GAVIN PUBLISHERS

Advances in Biochemistry and Biotechnology

Rouch A, et al. Adv Biochem Biotechnol 2: 110.

DOI: 10.29011/2574-7258.000010

Research Article

On the Acellular use of Lipoic Acid Ligase for Labeling Proteins

Anne Rouch, Benjamin Fould, Elodie Jeantet, Estelle Marcheteau, Mathias Antoine, Olivier Nosjean, Gilles Ferry, Jean A. Boutin*

Département de Chimie des Protéines, Pôle d'expertise Biotechnologie, Chimie & Biologie, Institut de Recherches Servier, 125 Chemin de Ronde, 78290 Croissy-sur-Seine, France

***Corresponding author:** Dr. JA Boutin, Département de Chimie des Protéines, Pôle d'expertise Biotechnologie, Chimie & Biologie, Institut de Recherches Servier, 125 Chemin de Ronde, 78290 Croissy-sur-Seine, France, Tel: +33155722748; Fax: +33155722810; Email : jean.boutin@servier.com

Citation: Rouch A, Fould B, Jeantet E, Marcheteau E, Antoine M et al. (2017) On the Acellular use of Lipoic Acid Ligase for Labeling Proteins. Advan Biochem Biotechnol 2: 110. DOI: 10.29011/2574-7258.000010

Received Date: 29 December, 2016; **Accepted Date:** 09 January, 2017; **Published Date:** 14 January, 2017

Abstract

Protein-specific labeling is becoming increasingly necessary in biology and pharmacology because it will help understand the role of these proteins in disease-related systems. Several technologies are available, ranging from enzymatic, specific acylation to chemically less-specific techniques. We chose to systematize the use of an *E. coli* ligase described by the Ting laboratory and called lipoid acid ligase A. This ligase and its mutants can recognize a small peptide sequence called LAP2 (Lipoic Acid Acceptor Peptide: H- GFEIDKVWYDLDA-NH2) and acylate the lysine of this peptide with different analogues of lipoic acid, depending on the mutant. We sought to assess independently the reactions of this ligase and its reported mutants and to use this method with protein substrates in vitro to qualify the conditions under which such reactions could be adapted to cellular conditions. The present paper deals with the acellular conditions of the ligase reactions (including its mutants) towards the peptide and towards two model proteins (rhoA and thioredoxin) fused with the peptide sequence recognized by the ligase. A limited acellular comparison was made between a chemical modification of thioredoxin and an enzymatic modification using the present system. It is obvious from our results that the enzymatic technique led to purer protein in which the acylation was perfectly achieved and localized.

Introduction

The capacity to catalyze or favor posttranslational modifications of proteins in the most precise possible site-specific manner holds enormous potential in biomedical research [1-4]. The inclusion of exogenous amino acids in a protein is one way to introduce new modifications of the molecules. Methods based on the modification of tRNA [2] or the chemical synthesis of proteins [3], as we have previously described [5], involve the problem of delivering the protein into a workable cellular model. An alternative is to develop a cell in which some exogenous components are added, such as ligases or other protein-modifying enzymes, as in the approach described in the last decade by Ting and colleagues [6] and by others [7].

Protein modification is a key step in cellular homeostasis. Few proteins remain unmodified once translated, and numerous enzymes catalyze these multiple modifications, such as kinases, transferases (farnesyl [8], myristoyl [9], methyl [10], lipids [11]), ligases (ubiquitin [12], SUMO [13]), and oxidative modifications [14]. The quest for inhibitors of these enzymes has been a rich area of research in recent decades, and some of them have turned out to be major players in therapeutics.

At the other end of the spectrum, another kind of protein modifier has broadly emerged from the field of non-ribosomal peptides: the ligases. Indeed, the synthesis-often by microorganisms-of peptides or pseudo peptides, cyclic or not, requires a series of tweaks because these peptides are not directly encoded by the host genetic material but rather are synthesized by a series of enzymes with a complexity that we have only begun to understand (see Marahiel et al. [15], Mootz et al. [16] for such examples). These processes were long thought to be a source of a rich diversity in terms of enzymes for production or research and were the basis for one of the most active areas of synthetic biology. Among these enzymes, some of the most fascinating can ligate small natural molecules to, for example, a lysine side chain, as long as the lysine is part of a sequence that the ligase recognizes. The most representative of the category is certainly the E. coli ligase that

Ting's group thoroughly identified and characterized, designating it as lipoic acid ligase (LplA; see Puthenveetil et al., and references therein [17]. Among the features of this enzyme are that it is expressed in a cell without noticeable problem(s) and that mutants can be designed to accept alternative co- substrate(s) in place of the natural lipoic acid. For example, N3-bearing compounds can be linked to the lysine via the ligase-catalyzed reaction [18]. The next step is to introduce, for instance, a fluorophore by click chemistry onto the azide group, labeling the given protein specifically in cellulo [19]. The chemistry complementing this approach is safely compatible within cellulo conditions [20]. The field of application is even larger than expected because the modifier introduced via click chemistry can be of almost any nature, including but not limited to radiolabeled chemicals [21], heavy atom-bearing compounds, or biotinyl derivatives [22].

Our goal here was to take the first steps towards the systematic use of an enzyme capable of modifying a given protein sequence with lipoic acid derivatives. Although the ligase LplA has been described by others, we thought that it was critical to revisit the original description and independently confirm and characterize this lipoic acid ligase. It is therefore important to understand that the present program led us to produce tools that could be well understood and mastered in vitro, including in the context of pure protein substrates. The ultimate steps were either to study the immobilization of the ligase for further modification of substrate proteins of interest or to express the whole system in cellulo to modify in situ those proteins with intracellular movements of interest. Here we extend the original observations by using pure, full-length proteins fused to sequences that the ligase recognizes. These steps are required to use the system in cellulo, which is our ultimate goal.

Materials and Methods

Chemicals

Reagents were purchased from Sigma-Aldrich, Alfa Aesar, Invitrogen, or GE Healthcare and used without further purification.

Expression and purification of His6-LplA wild type and mutants in *E. coli*

The production of wild-type (wt) His6-LplA, His6-LplA W37I, His6-LplA W37A, and His6-LplA E20A F147A H149G was achieved in the *E. coli* expression system (*E. coli* BL21 (DE3)) as described [19]. The bacterial cells were cultivated in LB medium with ampicillin (final concentration 100 µg/ml) until an absorbance at 600 nm of 0.6 was reached. The cells were then induced with 0.1 mM IPTG, and the temperature of the culture was reduced to 25°C. The bacterial cells were cultivated O/N in 2 L medium and harvested by centrifugation. The presence of target proteins in the soluble protein fraction was confirmed by SDS-PAGE.

Purification was performed on an AKTA Avant® system with a 1 ml His trap column (GE Healthcare). The sample was injected on a column equilibrated with buffer 1 (50 mM Tris, 300 mM NaCl, pH 7.8) and extensively washed with the same buffer. The column was then developed stepwise by 10 volumes of buffer 1 with 50 mM imidazole, then with 10 column volumes of buffer 1 with 1000 mM imidazole, and finally with 20 volumes of buffer 1 with 200 mM imidazole.

Fractions were analyzed by 12% SDS-PAGE gel electrophoresis followed by Coomassie staining. The fractions containing enzyme were pooled and dialyzed twice against a dialysis buffer (20 mM Tris, 1 mM DTT, 10% (V/V) glycerol, pH 7.5, or 20 mM HEPES, pH 8.0, 150 mM NaCl, 10% glycerol) at 4°C. The concentration of the purified proteins was determined spectrophotometrically at 280 nm using the calculated molar extinction coefficient of 42 180 M^{-1} cm^{-1} or 47 870 M^{-1} cm^{-1} for single mutants and triple mutants, respectively. The purity of recombinant proteins was estimated to exceed 95% based on the SDS-PAGE analyses.

Expression and purification of rhoA fused with LAP2 at the N-terminus, the C- terminus, or in a loop (His tag 1-180 /insertion D49-G50: [GGS-LAP2-GGS], F25N) in *E. coli*

Production of the three constructs of rhoA was achieved in the *E. coli* expression system (*E. coli* BL21 (DE3)). The bacterial cells were cultivated in LB medium with ampicillin (final concentration 100 µg/ml) until an absorbance at 600 nm of 0.8 was reached. The cells were induced with 0.1 mM IPTG, and the temperature of the culture was reduced to 22°C. The bacterial cells were cultivated O/N in 3.0 L medium and harvested by centrifugation. The presence of target proteins in the soluble protein fraction was confirmed by both SDS-PAGE and western blot analyses. The proteins were purified by affinity chromatography using NiNTA agarose according to the standard procedure provided by the manufacturer. A total of 20 mM Tris-HCl, pH 8, 250 mM NaCl, 10% glycerol, 10 mM $MgCl_2$, 50 µM GDP, and 250 mm imidazole, was used as elution buffer. To increase the purity of recombinant thioredoxin (Trx), the fractions collected from affinity chromatography column were loaded on a HiLoad 26/600 Superdex 200 pg column and eluted with 20 mM Tris-HCl, pH 8.0, 250 mm NaCl, 5% glycerol, 2 mm DTT, 1 mm $MgCl_2$, 50 µM GDP. After purification, the sample was concentrated, and the purity of recombinant proteins was estimated to exceed 95% based on the SDS-PAGE gel electrophoresis analyses.

Expression and purification of His6-TRX wt and His6-TRX LAP2 in *E. coli*

The production of the human Trx wt and LAP2 were done in the *E. coli* expression system (*E. coli* BL21 (DE3)). The bacterial cells were cultivated in LB medium with ampicillin (final concen-

tration 100 µg/mL) until an absorbance A600 of 0.8 was reached. The cells were induced with 0.1 mM IPTG, and the temperature of the culture was reduced to 25°C. The bacterial cells were cultivated O/N in 1.0 L medium and harvested by centrifugation. The presence of target proteins in the soluble protein fraction was confirmed by both SDS-PAGE gel electrophoresis and western blot analyses. The proteins were purified by affinity chromatography using NiNTA agarose according to the standard procedure provided by the manufacturer. A total of 50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 0.1 mM DTT, 250 mM imidazole was used as elution buffer. For increased purity of recombinant Trx the fractions collected from affinity chromatography column were loaded on a HiLoad 26/600 Superdex 200 pg column and eluted with 50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 5 mM DTT. After purification, the sample was concentrated and dialyzed against storage buffer (40 mM Tris- HCl, pH 7.5, 180 mM NaCl, 4 mM DTT, 10% glycerol). The purity of the recombinant proteins was estimated to exceed 95% based on the SDS-PAGE analyses.

Electrophoresis analyses

SDS-PAGE analyses were performed using Nu-PAGE 4-12% gels in MES or MOPS buffer (Invitrogen). For sample preparation under reducing conditions, samples were boiled for 5 min at 95°C with LDS sample buffer (Invitrogen) and 50 mM DTT. For non-reducing SDS-PAGE analysis, samples were first treated with 30 mM iodoacetamide for 2 min at 70°C. SeeBlue® Plus2 Pre-stained Protein Standard (Invitrogen) was used as a reference. Images were captured using a Chemi Doc XRS camera and Quantity One software (Bio- Rad).

Mass spectrometric analyses

Native mass spectrometry

Prior to Native ESI-MS measurements, LpIA and LAP2 were desalted in 150 mM ammonium acetate (pH 7.5) with Zeba Desalting Spin Columns (89849, Thermo Scientific) by three successive cycles. Proteins were then diluted to 10µM. Mass spectra were acquired in the positive V mode with a LCT "High Mass" mass spectrometer (MS Vision, Netherland) fittedwith an automated chip NanoESI system (Nanomate 200, Advion Biosciences, Ithaca, NY). Standard nanospray parameters were used (chip voltage 1700 V, gas pressure 0.75 psi). The mass spectrometer was carefully tuned with gentle desolvation parameters to preserve protein folding during the ionization/desorption process. The sample cone, extraction cone and RF Lens voltages were optimized to 260 V, 3 V, and 700 V, respectively.

The instrument was calibrated with Cesium Iodide diluted to water at 3 mg/ml. Data were acquired during two minutes and processed with MassLynx.

