ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

Ligation of Expressed Protein α-Hydrazides *via* Genetic Incorporation of an α-Hydroxy Acid

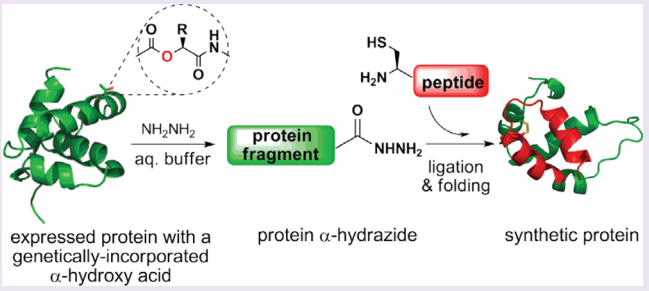
Yi-Ming Li,[†,§] Mai-Yun Yang,[‡,§] Yi-Chao Huang,[†] Yi-Tong Li,[†] Peng R. Chen,*[,‡] and Lei Liu*[,†]

[†]Tsinghua-Peking Center for Life Sciences, Department of Chemistry, Tsinghua University, Beijing 100084, China

[‡]Beijing National Laboratory for Molecular Sciences, Synthetic and Functional Biomolecules Center, Peking-Tsinghua Center for Life Sciences, and Department of Chemical Biology, Peking University, Beijing 100871, China

Supporting Information

ABSTRACT: Expressed protein ligation bridges the gap between synthetic peptides and recombinant proteins and thereby significantly increases the size and complexity of chemically synthesized proteins. Although the intein-based expressed protein ligation method has been extensively used in this regard, the development of new expressed protein ligation methods may improve the flexibility and power of protein semisynthesis. In this study a new alternative version of expressed protein ligation is developed by combining the recently developed technologies of hydrazide-based peptide ligation and genetic code expansion. Compared to the previous intein-based expressed protein ligation method, the new method does not require the use of protein splicing technology and generates recombinant protein α-hydrazides as ligation intermediates that are more chemically stable than protein α-thioesters. Furthermore, the use of an evolved mutant pyrrolysyl-tRNA synthetase (PylRS), ACPK-RS, from *M. barkeri* shows an improved performance for the expression of recombinant protein backbone oxoesters. By using HdeA as a model protein we demonstrate that the hydrazide-based method can be used to synthesize proteins with correctly folded structures and full biological activity. Because the PylRS-tRNA$^{Pyl}_{CUA}$ system is compatible with both prokaryotic and eukaryotic cells, the strategy presented here may be readily expanded to manipulate proteins produced in mammalian cells. The new hydrazide-based method may also supplement the intein-based expressed protein ligation method by allowing for a more flexible selection of ligation site.



Protein chemical synthesis enables a level of control on protein composition beyond that attainable by protein expression.[1−4] It can provide otherwise hardly accessible insights into protein's structure and function. The native chemical ligation (NCL) method invented by Kent *et al.* represents the most successful approach for protein chemical synthesis (Scheme 1).[5] This strategy involves a chemoselective amidation reaction of a C-terminal peptide α-thioester with an N-terminal Cys-peptide. Many small- to medium-sized proteins can be readily synthesized by using NCL, but the generation of large proteins through total synthesis is still a challenge. One practical solution to this size problem is the intein-based expressed protein ligation (EPL) developed by Muir *et al.*[6,7] This technology takes advantage of the protein splicing process and produces recombinant protein α-thioesters that can be used as precursors in the ligation reactions. In this manner EPL bridges the gap between synthetic peptides and recombinant proteins and thereby significantly increases the size and complexity of the synthetic target. Its great importance has been manifested by the synthesis and functional studies of proteins carrying various chemical modifications including fluorophores,[8] caging groups,[9,10] cross-linkers,[11] and post-translational modifications.[12−19]

