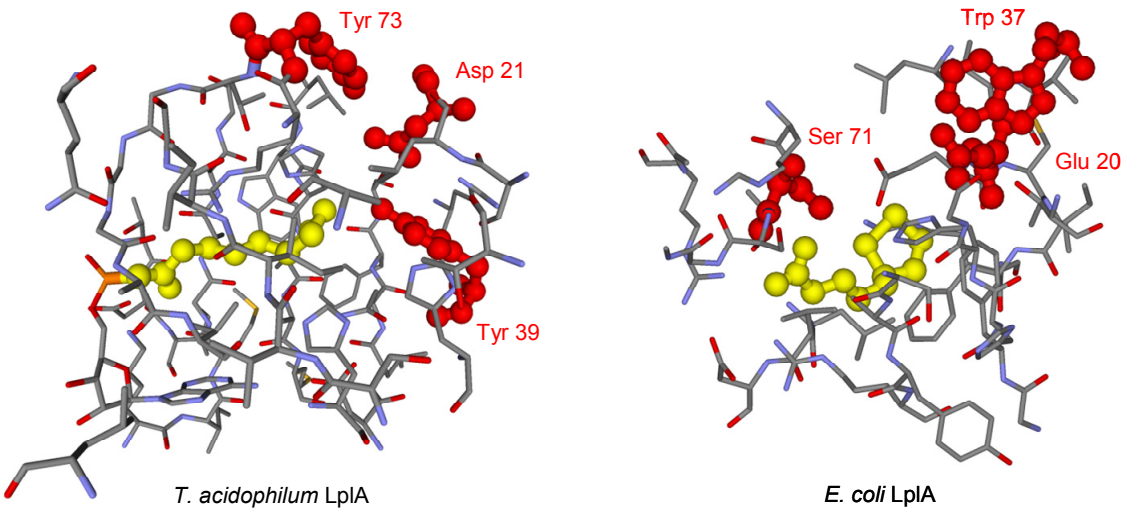
Supporting Figure S5: Measurement of aryl azide **1** ligation kinetics

Methods

References

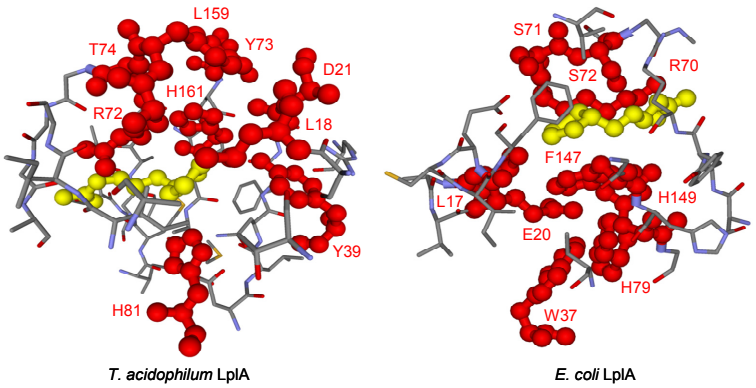


Supporting Figure 1. Syntheses of aryl azide probes **1** and **2**.