Denaturing liquid chromatography-mass spectrometry

Before LC-MS analysis, samples were diluted to 0.1 mg/ml in a vial with H_20. ESI-MS spectra were obtained in positive ion mode on a XevoG2XS QT of (Waters, Manchester, UK). Instrument settings were: source temperature 80°C, capillary voltage 2 kV, and cone voltage40 V. The instrument was calibrated with sodium iodide solution (2 µg/µl in 50/50 2- propanol/water). Protein samples were desalted on-line using a MassPrep™ Micro Desalting column (2.1x5mm). Column temperature was 80°C, flow rate 0.5ml/min. The following gradient was used:

Time	Flow (ml/min)	% water + 0.1% Acid Formic	% acetonitrile + 0.1% Acid Formic
Initial	0.5	95	5
0.5	0.5	95	5
0.51	0.2	95	5
2	0.2	10	90
2.1	0.5	95	5
2.7	0.5	10	90
2.8	0.5	95	5
3.4	0.5	10	90
4	0.5	95	5

MS spectra were acquired over the m/z range 500-4000. Data were processed with MassLynx.

Analytical Size-Exclusion Chromatography (SEC) with Multi-Angle Static Light Scattering (MALS)

Analyses were performed on an HPLC Agilent 1200 UV coupled with a MALS (DAWN 8+ Wyatt Technology) and a RI (Optilab T-rEX Wyatt Technology). For SEC, samples were centrifuged at 1400 rpm of 20 min before injection. Generally 20 µl of protein at approximately 1 mg.ml-1 was injected into the system. For all analyses, bovine serum albumin at 1 mg.ml-1was used as control.

Lipoic Acid Ligase

Proteins were analyzed on a KW Protein 803 column (Shodex). HPLC column, MALS and RI were equilibrated with a 20 mM Tris/HCl, 300 mM NaCl, pH 7.5, buffer at a flow rate of 1 ml.min-1, and the absorbance was recorded at 210 and 280 nm.

Rho A

These proteins were analyzed on a KW Protein 802.5 column (Shodex). HPLC column, MALS and RI were equilibrated with a 20 mM Tris/HCl, 300 mM NaCl, pH 7.4, buffer at a flow rate of 0.8 ml.min-1, and the absorbance was recorded at 210 and 280 nm.

Thioredoxine Analyses

These proteins were analyzed on a Superdex 75 5/150 GL column (GE Healthcare), HPLC column, MALS and RI were equilibrated with 50 mM Tris/HCl, 200 mM NaCl, 1 mM TCEP, pH 7.5, buffer at a flow rate of 0.2ml.min-1, and the absorbance was recorded at 280 nm and also at 210 and 650 nm.

Differential Scanning Fluorimetry

These analyses were performed on a CFX384 Real-Time PCR Detection System (Bio-Rad). Proteins were diluted at 0.1, 0.15, and 0.2 mg.ml^{-1}in analysis buffer and mixed with SYPRO® Orange Protein Gel Stain at 50X according to the manufacturer's protocol (Sigma). Samples were transferred into a 96-well plate (Hard-Shell® PCR Plates, Bio-Rad) and sealed with transparent foil (Microseal B, Bio-Rad). The plate was placed in the PCR instrument and the temperature scan run from 20 to 95°C, with a temperature ramp of 1°C min-1. Fluorescence was recorded (ex 483 nm/em 568 nm, FRET channel) as a function of temperature. For analyses, the melting temperature was defined as the minimum of the first derivative of the fluorescence intensity.

LplA-mediated enzymatic labeling of LAP2 peptide

The peptide LAP2 (H-GFEIDKVWYDLDA-NH_2) was synthesized by Genepep (St Jean de Védas, France). The peptide was solubilized at 2 mM in H_2O + 0.1% NH_4OH; lipoic acid (Fluka 04471) was solubilized at 5 mM in 50/50 H_2O/CH_3CN; 10-azidodecanoic acid and 7-azidoheptanoic acid (Piramal, Ahmedabad, India) were solubilized at 10 mM in CH_3CN; and LplA was produced according to the method described above. For a typical reaction, peptide LAP2 was added to the buffer (25 mM sodium phosphate, pH 8, 2 mM magnesium acetate, 1 mM ATP, 0.1% Tween 20) at 200 µM, lipoic acid or other substrate at 350 µM, and then LplA at 1 µM. The reaction was incubated at 37°C for 30 to 120 min and stopped with 300 mM EDTA, 50 mM Tris, pH 7.5, 30% CH_3CN. The reaction was analyzed by HPLC according to the method described below.

Determination of LAP2 Km and lipoic acid affinity constants

To measure the kinetic parameters Vmax, Kcat and KM for LplA ligation of the LAP2 peptide, LAP2 at 15, 25, 50, 75, 100, 200, 300, or 400 µM was mixed with 750 µM lipoic acid and 0.005 or 0.01 µM LplA, in 25 mM sodium phosphate, pH 8, 2 mM magnesium acetate, 1 mM ATP, 0.1% Tween 20. The reaction was incubated at 37°C and stopped after 30 min with 300 mm EDTA, 50 mm Tries, pH 7.5, 30% CH_3CN. The reaction was analyzed by the HPLC assay described below. Measurements were performed in duplicate. The amount of product obtained at each time point was plotted against time to obtain the initial velocity for each concentration of LAP2 peptide. The initial velocities (Vi) were plotted against LAP2 peptide concentration and fit to the Michaelis-Menten equation (Vi = Vmax [LAP2] / (KM + [LAP2])) using Prism software to obtain the Vmax and KM values for LAP2. Kcat is obtained according to the following equation: Kcat = Vmax / [Enz]. To measure the KM for LplA ligation of lipoic acid, LAP2 peptide at 200 µM was mixed with 10, 25, 50, 100, 300, 500, or 750 µM lipoic acid and 0.005, 0.01, or 0.025 µM LplA in 25 mm sodium phosphate, pH 8, 2 mM magnesium acetate, 1 mM ATP, 0.1% Tween 20. The conditions for measuring were as described above. The initial velocities (Vi) were then plotted against the lipoic acid concentration and fit to the Michaelis-Menten equation (Vi = Vmax [LAP2] / (KM + [lipoic acid])) using Prism software to obtain the Km for lipoic acid. Kcat is obtained according to the following equation: Kcat = Vmax / [Enz].

Reverse-phase chromatography assay for analysis of LAP2 labeling

HPLC assays were performed on an HPLC Agilent 1100 UV with buffer A corresponding to H_2O + 0.1% TFA and buffer B to CH_3CN + 0.1% TFA. Column (HPLC Chromolith Performance C18, 4.6 × 100 mm) was equilibrated at 2 ml.min-1with 75/25 buffer A/buffer B. Samples (5 µl) were injected using the following gradient: 0-1 min: 25% Buffer B; 1-9 min: 25-55% Buffer B; 9-10 min: 55% Buffer B; 10-11: 55-25% Buffer B; and 11-15 min: 25% Buffer B. The absorbance was recorded at 210 nm unless otherwise mentioned.

Reverse-phase chromatography coupled with mass spectrometric assay for analysis of LAP2 labeling

LC/MS assays were conducted using an HPLC Agilent 1200 UV equipped with a quadrupole 61400 with buffer A corresponding to H_2O + 0.1% TFA and buffer B to CH_3CN + 0.1% TFA. Column (XBridge peptide BEH C18, 2.1 × 100 mm) was equilibrated at 0.4 ml.min^{-1} with 75/25 buffer A/buffer B. Samples (4 µl) were injected following the same gradient as above. Mass spectra were recorded under the positive enhanced multi-charge mode of an ESI-MS. Absorbance was recorded at 210 and 280 nm, unless otherwise mentioned.

Reverse-phase chromatography assay for analysis of rhoA labeling

HPLC assays were performed on an HPLC Agilent 1100 UV with buffer a corresponding to H_2O + 0.1% TFA and buffer B to CH_3CN + 0.1% TFA. Column (HPLC Xbridge Protein BEH C4, 2.1 × 100 mm) was equilibrated at 0.3 ml.min^{-1} with 65/35 buffer A/buffer B. Samples (4 µl) were injected using the following gradient: 0-2 min: 35% Buffer B; 2-15 min: 35-65% Buffer B; 15-16 min: 65-35% Buffer B; and 16-20 min: 35% Buffer B. Unless otherwise mentioned, absorbance was recorded at 210 nm.

Enzymatic assay with Rho A LAP2 or Thioredoxine LAP2

RhoA LAP2, TRX LAP2, and LplA W37I were produced as described previously. 10- azidodecanoic acid (piramal) was solu-

bilized at 10 mM in CH_3CN and Alexa Fluor® 647 DIBO Alkyne (Molecular Probes) was solubilized at 10 mM in DMSO. For a classical reaction, LAP2 modified protein was added to the buffer (25 mM sodium phosphate, pH 8, 2 mM magnesium acetate, 1 mM ATP) at 10-20 µM, 10-azidodecanoic acid at 50 µM, and then LplA W37I at 1 µM. Reactions were incubated at 37°C for 60 to 120 min.

For the second step (click reaction), two strategies were possible according to the quantity of labeled protein. (1) For small batches (~0.2 mg of protein), the reaction was cooled to room temperature, and Alexa Fluor® 647 DIBO Alkyne was added directly (100-150 µM). The reaction was incubated at 25°C for 60 to 120 min. (2) For bigger batches (~2 mg of protein), the reaction was dialyzed twice against 50 mM Tris, 200 mM NaCl, 5 mM DTT, pH 7.5, to eliminate the unreacted substrate. The reaction mixture was concentrated (Ami con Ultra-4, PLBC, membrane Ultracel-3, 3 KD, Millipore) to approximately 70 µM. Alexa Fluor® 647 DIBO Alkyne was added (100 µM), and the reaction was incubated at 25°C for 60 to 120 min. In both cases, the reaction was stopped by injection in HPLC or by freezing.

Chemical labeling of Thioredoxin wt with NHS Alexa Fluor 647

Trx wt was dialyzed against PBS+1 mM TCEP because chemical labeling is not compatible with Tris content in storage buffer. Trx was then diluted to 156 µM (2 mg/ml) in PBS+1 mM TCEP and mixed with relevant equivalents of Alexa Fluor® 647 NHS Ester (10 mm in DMSO, Molecular Probes). The reaction was incubated at room temperature for 2 h and then stopped by addition of excess Tris salt.

Purification of labeled Thioredoxin-LAP2 by preparative SEC

LAP2-modified Trx labeled with 10-azidodecanoic acid and Alexa Fluor® 647 DIBO Alkyne or wt Trx labeled with NHS Alexa Fluor 647 were purified by preparative SEC using the AKTA Avant Gold Seal system. The reaction mixtures were injected on a Superdex 75 10/300 GL (GE Healthcare) equilibrated with 50 mM Tris/HCl, 200 mM NaCl, 1 mM TCEP, pH 7.5, buffer at 0.5 ml.min-1, and the absorbance was recorded at 210, 280, and 650 nm.

Results

Our goal was to clone, produce, purify, and characterize the ligases that Ting and colleagues largely used (see Uttamapinant et al [19] and references therein). Because our program included the use of synthetic co-substrates of LAP2, we generated mutated enzymes capable of recognizing those modified co-substrates (as described and partially characterized by Ting and colleagues [6, 23-25]).

Production and characterization of wt and mutated LplA

Using the sequence published by the original group [19], we had no major difficulty in obtaining bacteria overexpressing the transgene. Figure 1A shows the purification of the wt protein, as described in the material and methods section. It is clear that the purification process, by means of nickel affinity chromatography, went smoothly and led to several fractions of highly purified material. The two waves of elution were due to the increased concentrations of imidazole. Because the final analysis of the purified, pooled materials was cleaner in pool 3 (Figure 1B), we chose to use this material for the rest of the experiments. The same purification protocols were performed for the mutated enzymes and led to purified materials, as can be seen in Figure 1S. (supplementary figure).

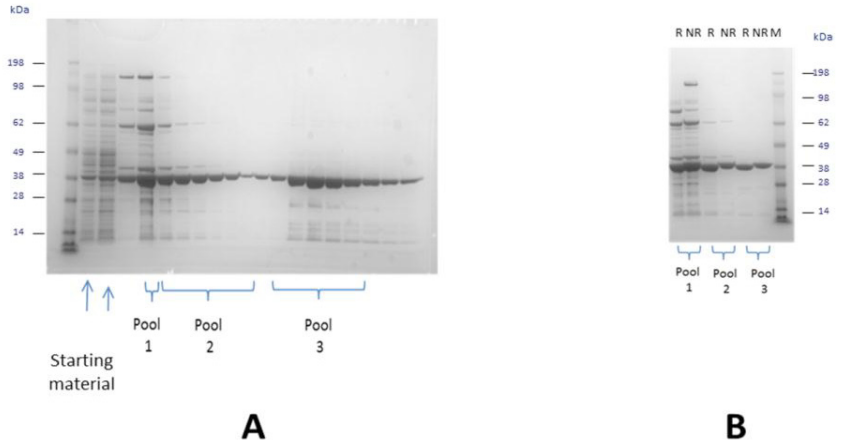


Figure 1: Purification steps of LplA wt enzyme.

A: SDS-PAGE analysis of the purification fractions of LplA. A nickel column was equilibrated and developed with three imidazole concentrations: 50 (Pool 1), 100 µM (Pool 2), and 200 µM (Pool 3). Fractions were independently combined into three pools. Analyses were done in reducing conditions.

B: SDS-PAGE analysis of the purified pools of the wt ligase. The three pools were individually dialyzed against a storage buffer (20 mM Tris base, 1 mM DTT, 10% (vol/vol) glycerol, pH 7.5). Analyses were performed in reducing and non-reducing conditions.

To qualify the purified proteins, we analyzed by SEC-MALS all the LplA enzymes, whether wild type or mutated. For example, in Figure 2A, the wt enzyme is shown to have been homogeneous and monomeric at 96% (P1); P2 represents oligomers (2%) and P3 aggregates (2%). Similar results were obtained for the other enzymes, particularly with the triply mutated lplA monomeric, which was pure at 99% (Figure 2D).

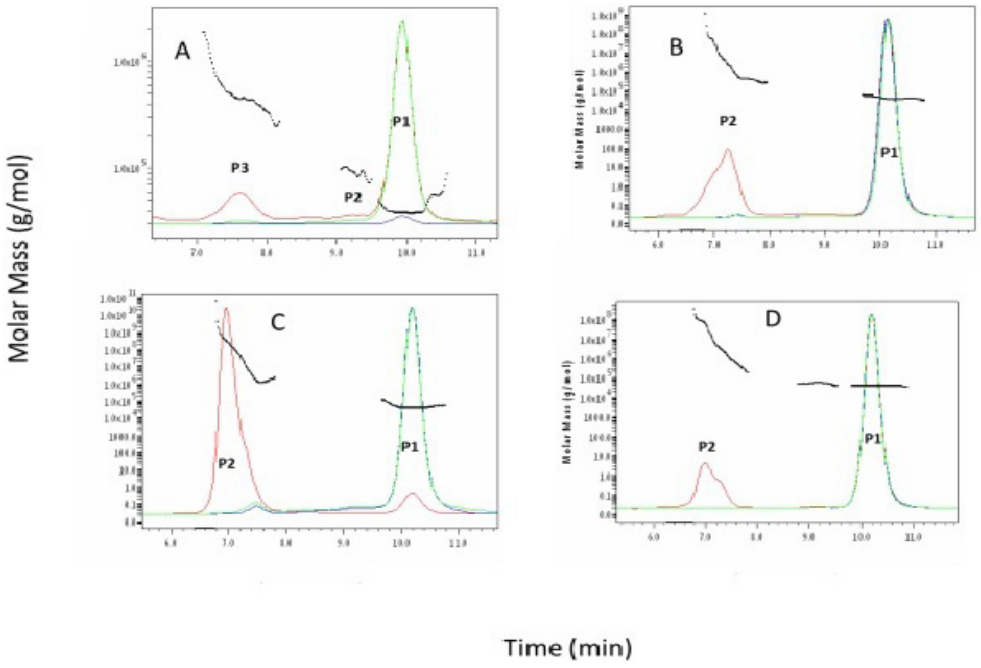


Figure 2: SEC-MALS analysis of LplA wild type, LplA W37I, LplA W37A, and LplA E20A_F147A_H149G enzymes.

A: LplA wild type: The enzyme was homogeneous and monomeric at 96% (P1); P2 represents oligomers (2%) and P3 aggregates (2%).

B: LplA W37I: The enzyme was homogeneous and monomeric at 99% (P1); P2 represents aggregates (1%).

C: LplA W37A: The enzyme was homogeneous and monomeric at 95% (P1); P2 represents aggregates (5%).

D: LplA LplA E20A_F147A_H149G: The enzyme was homogeneous and monomeric at 99% (P1); P2 represents aggregates (1%).

SEC-MALS analysis of wt LplA using a Shodex KW 803 column at 1 ml.min-1 with running buffer: Tris HCl 20 mM, NaCl 0.3 M, pH 7.4. The graphic shows UV signal (green), LS signal (red), RI signal (blue), and molecular weight (black) according to time. All signals are represented on a normalized scale.

We felt that plain characterization of pure proteins was necessary to confirm the usability of the enzyme. SDS-PAGE was therefore used to visualize the purity of the protein. Figure 3 shows the results of these experiments in native as well as reduced conditions for the wt enzyme (Figure 3A) and for the various mutants (Figure 3B). The gels clear indicated that the proteins were monomeric in their native state and almost pure without any significant amount of other proteins present in the preparations. It is also interesting to note that the stability of the proteins differed slightly from one mutant to the other despite the minimal nature of the mutations (one or three point mutations), as shown in Figure 3C. Indeed, while the Tm of the wt enzyme was 53°C, the W37A was 59°C, while the other mono-mutant dropped to 46°C, and the triple mutant was 45°C. We viewed this information as a warning concerning the possible stability of the enzyme at 37°C, the temperature usually used for enzymatic reactions. Related to this concern, it is interesting to point out the drastic differences between the mutant stabilities. A single mutation led to a drop of several degrees in the thermogram when compared to the wt enzyme. For LplA W37I, two Tm were observed: 46 and 49°C. This suggests that the enzyme was denatured in 2 times: one domain being more instable and denaturing at a lower temperature that the other domain.

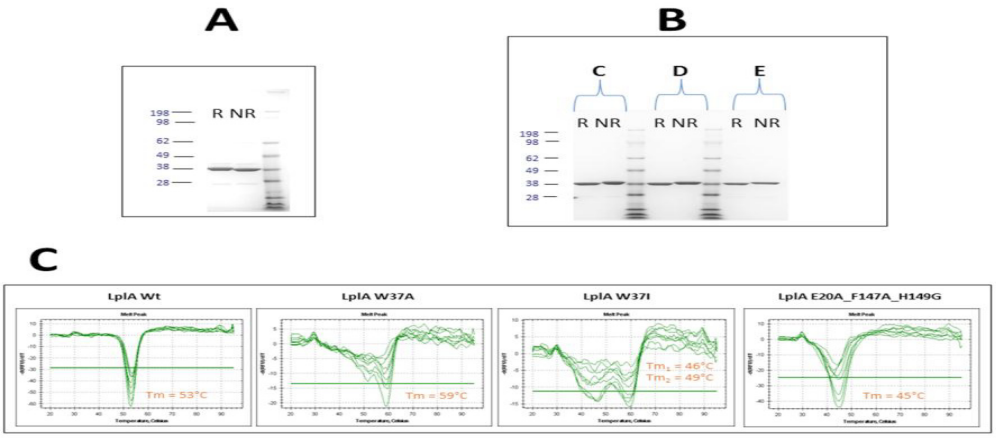


Figure 3: SDS-PAGE electrophoresis and thermal stability of LplA wild type (Wt), W37I, W37A, and E20A_F147A_ H149G enzymes.

SDS-PAGE electrophoresis of purified LplA wild type (A) and purified mutants (B) Mutated LplA were W37I (C), W37A (E), and E20A_F147A_H149G (F) enzymes, respectively. R represents reducing conditions; N R: non-reducing conditions; and M the SeeBlue® Plus2 Pre-stained protein standards. The analyses are in accordance with SEC-MALS analyses and theoretical molecular weights. Thermal stability of LplA Wt, W37A, W37I, and E20A_F147A_H149G were performed using differential scanning fluorimetry (C). Each sample was diluted in 25 mM sodium phosphate, pH 8.0, 2 mM magnesium acetate, 1 mM ATP, and analyzed at three concentrations (0.1, 0.15, and 0.2 mg.ml-1) in the presence of the SYPRO orange probe. Samples were slowly heated while the fluorescence emission was recorded. Curves show the fluorescence derivative according to temperature, and the melting temperature (Tm) is defined as the minimum of the curve.

Furthermore, we checked the four proteins by mass spectrometry analyses (Figure 4). For each construct, the theoretical mass was in accordance with experimental findings. As very often in the presence of the HAHA tag, a gluconoylation occurred in about 10% of the enzyme [26]. Overall, analytical the profiles of the enzymes were the same for the mutated LplAs and for the wild type (see table 1).

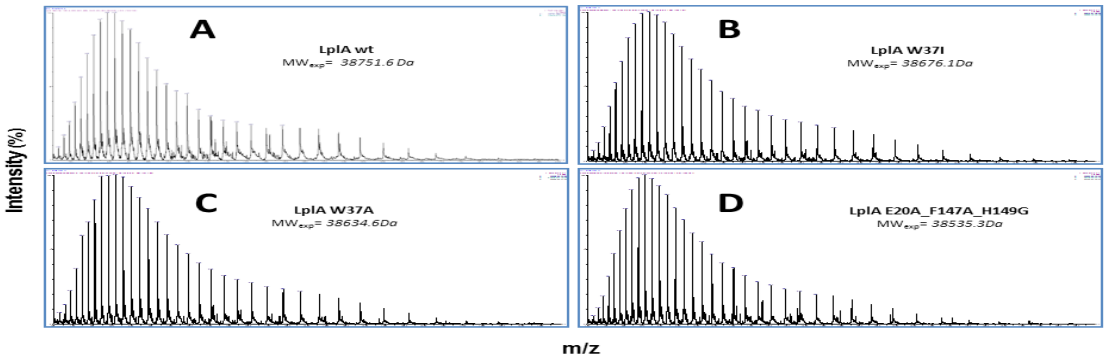


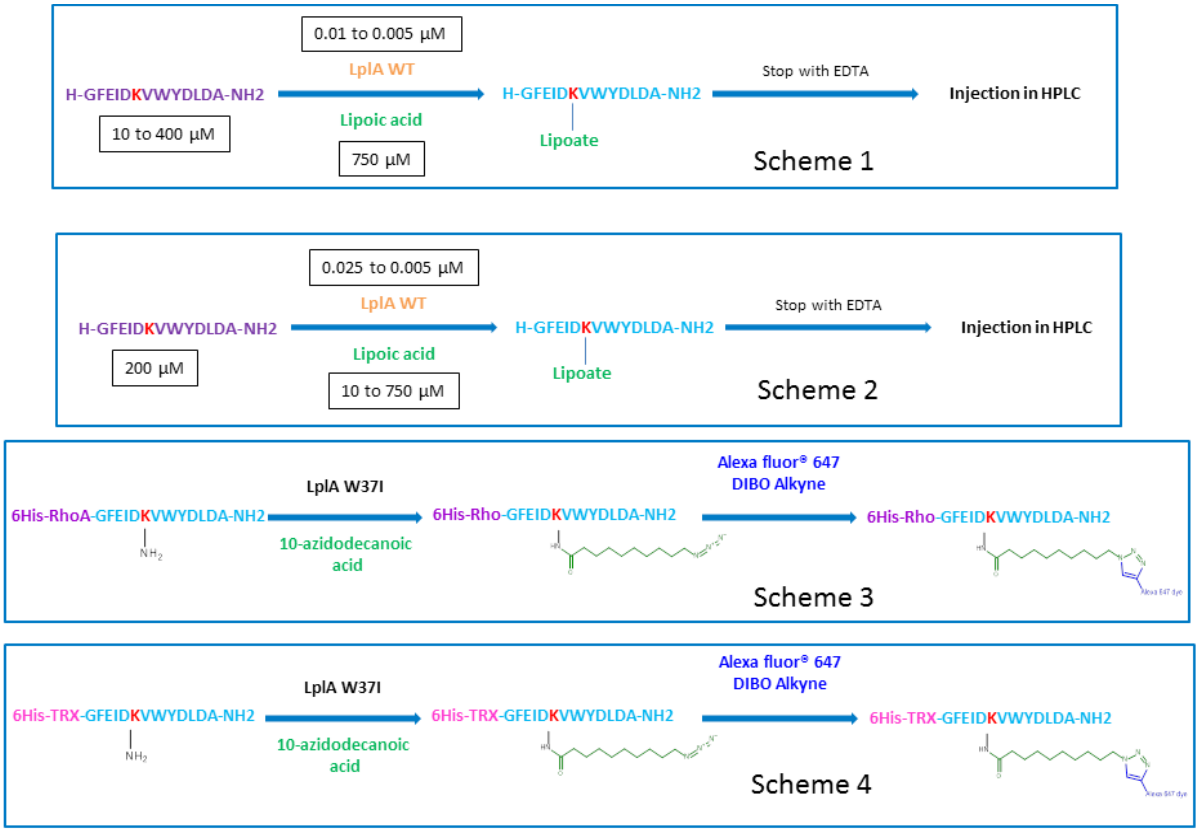
Figure 4: Characterisation of LplA.