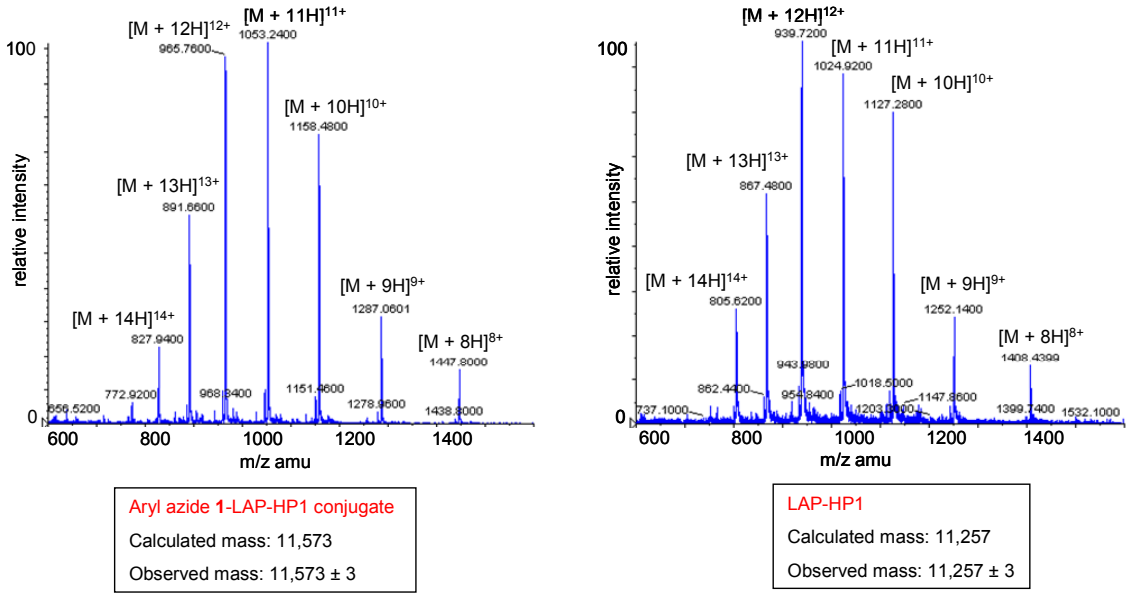


Supporting Figure 2. LplA active sites from *T. acidophilum* (left, PDB ID 2ART) and *E. coli* (right, PDB ID 1X2H). Lipoic acid is colored yellow. Red residues in the *E. coli* structure were targeted for mutagenesis in our second round of engineering. The corresponding residues in the *T. acidophilum* structure are also colored red.

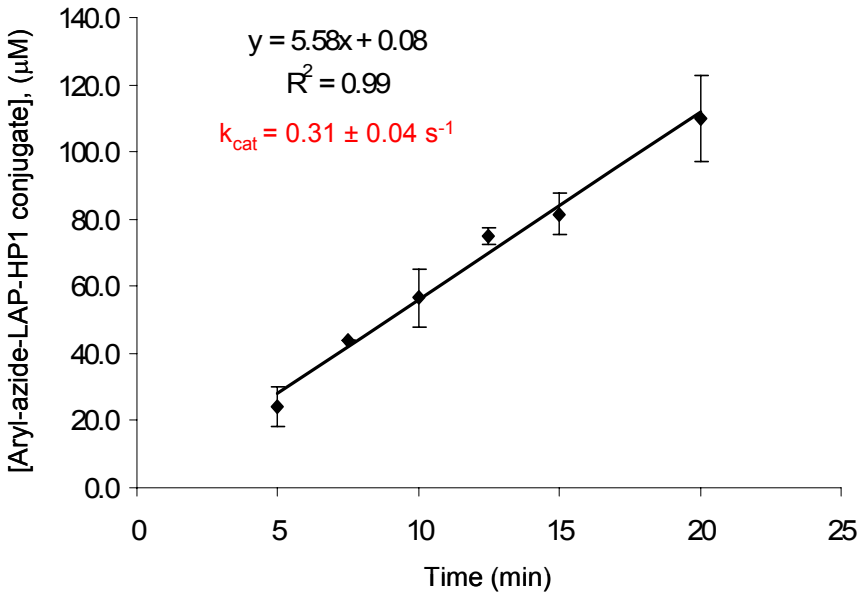
LplA mutant	% Conversion
W37A	100
E20A	66 ± 7
S71A	< 5
L17A	< 5
R70A	< 5
S72A	< 5
H79A	< 5
F147A	< 5
H149A	< 5



Supporting Figure 3. First-generation LplA mutants screened for aryl azide **1** ligation activity. Nine mutants were compared under identical reaction conditions (800 μM aryl azide **1**, 3 μM enzyme, 30 °C for 3 hours), and product (aryl azide-LAP-HP1 conjugate) was detected by HPLC. The % conversion to product obtained with W37A LplA was normalized to 100% and other conversion values are reported relative to this. On the right are structures of the *T. acidophilum* (PDB ID 2ART) and *E. coli* (PDB ID 1X2H) LplA active sites, with bound lipoic acid colored yellow. The 9 positions targeted for mutagenesis in the first round of engineering are colored red in the *E. coli* structure. The corresponding residues are also colored red in the *T. acidophilum* structure.