Electrospray ionization mass spectrometry analyses of LplA Wt (A), W37I (B), W37A (C) and E20A_F147A_H49G (D) at 0.1mg/ml in 25mM sodium phosphate pH 8, 2mM magnesium acetate.

The experiments were conducted on a Synapt G2-S instrument (Waters). For each construct theoretical mass is in accordance with experiment and ~10% of enzyme is gluconoylated. The profile is the same for mutated LplA than for wt.

Ligation to LAP2 peptide with lipoic acid and its analogs

We decided to re-assess the conditions under which the enzymes worked. The team that originated the enzymes reported some of these features [6-17], but we wanted to complete these observations, particularly regarding the specificity vis-à-vis synthetic co- substrates. At first, we focused our attention on the wt enzyme, so a first step was a classical assessment of enzymatic activity using both a peptide as a substrate and lipoic acid as the natural co-substrate, as summarized on Scheme 1.



Schemes: HPLC assay for enzyme-catalyzed site-specific ligation of lipoic acid to LAP2 peptide.

Scheme 1 and 2: Enzyme -catalyzed site-specific ligation of lipoic acid to the lysine (in red) of LAP2 peptide (H-GFEIDKVWYDLDA-NH2). LAP2 is combined to lipoic acid and LplA wt in the presence of ATP as described in the experimental section. After incubation at 37°C, the reaction was stopped with EDTA and analyzed by reverse-phase chromatography (see HPLC assay in the experimental section). Note the changes in reactant concentrations between scheme 1 & 2.

Schemes 3 and 4: Enzyme-catalyzed site-specific ligation of lipoic acid to the lysine (in red) of LAP2 fused with the protein RhoA (Scheme 3) or Trx (Scheme 4). The fused protein was combined to 10- azidododecanoic acid and the mutant ligase LplA W37I in the presence of ATP as described in the experimental section. In a second step, the fluorophore Alexa Fluor 647 was introduced via a click chemistry reaction between the azido and cycloalkyne.

Figure 5A summarizes the results of the different steps in this process. To the best of our knowledge, a systematic survey of this activity previously has been only partially reported. The various steps of assay validation are shown. We thoroughly checked the conditions of the assay in a step-by-step approach. With HPLC measurement of the peak corresponding to the product of the reaction and by validating its molecular weight (879.9 Da) by mass spectrometry (Figure 5B), the nature of the enzymatically catalyzed reaction was ensured, including the absolute necessity for the presence of ATP in the reaction (see step 6, Figure 5A).

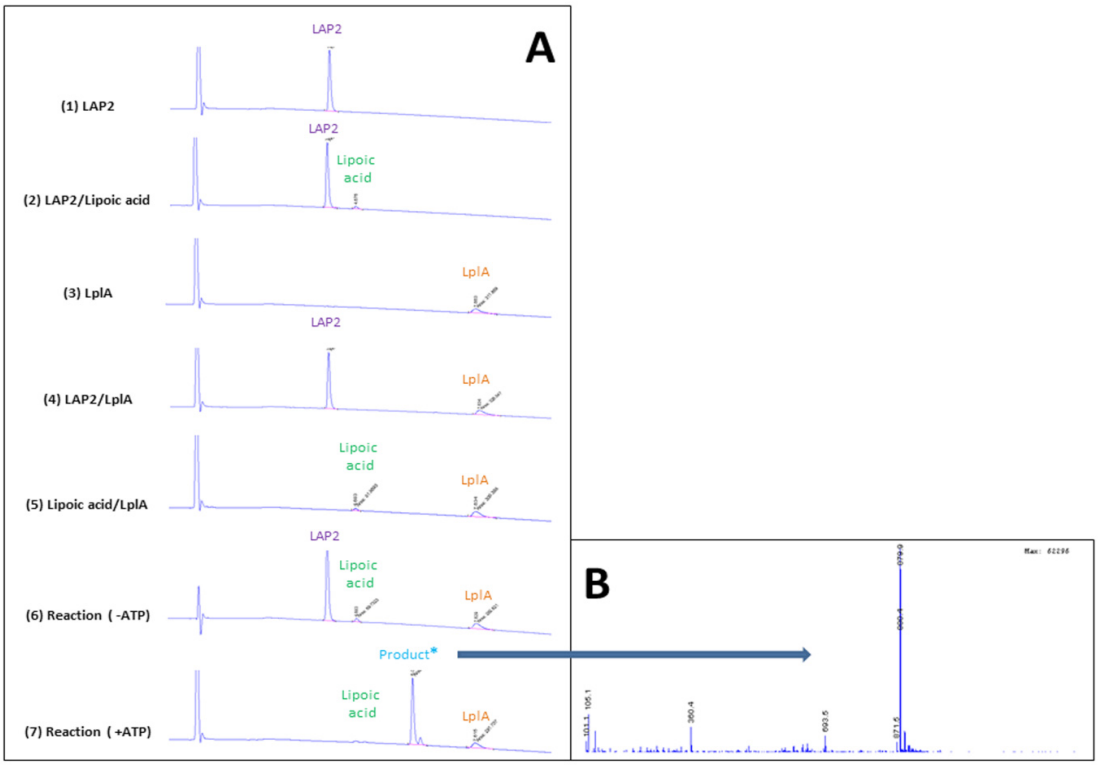


Figure 5: Confirmation of LplA-specific activity with ATP.

Reverse-phase chromatography assay showing the Lpla-mediated ligation of lipoic acid to LAP2 peptide

(A). Chromatograms were registered at 210 nm. (1) LAP2 peptide at 200 µM in reaction buffer (25 mM sodium phosphate, pH 8, 2 mM magnesium acetate, 1 mM ATP). (2) LAP2 peptide at 200 µM combined with lipoic acid at 350 µM in reaction buffer. (3) LplA at 2 µM in reaction buffer. (4) LAP2 peptide at 200 µM combined with LplA at 2 µM in reaction buffer. (5) Lipoic acid at 350 µM and LplA at 2 µM in reaction buffer. (6) LAP2 peptide at 200 µM combined with lipoic acid at 350 µM and LplA at 2µM in reaction buffer without ATP. (7) LAP2 peptide at 200 µM combined with lipoic acid at 350 µM and LplA at 2 µM in reaction buffer. These chromatograms confirmed the specific activity of LplA with ATP. The starred peak was analyzed by mass spectrometry (see Figure 5B). LC/MS analysis of lipoic acid coupling to LAP2 peptide mediated by LplA

(B): LAP2 peptide at 200 µM was combined with lipoic acid at 350 µM and LplA at 2 µM in reaction buffer. Reverse-phase chromatography assay coupled with mass spectroscopy confirmed the molecular weight of the reaction product.

Finally, using the peptide substrate, we completed the characterization of the enzyme by determining its kinetic constants from the standard Michaelis-Menten curves built from the experiments (Figure 6). We obtained the kinetic parameters of the reaction for the wt lipase by varying the concentrations of substrate (lipoic acid) at a fixed concentration of co- substrate (peptide LAP2) or vice versa. From these experiments, we calculated that the Kcat = $0.27s^{-1}$ and $0.32s^{-1}$ while the KM for the substrate was 19 µM and the KM for the co-substrate was 25 µM. The confidence in these results (n = 2) was calculated to be $R^2 = 0.98$ and 0.97, respectively. Similar kinetic experiments were run with the mutated enzymes (not shown). Their characteristics were only marginally affected by their respective mutation(s).

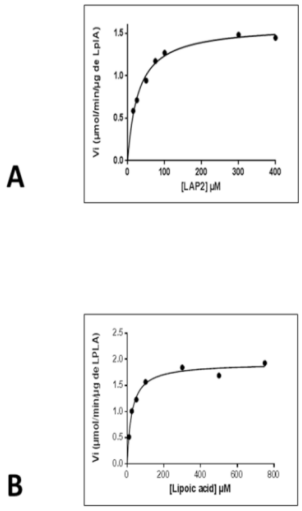


Figure 6: Kinetic characteristics of LplA towards its natural co-substrate, lipoic acid. Various concentrations of LAP2 peptide were lipoylated with 0.01 to 0.005 µM LplA and 750 µM lipoic acid

(A). The amount of product obtained at each time point was plotted against time to obtain the initial velocity for each concentration of LAP2 peptide. The peptide LAP2 at 200µM was lipoylated with various concentrations of lipoic acid and 0.025 to 0.005 µM LplA

(B). The amount of product obtained at each time point was plotted against time to obtain the initial velocity for each concentration of LAP2 peptide. The collection of initial velocities (Vi) was then plotted against LAP2 peptide concentration and fit to the Michaelis-Menten equation (Vi = Vmax [LAP2]/(Km + [LAP2])) to obtain the Km for LAP2. In both cases, the reactions were stopped with EDTA 0.3 M, Tris pH 7.5, 30% CH3CN. Samples were injected in HPLC as described in the experimental section.

Furthermore, the capacity of each enzyme, including the mutated ones, to recognize each of the various unnatural co-substrates (7-azidoheptanoic acid, 10-azidodecanoic acid, and 11-bromo-undecanoic acid) was tested. The results are summarized in Figure 7. These measurements confirmed that W37I is the most efficient mutant for the ligation of 10- azidodecanoic acid [18]. For 11-bromo-undecanoic acid, W37I was more efficient than W37A, which has been used in the literature for this substrate [27]. For 7-azidoheptanoic acid, LplA wild type was the most efficient mutant. This substrate has not been published, but the current results confirmed that for small azido co-substrates, the wt enzyme is the better enzyme whereas for longer co-substrates, the W37I mutant is more efficient.

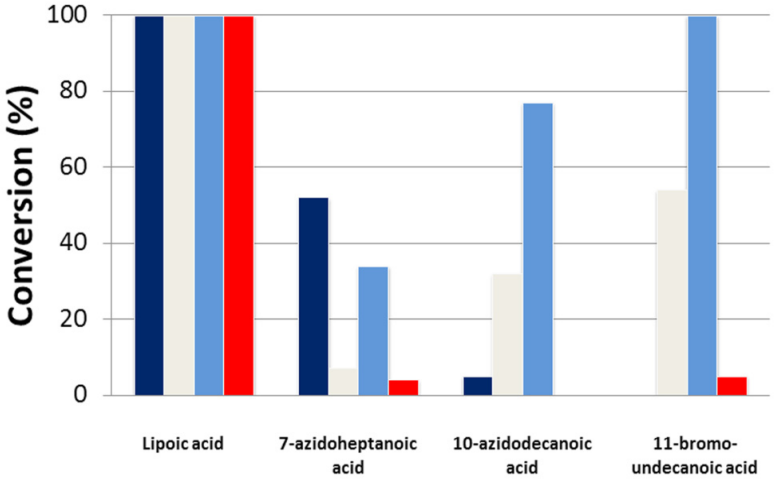


Figure 7: Wt and mutated ligase Lpa2-catalyzed ligation of different co-substrates of LAP2 peptide.LplA-mediatedligation kinetics on LAP2 peptide with lipoic acid, 7-azidoheptanoic acid, 10- azidodecanoic acid, and 11-bromo-undecanoic acid. Kinetics measurements were done with 1 µM LplA, 350 µM co-substrate, and 200 µM LAP2 peptide. Conversions were measured by HPLC assay as previously described. The conversion into product after 60 min is reported for each co-substrate with each LplA.

To ensure the validity of these observations, as with the natural substrate/co- substrate couple, the mass spectrometry analyses of those products were obtained, as reported in Figure 2S, to ensure the nature of the product generated by the various enzymes.

Enzyme-catalyzed site-specific ligation to rhoA LAP2

As a systematic approach to translating our observation first into a full-size protein and ultimately into a cellular context, we designed two model proteins (the GTPase rhoA and Trx) fused with the Lap2 peptide recognized by the LplA enzymes, as described above. These proteins were chosen because our laboratory has previous experience working with Trx [28] and rhoA (Petit et al., in preparation). We designed the fused proteins according to a triple scheme: the peptide was fused to the N-terminus, to the C-terminus, or finally in the middle of a sequence taking into account its tridimensional structure so that this extra sequence would be positioned inside a loop reachable from the outside of the protein, as in an in situ situation.

Figure 8 shows the results of the expression and purification of the three proteins by SDS-PAGE electrophoresis. Despite the changes - especially in the in-loop addition of the sequence - the three fused rhoA proteins were expressed in bacteria to workable levels. At the final step, the proteins, obtained using standard expression/extraction/purification processes, were pure at more than 95% (Figure 8).

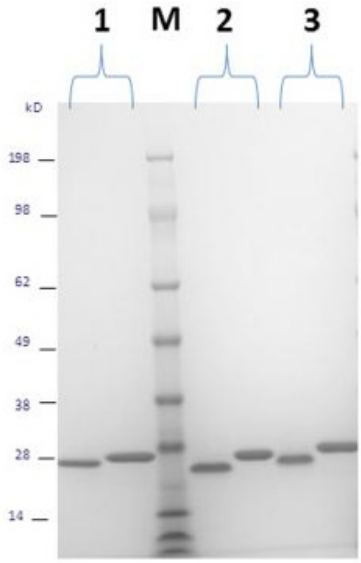


Figure 8: SDS-PAGE analysis of the three rhoA proteins fused with the LAP2 peptides. The rhoA proteins were fused with the LAP2 peptide, at the N-terminus, the C-terminus, or in the middle of the sequence. They were expressed in *E. coli*, extracted, and purified as described in the material and methods section. Lane 1: The LAP2-fused rhoA in the N-terminus, in reduced and non- reduced conditions. Lane 2: The rhoA-fused LAP2 at the C-terminus, in reduced and non-reduced conditions. Lane 3: The fused LAP2-rhoA protein, in which the LAP2 sequence is embedded in the rhoA sequence, in a loop. Lanes represent the reduced (left) or the non-reduced (right) samples.

Because massive expression of mammal proteins is not a guarantee of stability and activity, we systematically analyzed the three newly produced analogues of rhoA by both mass spectrometry and SEC-MALS. The results of both experiments are shown in Figure 3S. In brief, the mass spectrometry analyses showed that the three proteins were in line with the expected sequences (Figure 3SA). Nevertheless, the SEC-MALS experiments (Figure 3SB) clearly showed that only the rhoA fused at the C-terminus with the LAP2 peptide did not aggregate, rendering the use of the two other proteins unsuitable for our purpose. Therefore, we used this whole fused protein as a possible substrate of LplA, according to Scheme 2. The result of the incubation was analyzed by HPLC, as described in the materials and methods section. As shown in Figure 9, the protein was clearly modified by the LplA ligase to lead to a rhoA fused to the LAP2 peptide on which the azidodecanoic acid was transferred to the lysine of the LAP2 peptide.

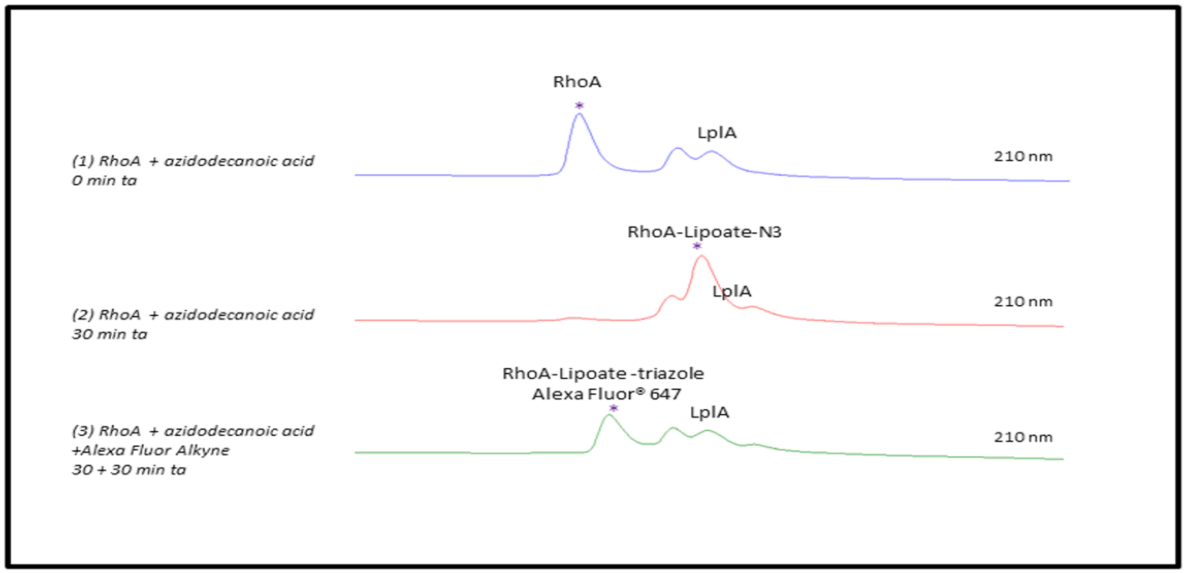


Figure 9: Enzyme-catalyzed ligation of 10-azidodecanoic acid to rhoA LAP2. The first step corresponds to the LplA W37I enzyme-catalyzed ligation of 10-azidodecanoic acid onto the lysine of the rhoA-LAP2. The second step is a click reaction between the azido derivative and Alexa Fluor® 647 DIBO Alkyne. Lane (1): starting conditions, analyzed at t = 0. Lane (2): analysis of the reaction medium after 30 min incubation. Lane (3): reaction as in (2) reacting with the Alexa Fluor 647 alkyne for 30 more minutes.

RhoA has a feature that guided our choice: It is a protein that, once farnesylated, is transferred to the membrane (see Nguyen et al. [8] for general information on prenylation). Given that our main goal was to trace the transport of the protein inside the cell, maintaining this feature in the final construct (the fused protein with the LAP2 peptide) was critical. Unfortunately, as has been known for decades, farnesylation is the first step in the cleavage of the remaining part of the CAAX box, downstream from the farnesylation site (the C of the CAAX box). It became obvious that the sole usable protein bearing this peptide was the one fused at the C-terminus. This construct would not allow for farnesylation [29], much less the cleavage and therefore transfer of the protein from the cytosol to the membrane. Despite the endpoint being a failure,we learned enough in the fusion, expression, and production processes to prepare the next protein to be tested in the system: Trx.

Enzyme-catalyzed site-specific ligation to Thioredoxin-LAP2

As described above for rhoA, our first task was to produce the three candidate proteins: the one with the peptide tag fused at the N-terminus, one with the C-terminus, and one in which the sequence was added to an internal loop of Trx. In contrast with results with the preceding protein, Trx was produced by the bacteria only when the tag was fused on the C-terminus. Because the protein is also His-tagged at the N-terminus, we purified the fused protein to homogeneity and characterized it by mass spectrometry. As Figure 10 shows, the protein was almost homogeneous by SDS-PAGE while the SEC-MALS gave a 98% purity result (Figure 10A). Nevertheless, the mass spectrometry analysis of the fused protein showed that 28% of the protein was gluconoylated, a feature that is frequent in bacterially expressed proteins bearing an HAHA tag [26] (Figure 10B). The measured molecular weight (14426.1 $g.mol^{-1}$) is consistent with the expected one (14428.3 $g.mol^{-1}$) considering deletion of 3 amino-acids: 1 amino-acids in N-ter corresponding to methionine deletion and 2 amino acids in C-ter corresponding to “LE” deletion.

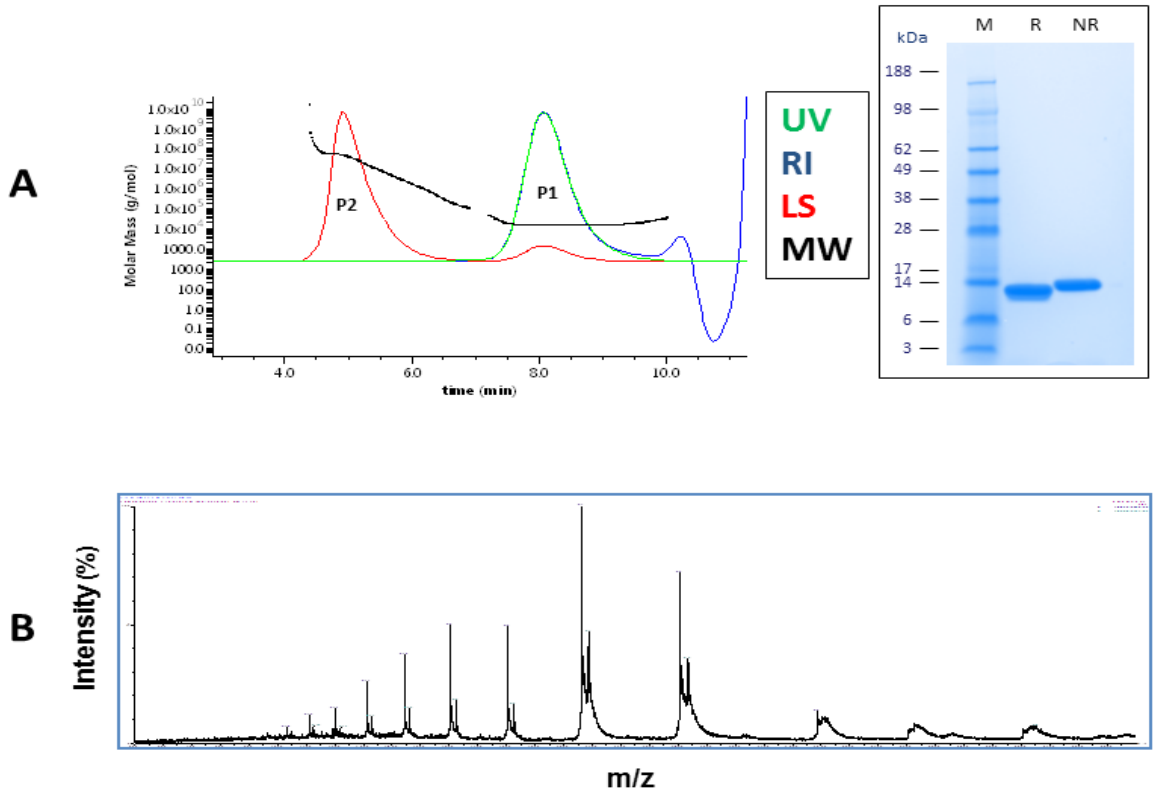


Figure 10: Characterization of Thioredoxin LAP2: Nano-ESI spectra of Thioredoxin LAP2 in native conditions. Peak A corresponds to 6His-Trx-LAP2, peak B corresponds to posttranslational modification (gluconoylation + 178Da)

Pannel A, Left: SEC-MALS analysis of Trx-LAP2 using a Superdex 75 5/150 column at 0.2 ml.min-1 with running buffer: Tris HCl 50 mM, NaCl 200 mM, TCEP 1 mM, pH 7.5. The graphic shows UV signal (green), LS signal (red), RI signal (blue), and molecular weight (black) according to time. All signals are represented on a normalized scale. The enzyme is homogeneous and monomeric at 98% (P1). P2 represents aggregates.