Supporting Figure 4. Mass spectra of the product of aryl azide **1** ligation to LAP-HP1 (left), and unmodified LAP-HP1 (right). These samples were obtained from the HPLC traces shown in Figure 2b.



Supporting Figure 5. Measurement of aryl azide **1** ligation kinetics. k_{cat} was determined with 300 nM W37V LplA, 720 μM LAP-HP1, and 800 μM aryl azide **1** in sodium phosphate buffer at pH 7.0. Product conjugate was detected by HPLC. All measurements were performed in triplicate.

Methods

General synthetic methods

Reagents were purchased from Alfa Aesar and Invitrogen and used without further purification. Analytical thin layer chromatography was performed using 0.25 mm silica gel 60 F_{254} plates. Flash column chromatography was carried out using silica gel (ICN SiliTech 32-63D). Mass spectra were recorded on an Applied Biosystems 200 QTRAP Mass Spectrometer using electrospray ionization. HPLC was performed on a Varian Prostar Instrument equipped with an autosampler, using a reverse-phase 250 × 4.6 mm Microsorb-MV 100 C18 column. Chromatograms were recorded at 210 nm. ^{1}H NMR spectra were recorded on a Varian Mercury 300 MHz instrument. Chemical shifts are reported in delta (δ) units, parts per million (ppm), and referenced to the residual solvent peak. Coupling constants (*J*) are reported in hertz (Hz). The following abbreviations for multiplets are used: s, singlet; dt, doublet of triplets; t, triplet; m, multiplet. Probes were stored as 100 mM stock solutions at –20 °C in DMSO.

*Synthesis of aryl azide (**1**)*: To a solution of 5-aminovaleric acid (12 mg, 100 μmol) in dry DMSO (250 μL) was added N-succinimidyl 4-azido-2,3,5,6-tetrafluorobenzoate (25 mg, 75 μmol; Invitrogen). The reaction was allowed to proceed overnight at 45 °C in the dark. The mixture was then acidified with 0.5 N HCl (250 μL) and extracted with ethyl acetate (400 μL) three times. The combined extracts were concentrated *in vacuo* and purified by silica chromatography (10% methanol in ethyl acetate) to afford **1** as a white powder. Further recrystallization in ethyl acetate-methanol mixture yielded **1** as white needles (3 mg, 9 μmol, 12%). TLC: R_f = 0.5 (9: 1 ethyl acetate-methanol). ^{1}H-NMR (300MHz, $(CD_3)_2SO$): δ 12.03 (s, 1H), 8.89 (t, 1H, *J*= 5.4), 3.23 (m, 2H, *J*=6.3), 2.22 (t, 2H, *J*= 7.2), 1.497 (m, 4H). ESI-MS: calculated for $[M-H]^-$ 333.07; observed 333.18.

*Synthesis of aryl azide (**2**)*: To a solution of 4-aminobutanoic acid (10.3 mg, 100 μmol) in dry DMSO (250 μL) was added N-succinimidyl 4-azido-2,3,5,6-tetrafluorobenzoate (25 mg, 75 μmol; Invitrogen). The reaction was allowed to proceed overnight at 45 °C in the dark. The mixture was then acidified with 0.5 N HCl (250 μL) and extracted with ethyl acetate (400 μL) three times. The combined extracts were concentrated *in vacuo* and purified by silica chromatography (10% methanol in ethyl acetate) to afford **2** as a white powder. ESI-MS: calculated for $[M-H]^-$ 319.05; observed 319.08.