Pannel A, Right: SDS-PAGE analysis of Trx-LAP2, R = reduced condition, N R = non-reduced condition, M = SeeBlue® Plus2 Pre-stained Protein Standard. Analyses are in accordance with SEC- MALS analysis and theoretical molecular weight.

Pannel B, Mass spectrometry analysis of the Trx-LAP2 fused protein: MS analysis of TrxLAP2. Peak A corresponds to Trx-(6His-Trx-LAP2) truncated from its 3 amino-acids: 1 amino-acids in N-ter corresponding to methionin deletion and 2 amino acids in C-ter corresponding to "LE" deletion. Peak B corresponds to posttranslational modification of Trx-LAP2 (gluconoylation). Population A is estimated at 73% and population B at 27%. (150 mM ammonium acetate pH 7.4)

As summarized in Scheme 3, we then used this protein as a substrate for the LplA enzyme. The first step was an enzymatically driven transfer of 10-azidodecanoic acid onto the tagged protein, and then a click-chemistry modification of the 10- azidodecanoic acid with Alexa Fluor® 647 DIBO Alkyne. The successive HPLC analyses of the reaction are given in Figure 11, clearly showing the appearance of the end product of the reactions (see chromatogram 4, Figure 11). Indeed, these analyses show that after labeling with 10-azidodecanoic acid and Alexa Fluor® 647 DIBO Alkyne, the peak corresponding to Trx absorbed at 638 nm, the specific absorbance of the fluorophore.

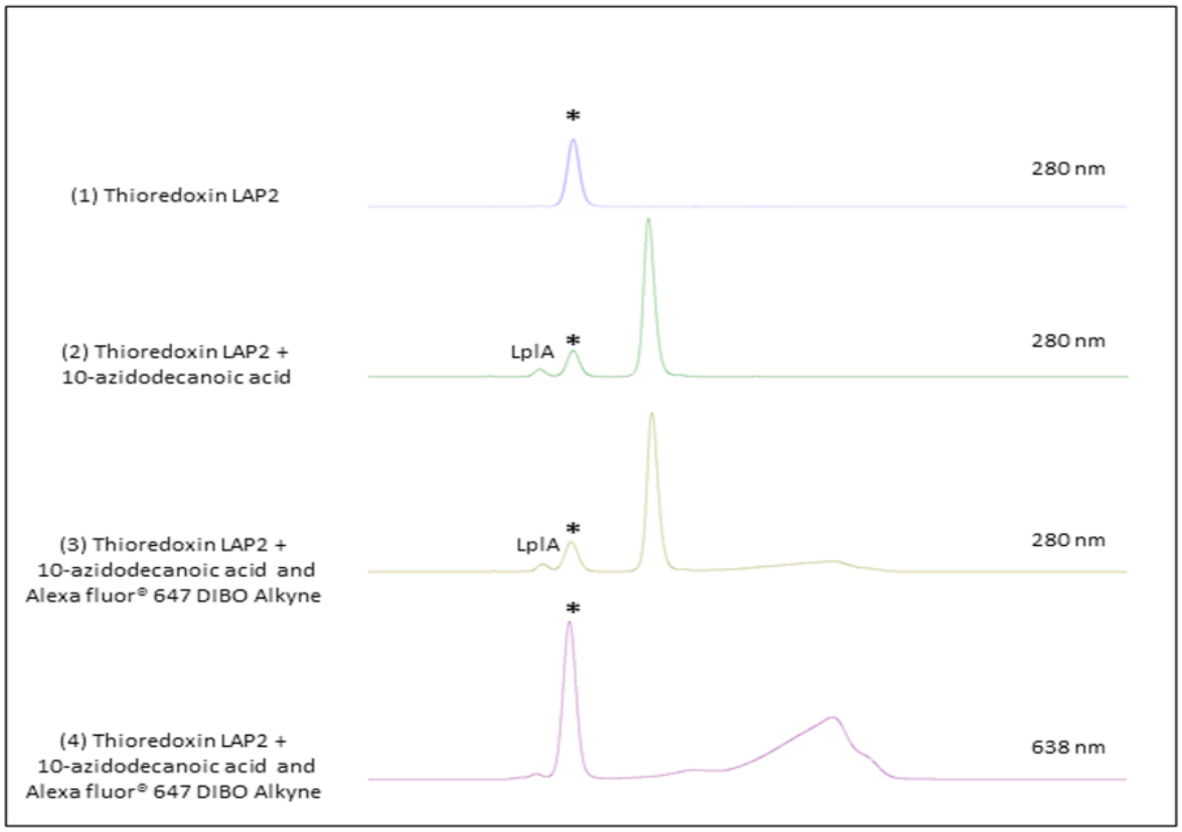


Figure 11: SEC analysis of labeled thioredoxin-LAP2. Size exclusion chromatography of:

1. Trx-LAP2 diluted at 69 µM (1 mg/ml) in Tris HCl 50 mM, NaCl 200 mM, TCEP 1 mM, pH 7.5. Spectrum was recorded at 280 nm.
2. Trx-LAP2 at 20 µM labeled with 10-azidodecanoic acid at 50 µM and LplA W37I at 1 µM. After labeling, the reaction was dialyzed against Tris HCl 50 mM, NaCl 200 mM, DTT 5 mM, pH 7.5, and concentrated to ~1 mg/ml. The spectrum was recorded at 280 nm.

(3, 4) Same as sample (2) and after click reaction with Alexa Fluor® 647 DIBO Alkyne at 100 µM. Spectra were recorded at 280 nm (lane 3) and 638 nm (lane 4).

For each spectrum, the Trx peak is indicated by a star. The analysis was performed with a Superdex 75 5/150 column at 0.2 ml.min-1 with running buffer: Tris HCl 50 mM, NaCl 200 mM, pH 7.5.

At the last step of the experiment, the reaction cocktail was purified by gel filtration (essentially to eliminate the non-reactivated fluorophore) and the Alexa-labeled protein was further analyzed by mass spectrometry. Figure 12 summarizes the two-step analyses. Purification led to a protein that was clearly homogeneous (Figure 12A). The protein was effectively labeled by the fluorophore and almost no non-site labeling was observed, proving (if necessary) that this approach can lead to a perfectly mastered labeling of the target protein at a given place, in contrast to what happens when chemical labeling is performed (see below). MS analyses show one peak corresponding to Trx-LAP2 + Alexa Fluor® 647 DIBO (15781.8 g.mol-1) (peak A, figure 12B) and one peak corresponding to Trx-LAP2 + Alexa Fluor® 647 DIBO with posttranslational modification (gluconoylation: 15797.4 g.mol-1) (peak B, figure 12B).

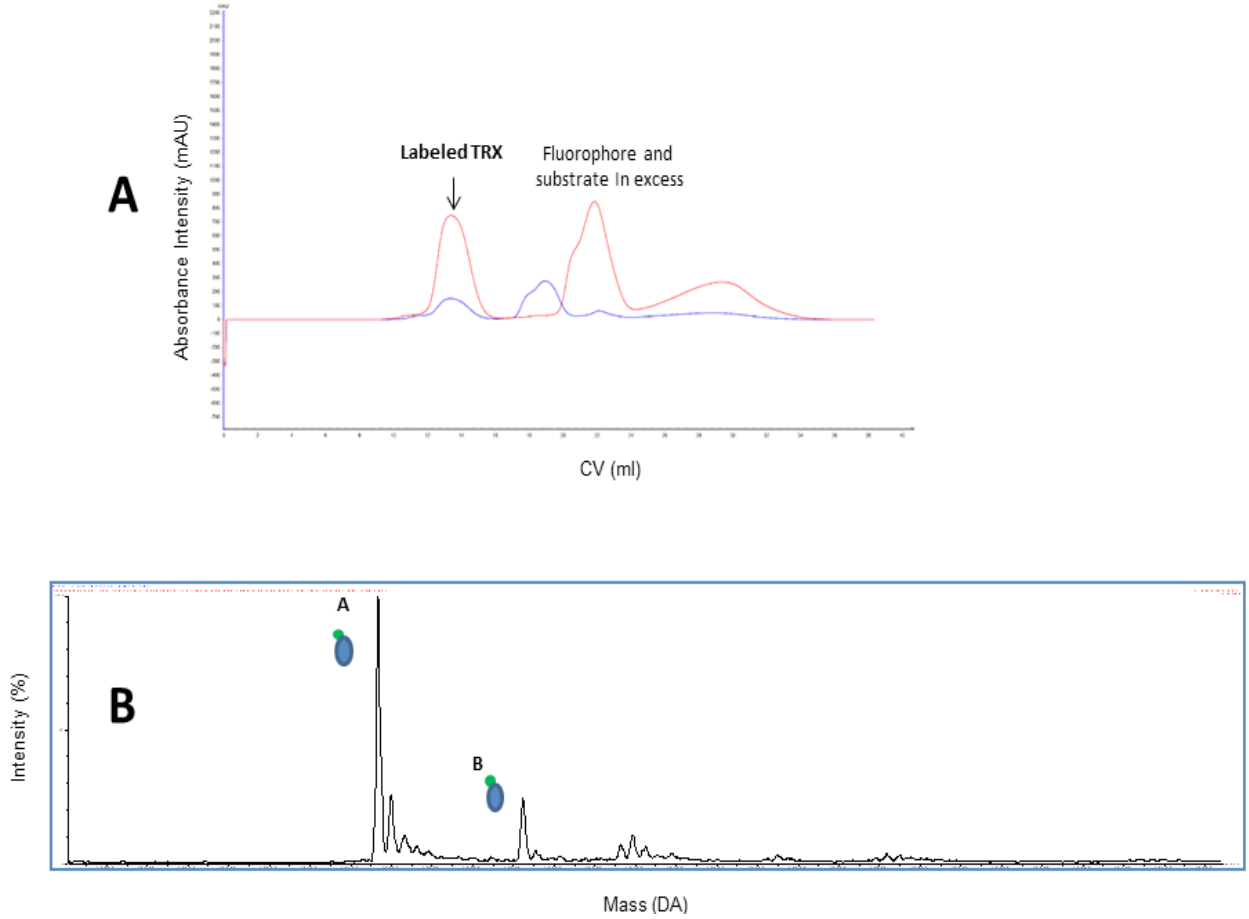


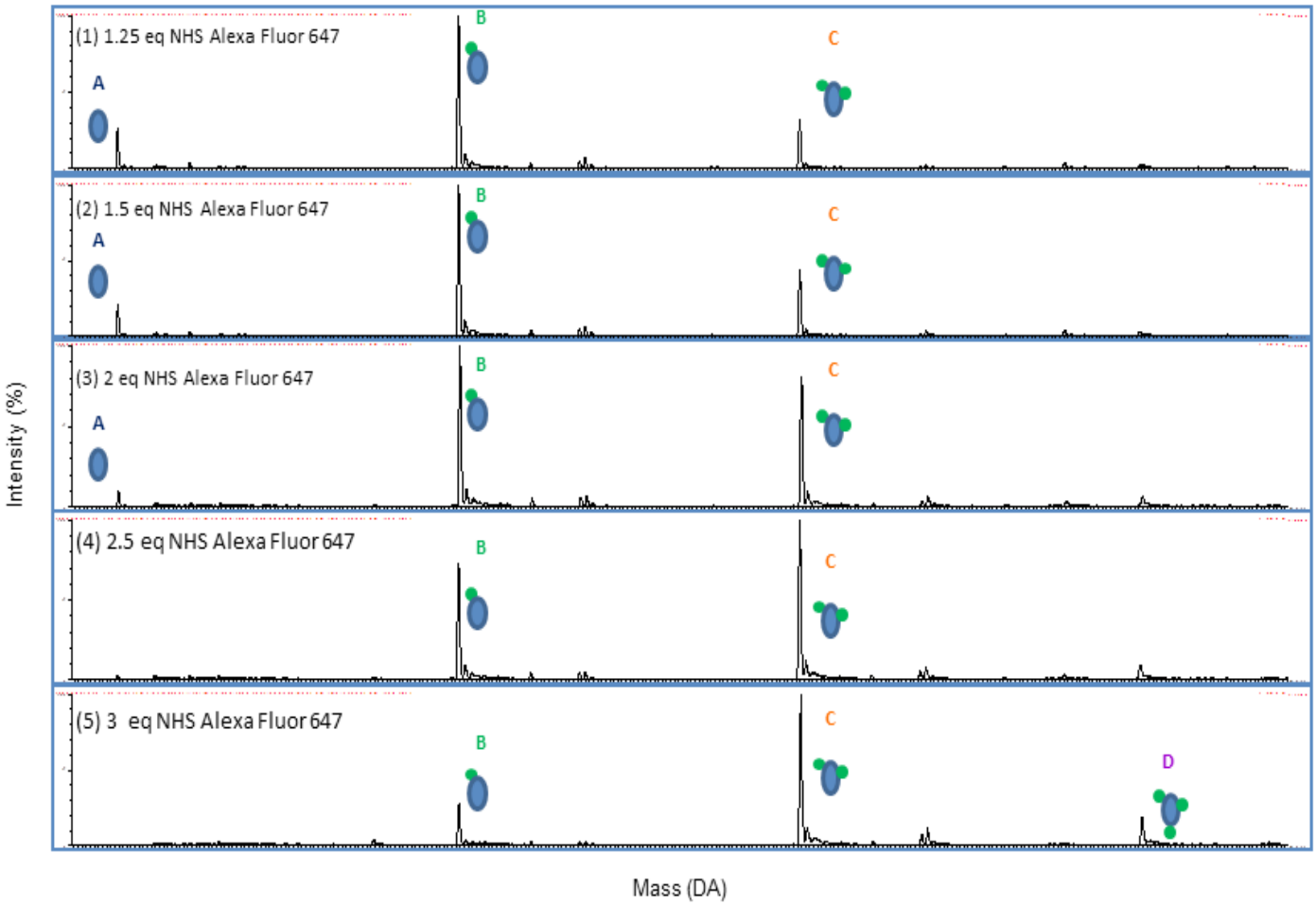
Figure 12: Purification chromatogram of labeled TRX-LAP2.