Comparison of aryl azide ligation efficiencies for second-generation LplA mutants (Figure 2a)

To compare the rates of aryl azide **1** ligation to LAP-HP1, as catalyzed by different LplA mutants, reactions were assembled as follows: 400 nM LplA mutant, 625 μM LAP-HP1, 800 μM aryl azide **1**, 2 mM ATP, and 4 mM magnesium acetate in 25 mM sodium phosphate, pH 7.0. LplA and LAP-HP1 proteins were expressed and purified as previously described.[1] Ligation reactions were incubated at 30 °C for 30 minutes, then quenched with 100 mM EDTA (final concentration). Reactions were analyzed by C18 reverse-phase HPLC using a gradient of 30–45% acetonitrile in water with 0.1% trifluoroacetic acid over 20 minutes with a 1 mL/minute flow rate. LAP-HP1 had a

retention time of 9 minutes, but after conjugation to aryl azide **1**, the retention time increased to 16 minutes. All measurements were performed in triplicate. The extent of conversion of LAP-HP1 from the unmodified form to the modified form was calculated from the ratio of LAP-HP1-aryl azide peak area to the sum of (LAP-HP1 + LAP-HP1-aryl azide) peak areas. Conversion by the best mutant, LplA W37V, was normalized to 100%, and other conversions were reported relative to this.

HPLC assay for probe ligation to LAP-HP1 (Figure 2b)
Enzymatic reactions for this figure were assembled as follows: 2 μM LplA mutant, 200 μM LAP-HP1, 800 μM aryl azide **1**, 2 mM ATP, and 4 mM magnesium acetate in 25 mM sodium phosphate, pH 7.0. For the negative controls, ATP was omitted from the reaction, or wild-type LplA was substituted for W37V LplA. Reactions were incubated at 30 °C for 6 hours, then quenched with 100 mM EDTA (final concentration) and analyzed by HPLC as described above.

Testing ligation specificity on mammalian cell lysate (Figure 2c)
HEK 293T cells expressing the LAP-CFP or LAP(Ala)-CFP gene[1] were lysed under hypotonic conditions in 1 mM HEPES pH 7.5, 5 mM magnesium chloride, 1 mM phenylmethylsulphonyl fluoride and protease inhibitor cocktail. After 10 minute incubation at 4 °C, cells were vortexed for 2 minutes at room temperature. Crude cell lysate was cleared by centrifugation and stored at −80 °C. The expression levels of LAP-CFP and LAP(Ala)-CFP were estimated by measuring the fluorescence of CFP. Lysate labeling was performed under the following conditions: 1 μM LplA mutant, 200 μM aryl azide **1**, 4 mM ATP, 5 mM magnesium acetate, 25 mM sodium phosphate pH 7.0, for 5 hours at 30 °C. To visualize ligated aryl azide, Staudinger ligation was performed by addition of FLAG-phosphine[2] to a final concentration of 700 μM for 12 hours at 30 °C. The labeled lysate was run on 12 % SDS PAGE gel, then blotted with anti-FLAG(M2)-peroxidase antibody conjugate (Sigma-Aldrich, 1: 2000 dilution).

Live cell labeling with aryl azide and cyclooctyne-Cy3 (Figure 3)
HeLa cells were transfected with the LAP-CFP-TM expression plasmid[1] using Lipofectamine 2000. After 24 hours, the cells were washed twice with Dulbecco's Phosphate-Buffered Saline (DPBS) pH 7.4. Enzymatic ligation of aryl azide **1** was performed in DPBS with 10 μM W37V LplA, 1 mM aryl azide **1**, 1.5 mM ATP and 5 mM magnesium acetate for 40 min at 30 ºC. Cells were rinsed three times with DPBS, then incubated for 20 minutes at room temperature with 400 μM cyclooctyne-Cy3.[1,3] Cells were rinsed three times with cold DPBS before imaging in the same buffer using a 63X oil-immersion lens on a Zeiss Axio-Observer Z1 inverted epifluorescence microscope as previously described.[1]

Photocrosslinking of FKBP and FRB in mammalian cell lysate (Figure 4)
HEK 293T cells expressing FKBP-LAP, FKBP-LAP(Ala) or CFP-FRB were lysed under hypotonic conditions in 1 mM HEPES pH 7.5, 5 mM magnesium chloride, 1 mM phenylmethylsulphonyl fluoride and protease inhibitor cocktail. Cells were frozen and thawed three times, before vigorous vortexing for 2 minutes at room temperature. Crude lysate was then cleared by centrifugation and stored at -80 °C. To perform labeling with

aryl azide **1**, lysate containing FKBP-LAP and CFP-FRB were combined with 2 μM LplA enzyme, 200 μM aryl azide **1**, 4 mM ATP, and 5 mM magnesium acetate in 25 mM sodium phosphate pH 7.0, for 5 hours at 30 °C. To confirm ligation, we performed Staudinger ligation as described above with FLAG phosphine. To some reactions 1 μM rapamycin was added. Photocrosslinking was performed on 40 μL samples, placed 1.5 inches from a 800 W Hanovia UV lamp separated by Pyrex glass to filter out light <300 nm. After irradiation for 4 minutes at room temperature, we analyzed the samples by Coomassie stain and in-gel fluorescence visualization. Fluorescence was visualized on a Storm 860 instrument (Amhersham).

Comparison of aryl azide ligation efficiencies for first-generation LplA mutants (Supporting Figure 3)

To compare the rates of aryl azide **1** ligation to LAP-HP1, as catalyzed by different LplA mutants, reactions were assembled as follows: 3 μM LplA mutant, 625 μM LAP-HP1, 800 μM aryl azide **1**, 2 mM ATP, and 4 mM magnesium acetate in 25 mM sodium phosphate, pH 7.0. Ligation reactions were incubated at 30 °C for 3 hours, then quenched with 100 mM EDTA (final concentration). Reactions were analyzed by C18 reverse-phase HPLC as mentioned above. All measurements were performed in triplicate. The extent of conversion of LAP-HP1 from the unmodified form to the modified form was calculated from the ratio of LAP-HP1-aryl azide peak area to the sum of (LAP-HP1 + LAP-HP1-aryl azide) peak areas. Conversion by the best first generation mutant, LplA W37A, was normalized to 100%, and other conversions were reported relative to this.

Mass spectrometric detection of aryl azide-protein conjugate (Supporting Figure 4)

Reactions were performed as described above for Figure 2b. We collected the starred peaks manually and injected them onto the mass spectrometer at a flow rate of 10 μL/min. Mass spectra were recorded under the positive-enhanced multi-charge mode.

Kinetic analysis of aryl azide ligation onto LAP-HP1 (Supporting Figure 5)

To measure the ligation k_{cat}, 300 nM W37V LplA was incubated with 720 μM LAP-HP1, 800 μM aryl azide **1**, 2 mM ATP, and 4 mM magnesium acetate in 25 mM sodium phosphate pH 7.0. ATP was added last to initiate the reaction, then the mixture was incubated at 30 °C. 10 μL aliquots were removed at various timepoints and quenched with 100 mM EDTA (final concentration). Samples were analyzed using HPLC as described above. All measurements were performed in triplicate. A calibration curve was prepared, to relate the ratio of the integrated HPLC peak areas (LAP-HP1: LAP-HP1-aryl azide **1**) to the actual protein ratio, in order to compensate for differences in the extinction coefficients of LAP-HP1 and LAP-HP1-aryl azide **1**. The amount of product obtained at each time point was plotted against time, to obtain the V_{max} of the reaction. From this, the k_{cat} value was extracted using the equation $V_{max} = k_{cat}\,[E]_{total}$.

Estimation of K_m for aryl azide 1 (data not shown)

We could not obtain an accurate K_m for aryl azide **1**, because our HPLC detection limit prevented us from accurately measuring initial velocities when using low micromolar concentrations of aryl azide **1**. However, we were able to estimate the K_m, based on measurements of initial velocities at aryl azide **1** concentrations of 25, 50, 75, 100, and 200 μM. For these measurements, reactions were quenched after 5 minutes (during which

the reaction rate was still linear) using 100 mM EDTA. Samples were analyzed by HPLC as above. The plot of initial velocities (V_0) versus aryl azide **1** concentration was fit to the Michaelis-Menten equation ($V_0 = V_{max}$ [aryl azide **1**]/(K_m + [aryl azide **1**])), using Origin 7.0 software, to give an upper bound estimate for the K_m of 80 μM.