(A) Purification of the Alexa-labeled Trx-LAP2 protein. After the incubation of the Trx- LAP2 fused protein with LplA ligase in the presence of 10-azidododecanoic acid, the labeled protein was incubated with Alexa Fluor 647 DIBO alkyne that reacted with the azido moiety of the dodecanoic acyl moiety. The material was subjected to chromatogram purification. This material should correspond to the Trx-LAP2 fused protein labeled (at the lysine of LAP2) with 10-azidodecanoic acid. This moiety should react with Alexa Fluor® 647 DIBO Alkyne by click chemistry, leading to a final, specifically Alexa-labeled protein on the LAP2 peptide. The analysis was performed with a Superdex 75 10/300 column at 0.5 ml.min-1 with the running buffer (Tris HCl 50 mM, NaCl 200 mM, 1 mM TCEP, pH 7.5). Absorbance wavelengths were recorded at 280 nm (blue) and 650 nm (red). The main labeled Trx peak was collected for further analysis.

(B) Nano-ESI spectra of labeled thioredoxin LAP2 in native conditions. The characterization of the purified labeled protein was done by mass spectrometry analysis. Trx-LAP2 was labeled as described previously with 10-azidodecanoic acid and Alexa Fluor® 647 DIBO Alkyne. Peak A corresponds to labeled thioredoxin LAP2, peak B corresponds to labeled thioredoxin LAP2 with posttranslational modification (gluconoylation + 178Da). The analysis shows that all of the protein is labeled (no unlabeled product remaining, suggesting 100% yield + purification.

Comparison of the ligase-catalyzed site-specific ligation of thioredoxine-LAP2 with chemical ligation

Proteins are generally labeled using several technologies that often lack controlled specificity. To compare our partially enzymatic approach with a standard technique of protein labeling, we performed an experiment in which the same protein was used, but the labeling was done by a chemical approach. The conditions are summarized in the materials and methods section. In brief, the labeling was performed using five different amounts of Alexa: from 0.5 to 3 equivalents (Figure 13).



Electrospray mass spectrometry of labeled TRX Wt with NHS Alexa fluor 647 Labeling was done with 1.25 (1), 1.5 (2), 2 (3), 2.5 (4) or 3 (5) equivalents of fluorophore to evaluate the fluorophore protrein ratio (FPR). Wt Trx was labeled chemically with NHS Alexa Fluor® 647 DIBO Alkyne as described in the experimental section. Labeling was done with 1.25 (lane 1), 1.5 (lane 2), 2 (lane 3), 2.5 (lane 4), or 3 (lane 5) equivalents of fluorophore to evaluate the fluorophore protein ratio. All reactions were analyzed by mass spectrometry. Peak A is unlabeled wt Trx; peak B, Trx labeled with one molecule of NHS Alexa Fluor 647; peak C, Trx labeled with two molecules of NHS Alexa Fluor 647, and peak D, Trx labeled with three molecules of NHS Alexa Fluor 647.

In all cases, at least two species could be found: one labeled with one Alexa, the other with two. The nature of the link between the fluorophore and the amino acid seems to be unique, but the respective amounts of each species are variable according to the amount of Alexa present in the labeling system. In other words, there is no way to obtain a unique labeling of the protein under our conditions whereas with the enzyme-directed labeling, the specificity of the labeling was beyond doubt.

Discussion

It is noteworthy that there is a huge need for protein-modifying techniques that are usable across a range of situations, from the acellular to the cellular contexts [see reviews by Rashidian et al. [30] and Hinner and Johnson [4]. For example, one can 'easily' introduce exotic amino acids into a protein by using direct, complete chemical synthesis of proteins [3,5], using semi-synthetic approaches (intent) [31], or manipulating the cellular genetic ma-

terial of bacteria [2] to make it able to integrate into a newly synthesized protein an extra amino acid encoded by mutated codons. All of these techniques have been developed in the last two decades and provide researchers with a palette of tools adapted to their questions. Nevertheless, none of these approaches is straightforward or devoid of downsides. For example, the manipulation of tRNA is not easy and is often limited to one or two exotic amino acids incorporated per protein. Furthermore, their amount and the homogeneity of the final product are often underestimated or disregarded. The synthetic chemical proteins that we favored in our laboratory supply enough material to perform quality-control experiments [5], but the study of the final enzyme in a cellular context is complicated and requires either patch-clamp techniques (and therefore large cells) or detergent treatments to produce pores at the surface, allowing insertion of the protein in a cellular context. Again, none of these techniques is simple or easy to use, even though they do yield splendid results and beautiful experiments (see review by Xiao and Schultz [32]).

We recently decided to complete our panel of techniques in these new areas of biochemistry with in situ, specific techniques that, rather than introducing exotic amino acids into a protein as we have already reported [5], result in modification of a protein at a perfectly appropriate position in the protein sequence, provided that the protein can be fused to a small, well-chosen peptide bearing a ligase recognition sequence. The techniques that we extrapolated from the elegant work of Ting and colleagues (see Liu et al. [33] and references therein) allowed mastery of the conditions under which the ligase can be manipulated, particularly the mutant that can recognize co-substrates other than the wild type, such as Loose azide derivatives of lipoic acid or of other fatty acids [34]. From there, it is possible to introduce almost any type of compound onto this azide residue by way of click chemistry. The possibility of introducing in a cellular-friendly manner a fluorophore [34], radioactive compounds [21], or simple chemicals with special biophysical characteristics [35] is an open door towards many experiments in a cellular context. These approaches were dramatically lacking from our available tools; therefore, we believed that a systematic validation of this approach in an acellular context would support confidence in their transposition in cells, particularly regarding the position of the tag in the fused protein; we noted that adding this particular sequence impaired the expression of this protein in bacteria, suggesting that a similar situation would arise in a cellular host, despite pioneering works with the model green fluorescent protein [34]. The next step will be to translate the current knowledge to TRX intracellular trafficking.

Hinner and Johnson [4] have reviewed the top features of protein labeling. Among the highlights, the minimal size of the labeled molecule (to limit interference with protein function), absence of non-target labeling (specificity), and possibility of introducing any kind of chemicals through this particular labeling process are the main features to expect from any such methodology. Our feeling is that with the sole necessity of having to have the enzyme present in the targeted tissues or cells, the use of lplA addresses these requirements with success; of course, however, other enzymes or approaches are also usable (see review by O'Hare et al. [36]). The fact that one can introduce an alkyne function suggests that any further modification of the protein bearing this function will be limited only by the chemistry, and more particularly by the penetration of the various chemicals into the cells. This particular characteristic has been retained for many applications [6-18 ,23,27,33] in which the mutagenesis of the lplA led to mutants that can accept more broadly other chemicals distantly derived from lipoic acid, the natural co-substrate of the enzyme.

In the present work, we confirmed the basic observations for the use of the lplA ligase as a follow-up of the pioneering work of Ting and colleagues. We have extended those observations here and believe these findings will help the community apply this versatile system for such approaches and many others. We used a new approach in the sense that we demonstrated in an acellular system that a protein fused with the peptide substrate of the ligase can be expressed, purified, characterized, and successfully used in an enzymatic assay with the full-length protein (as opposed to a model peptide).

It seems that the path pioneered by Ting's group regarding this particular ligase also could be explored and extended to different protein contexts, such as proteins involved in the microorganism synthesis of natural products like peptide synthetic's [16] or cyclases [37]. From there, one can anticipate that many tools will become available in the general area of synthetic biology, either for labeling proteins at different points of their sequences in a very specific manner or for providing alternative pathways for microorganism production of natural or new diverse chemicals.

Supplementary Figures

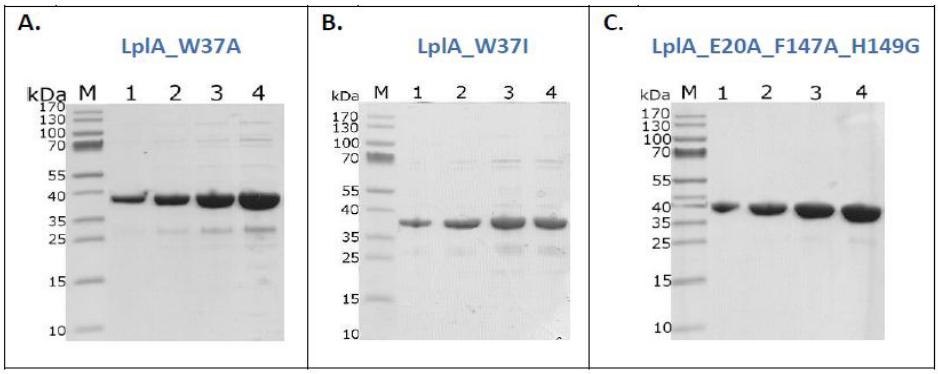


Figure 1S: SDS-PAGE analyses of the purified mutated versions of LpIA: LpIA_W37A (A), LpIA_W37I (B), and LpIA_E20_F147A_H149G (C). Lanes: M - molecular mass standard; 1-4, 2, 5, 10, and 15 µg of protein were added per lane, respectively.

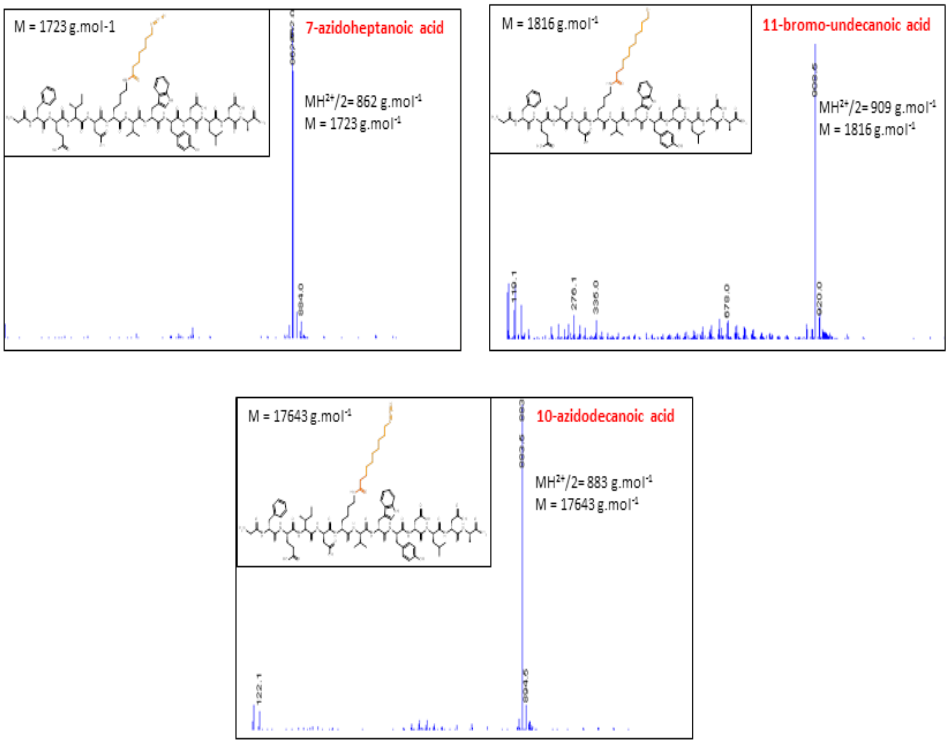


Figure 2S: Mass spectrometry-coupled liquid chromatography analyses of substrates ligation on LAP2. LC/MS analysis of LAP2 peptide coupling with 7-azidoheptanoic acid, 10-azidodecanoic acid, and 11-bromo-undecanoicacid. Enzymatic reactions were done with 1 µM LplA W37I, 200µM LAP2, and 350 µM of substrate. The reaction was stopped with 300 mM EDTA, 50 mM Tris, pH 7.5, 30% CH3CN, and the samples were injected in LC/MS as described in the experimental section.

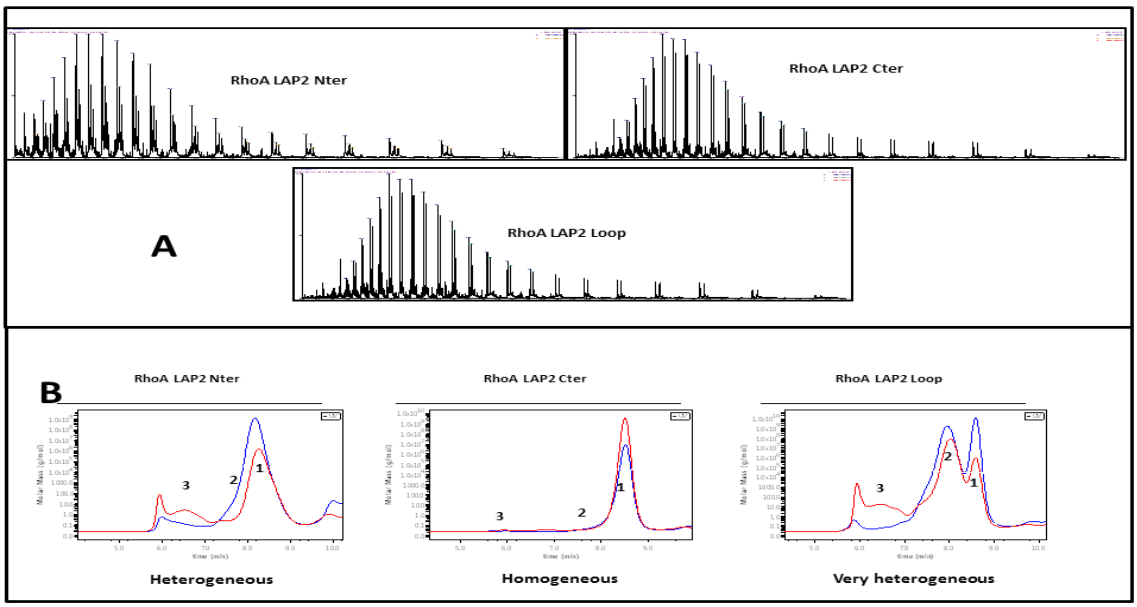


Figure 3S: Mass spectrometry analyses of LAP2-fused rhoA.

A: Mass spectrometry analysis of LAP2-fused rhoA mutants. RhoA LAP2 Niter refers to 8His-GFEIDKVWYDLDA-Rho, RhoA LAP Cater to 8His-RhoA-GFEIDKVWYDLDA, and RhoA LAP2 Loop to 8His-RhoA-D49 [GGSGFEIDKVWYDLDAGGS] G50. For all spectra, the major fraction detected is attributed to RhoA deleted from methionine (22568.4 Da for RhoA LAP2 Niter, 22639.6 Day for RhoA LAP2 Cater, and 23042 Day for RhoA LAP2 Loop). A secondary fraction was observed and assigned to posttranslational modification (gluconoylation + 178 Da). Other ions were present and corresponded to oxidized forms (+16 Day).

B: SEC-MALS analyses of RhoA LAP2.SEC-MALS, analysis of RhoA LAP2 Niter, RhoA LAP2 Cater, and RhoA LAP2 Loop. SEC-MALS analysis was done using a Shodex KW 802.5 column at 0.8 ml.min-1with running buffer: Tris HCl 20 mM, NaCl 0.3 M, pH 7.4. The graphic shows UV signal (green), LS signal (red), RI signal (blue), and molecular weight (black) according to time. All signals are represented on a normalized scale. For each construct, analyses were done at t0 and after 1 h of incubation at 37°C. For each graph, P1 represents the monogenic protein, P2 the diametric protein and P3 aggregates and oligomers.

References

1. Vila-Perello M and Muir TW (2010) Biological applications of protein splicing. Cell 143: 191-200.
2. Liu CC and Schultz PG (2010) Adding new chemistries to the genetic code. Annu Rev Biochem 79: 413-444.
3. Kent S, Sohma Y, Liu S, Bang D, Pentelute B, et.al (2012) Through the looking glass--a new world of proteins enabled by chemical synthesis. J Pept Sci 18: 428-436.
4. Hinner MJ and Johnsson K (2010) How to obtain labeled proteins and what to do with them. Curr Opin Biotechnol 21: 766-776.
5. Bacchi M, Jullian M, Sirigu S, Fould B, Huet T, et.al (2016) Total chemical synthesis, refolding, and crystallographic structure of fully active immunophilin calstabin 2 (FKBP12.6). Protein Sci 25: 2225-2242.
6. Fernandez-Suarez M, Baruah H, Martinez-Hernandez L, Xie KT, Baskin JM, et.al (2007) Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. Nat Biotechnol 25: 1483-1487.
7. Fujiwara K, Toma S, Okamura-Ikeda K, Motokawa Y, Nakagawa A, et.al (2005) Crystal structure of lipoate-protein ligase A from Escherichia coli. Determination of the lipoic acid-binding site. J Biol Chem 280: 33645-33651.
8. Nguyen UT, Goody RS, Alexandrov K (2010) Understanding and exploiting protein prenyltransferases. Chembiochem 11: 1194-1201.
9. Boutin JA (1997) Myristoylation. Cell Signal 9: 15-35.
10. Pelletier N, Gregoire S, Yang XJ (2008) Analysis of protein lysine acetylation in vitro and in vivo. Curr Protoc Protein Sci Chapter 14.
11. Zeidman R, Jackson CS, Magee AI (2009) Analysis of protein acylation. Curr Protoc Protein Sci Chapter 14.
12. Laney JD, Hochstrasser M (2011) Analysis of protein ubiquitination. Curr Protoc Protein Sci Chapter 14.
13. Sarge KD (2016) Analysis of protein sumoylation. Curr Protoc Protein Sci 83 : 1-8.
14. Yan LJ (2009) Analysis of oxidative modification of proteins. Curr Protoc Protein Sci Chapter 14: 4.1-4.28.

15. Marahiel MA, Stachelhaus T, Mootz HD (1997) Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. Chem Rev 97: 2651-2674.

16. Mootz HD, Schwarzer D, Marahiel MA (2002) Ways of assembling complex natural products on modular nonribosomal peptide synthetases. Chembiochem 3: 490-504.

17. Puthenveetil S, Liu DS, White KA, Thompson S, Ting AY (2009) Yeast display evolution of a kinetically efficient 13-amino acid substrate for lipoic acid ligase. J Am Chem Soc 131: 16430-16438.

18. Yao JZ, Uttamapinant C, Poloukhtine A, Baskin JM, Codelli JA, et.al (2012) Fluorophore targeting to cellular proteins via enzyme-mediated azide ligation and strain-promoted cycloaddition. J Am Chem Soc 134: 3720-3728.

19. Uttamapinant C, Sanchez MI, Liu DS, Yao JZ, Ting AY (2013) Site-specific protein labeling using PRIME and chelation-assisted click chemistry. Nat Protoc 8: 1620-1634.

20. Liu DS, Tangpeerachaikul A, Selvaraj R, Taylor MT, Fox JM, et.al (2012) Diels-Alder cycloaddition for fluorophore targeting to specific proteins inside living cells. J Am Chem Soc 134: 792-795.

21. Drake CR, Sevillano N, Truillet C, Craik CS, VanBrocklin HF, et al. (2016) Site- Specific Radiofluorination of Biomolecules with 8-[(18)F]-Fluorooctanoic Acid Catalyzed by Lipoic Acid Ligase. ACS Chem Biol 11: 1587-1594.

22. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR (2006) A comparative study of bioorthogonal reactions with azides. ACS Chem Biol 1: 644-648.

23. Uttamapinant C, White KA, Baruah H, Thompson S, Fernandez-Suarez M, et al. (2010) A fluorophore ligase for site-specific protein labeling inside living cells. Proc Natl Acad Sci U S A 107: 10914-10919.

24. Cohen JD, Thompson S, Ting AY (2011) Structure-guided engineering of a Pacific Blue fluorophore ligase for specific protein imaging in living cells. Biochemistry 50: 8221-8225.

25. Jin X, Uttamapinant C, Ting AY (2011) Synthesis of 7-aminocoumarin by Buchwald- Hartwig cross coupling for specific protein labeling in living cells. Chembiochem 12: 65-70.

26. Geoghegan KF, Dixon HB, Rosner PJ, Hoth LR, Lanzetti AJ, et al. (1999) Spontaneous alpha-N-6-phosphogluconoylation of a “His tag” in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. Anal Biochem 267: 169-184.

27. Liu DS, Phipps WS, Loh KH, Howarth M, Ting AY (2012) Quantum dot targeting with lipoic acid ligase and HaloTag for single-molecule imaging on living cells. ACS Nano 6: 11080-11087.

28. Fould B, Lamamy V, Guenin SP, Ouvry C, Coge F, et al. (2012) Mutagenic analysis in a pure molecular system shows that thioredoxin-interacting protein residue Cys247 is necessary and sufficient for a mixed disulfide formation with thioredoxin. Protein Sci 21: 1323-1333.

29. Boutin JA, Marande W, Petit L, Loynel A, Desmet C, et al. (1999) Investigation of S-farnesyl transferase substrate specificity with combinatorial tetrapeptide libraries. Cell Signal 11: 59-69.

30. Rashidian M, Dozier JK, Distefano MD (2013) Enzymatic labeling of proteins: techniques and approaches. Bioconjug Chem 24: 1277-1294.

31. Muir TW, Sondhi D, Cole PA (1998) Expressed protein ligation: a general method for protein engineering. Proc Natl Acad Sci U S A 95: 6705-6710.

32. Xiao H, Schultz PG (2016) At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code. Cold Spring Harb Perspect Biol 9.

33. Liu DS, Nivon LG, Richter F, Goldman PJ, Deerinck TJ, et al. (2014) Computational design of a red fluorophore ligase for site-specific protein labeling in living cells. Proc Natl Acad Sci U S A 111: E4551-E4559.

34. Plaks JG, Falatach R, Kastantin M, Berberich JA, Kaar JL (2015) Multisite clickable modification of proteins using lipoic acid ligase. Bioconjug Chem 26: 1104- 1112.

35. Best M, Degen A, Baalmann M, Schmidt TT, Wombacher R (2015) Two-step protein labeling by using lipoic acid ligase with norbornene substrates and subsequent inverse-electron demand Diels-Alder reaction. Chembiochem 16: 1158-1162.

36. O’Hare HM, Johnsson K, Gautier A (2007) Chemical probes shed light on protein function. Curr Opin Struct Biol 17: 488-494.

37. Zhang J, Liu N, Cacho RA, Gong Z, Liu Z, Qin W, Tang C, Tang Y, Zhou J (2016) Structural basis of nonribosomal peptide macrocyclization in fungi. Nat Chem Biol 12: 1001-1003.