Site-directed mutagenesis of LplA

The wild-type LplA expression plasmid PYFJ16 (encoding the LplA gene with an N-terminal His_6 tag inside the pQE-2 vector from Qiagen) was a kind gift from John E. Cronan.[4] We performed site-directed mutagenesis of the LplA gene using the QuikChange kit (Stratagene), with the primers listed below (reverse complement sequences not shown).

LplA mutation	Forward primer sequence
W37V	5′-CAGCGCGTTCTGTTTCTCGTTCGCAATGCCGACACGGTA
W37S	5′-CAGCGCGTTCTGTTTCTCTCCCGCAATGCCGACACGGTA
W37G	5′-CAGCGCGTTCTGTTTCTCGGTCGCAATGCCGACACGGTA
W37A	5′-GCGCGTTCTGTTTCTCGCTCGCAATGCCGACACGG
E20A	5′-GGTTTAACCTGGCGGTGGCAGAGTGTATTTTTCGC
S71A	5′-GCGCGGCGCGCTAGCGGTGGCGGC

Cloning of CFP-FRB and FKBP-LAP for mammalian cell expression

To clone CFP-FRB in pcDNA3, we first amplified the FRB gene using the primers 5′ CTCGGATCCGATGTATCCGTACG and 5′ TTTTGAATTCCTTTAGCCGTCTGCTGATGCCGGTACTTCCAGAACTGCTTTGTCG to introduce an N-terminal BamHI site and a C-terminal 10-amino acid linker (GSGSTSGSGK) followed by LAP (DEVLVEIETDKAVLEVPASADG) followed by an EcoRI site. The PCR product was digested with BamHI and EcoRI enzymes and ligated into similarly-digested pcDNA3 vector to obtain the FRB-LAP-pcDNA3 plasmid. We next amplified the FRB-LAP gene using the primers 5′ GACGTGTACAAGTCCTCCGGCGGTCCGTATCCGTACGACGTACCAG and 5′ GCGGAATTCCTTTAACCCTGTCCAGTCATGCCGTCTGCTGATGCCGGTAC to introduce an N-terminal BsrGI site and a C-terminal 5-amino acid linker (MTGQG) followed by an EcoRI site. The PCR product was digested with BsrGI and EcoRI and ligated into similarly-digested CFP-AP-pcDNA3[5] vector. To inactivate the LAP sequence, we used QuikChange (Stratagene) to mutate the lysine of LAP to alanine using the primer 5′ GTTGAAATCGAAACCGACGCCGCAGTTCTGGAAGTAC and its reverse complement.

To clone FKBP-LAP-pcDNA3 for mammalian cell expression, we PCR-amplified the FKBP gene using the primers 5′ CTCGGATCCGATGGAACAAAAAC and 5′ TTTTGAATTCCTTTAGCCGTCTGCTGATGCCGGTACTTCCAGAACTGCTTTGTCG to introduce an N-terminal BamHI site and a C-terminal 10-amino acid linker (GSGSTSGSGK) followed by a LAP sequence (DEVLVEIETDKAVLEVPASADG) and an EcoRI site. The PCR product was digested with BamHI and EcoRI and pasted into similarly-digested pcDNA3 vector to obtain the FKBP-LAP-pcDNA3. To prepare the Ala mutant, we use QuikChange with the primer 5′ GTTGAAATCGAAACCGACGCCGCAGTTCTGGAAGTAC and its reverse complement.

References

[1] M. Fernandez-Suarez, H. Baruah, L. Martinez-Hernandez, K. T. Xie, J. M. Baskin, C. R. Bertozzi, A. Y. Ting, *Nat. Biotechnol.* **2007,** *25,* 1483-1487.

[2] D. J. Vocadlo, H. C. Hang, E. J. Kim, J. A. Hanover, C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A* **2003,** *100,* 9116-9121.

[3] N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo, C. R. Bertozzi, *ACS Chem. Biol.* **2006,** *1,* 644-648.

[4] Y. Jiang, J. E. Cronan, *J. Biol. Chem.* **2005,** *280,* 2244-2256.

[5] I. Chen, M. Howarth, W. Lin, A. Y. Ting, *Nat. Methods* **2005,** *2,* 99-104.