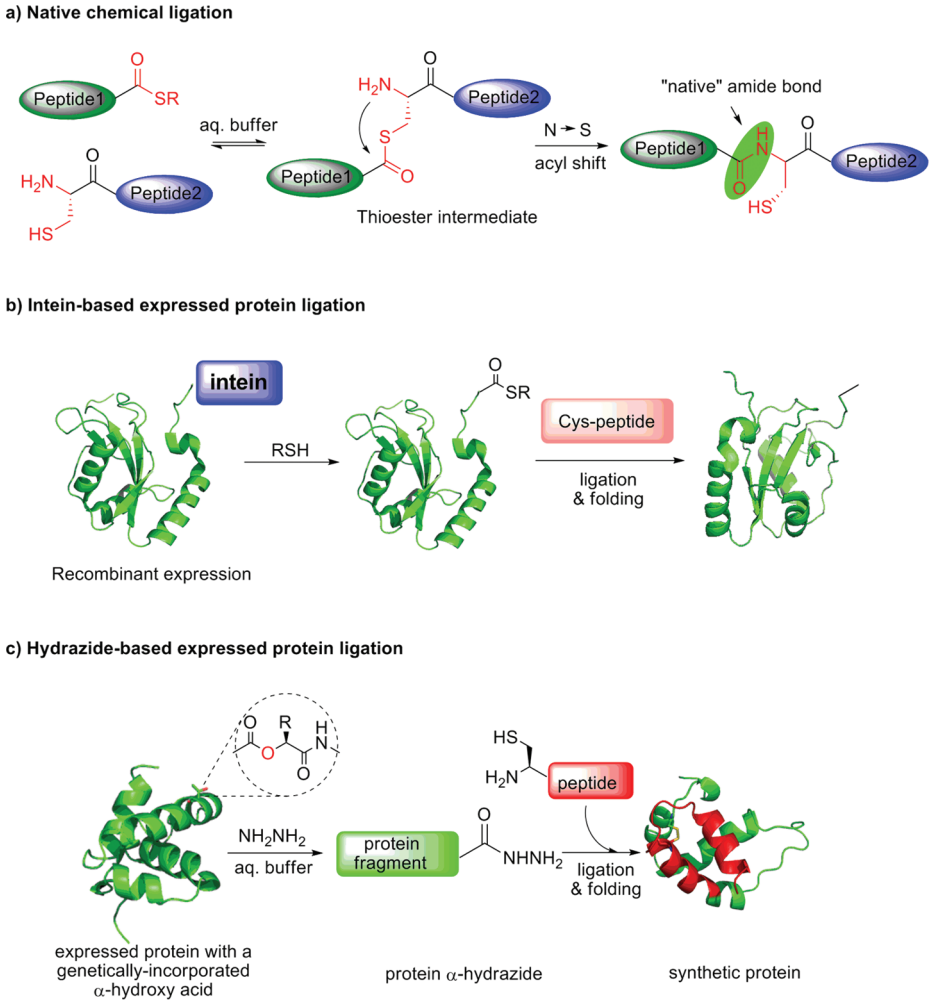
In the present study we wish to describe an alternative new version of EPL that uses recombinant protein α-hydrazides produced *via* genetic incorporation of an α-hydroxy acid (Scheme 1). This method is developed on the basis of our recent finding that peptide α-hydrazides may be used as the reagent equivalent of a "thioester synthon" in NCL.[20−24] An important advantage of the hydrazide-based ligation approach is that peptide α-hydrazides can be readily prepared by using both Boc and Fmoc solid-phase peptide synthesis. To further expand the scope and utility of this approach, we now report that protein α-hydrazides can be readily produced through hydrazinolysis of a protein bearing a site-specifically incorporated oxoester in its backbone.[25−28] Such production of recombinant protein α-hydrazides allows for the development of a conceptually new version of EPL that does not involve protein splicing. The new EPL method may also provide several utilities complementary to the intein-based technology for protein semisynthesis. In addition, the hydrazide-based EPL method represents a new and useful application for the genetic code expansion technology pioneered by Schultz *et al.*[29−36]

RESULTS AND DISCUSSION

Genetic Incorporation of α-Hydroxy Acids. We started by introducing an oxoester moiety into the backbone of

Received: January 16, 2012
Accepted: March 16, 2012
Published: March 16, 2012

Scheme 1. Native Chemical Ligation, Intein-Based Expressed Protein Ligation, and Hydrazide-Based Expressed Protein Ligation



proteins *via* an α-hydroxy acid. This had been achieved by Schultz *et al.*, who used the genetic code expansion method to incorporate *p*-hydroxy-L-phenyllactic acid into proteins with an evolved mutant tyrosyl-tRNA synthetase from *M. jannaschii*.[37] More recently, Yokoyama *et al.* also successfully incorporated the α-hydroxy acid analogue of Boc-lysine into protein backbones by using the wild type *M. mazei* pyrrolysyl-tRNA synthetase (PylRS) and its cognitive $tRNA^{Pyl}_{CUA}$.[38] This PylRS-$tRNA^{Pyl}_{CUA}$ pair, encoding the 22nd naturally occurring amino acid pyrrolysine in response to an amber codon in archaea species, has recently been adapted for incorporation of a variety of unnatural amino acids into proteins in bacteria, yeast, and mammalian cells.[39−42] Such an approach is more advantageous than the previous systems because it works as an "one-stop shop" for introducing desired functionalities into the side chain and main chain of proteins in both prokaryotic and eukaryotic cells.[43] The results we obtained on proteins produced in *E. coli* cells could therefore be easily expanded into proteins expressed in mammalian cells. Here, we decided to employ this PylRS-$tRNA^{Pyl}_{CUA}$ system for introducing the oxoester moiety into the backbone of proteins and first focused on improving the incorporation efficiency of α-hydroxy acid analogues of pyrrolysine. We have recently evolved a mutant PylRS, ACPK-RS, from *M. barkeri* that can encode an azide-bearing pyrrolysine analogue with enhanced efficiency compared to the previous linear azide analogues.[44] In this work, we tested the incorporation efficiency of ACPK-RS with four different α-hydroxy acid analogues. The *E. coli* acid chaperone HdeA, which plays essential roles in supporting the acid resistance of *E. coli* cells by suppressing the aggregation of denaturated periplasmic proteins at pH < 3,[45,46] was used as a model protein in our study.

As shown in Figure 1 and Supplementary Figure S2, for optimizing the best incorporation site, four α-hydroxy acids (**1**−**4**) were tested for incorporation into HdeA at different backbone sites including S27, A38, V58, A61, and I85. Protein expression was carried out in BL21-DE3 cells co-transformed with orthogonal plasmids expressing the ACPK-RS-$tRNA^{Pyl}_{CUA}$ pair and HdeA harboring an inframe amber mutation (TAG). The medium was supplemented with or without the α-hydroxy acid (**1**−**4**) at 1 mM final concentration. Protein was purified by Ni-NTA column and concentrated by using centrifugal filters. Through SDS-PAGE analysis we found that compounds **1**−**4** can all be recognized by the ACPK-RS-$tRNA^{Pyl}_{CUA}$ system, and compound **3** (designated as BKOH below) showed the highest expression efficiency at the V58 backbone site. This conclusion was also confirmed by the Western blot analysis (Supporting Information). The isolated yield of the full-length HdeA-myc-His_6 containing BKOH at the V58 position (*i.e.*, HdeA-myc-His_6-V58BKOH) was estimated to be 3−4 mg L^{-1}.

dx.doi.org/10.1021/cb300020s | *ACS Chem. Biol.* 2012, 7, 1015−1022

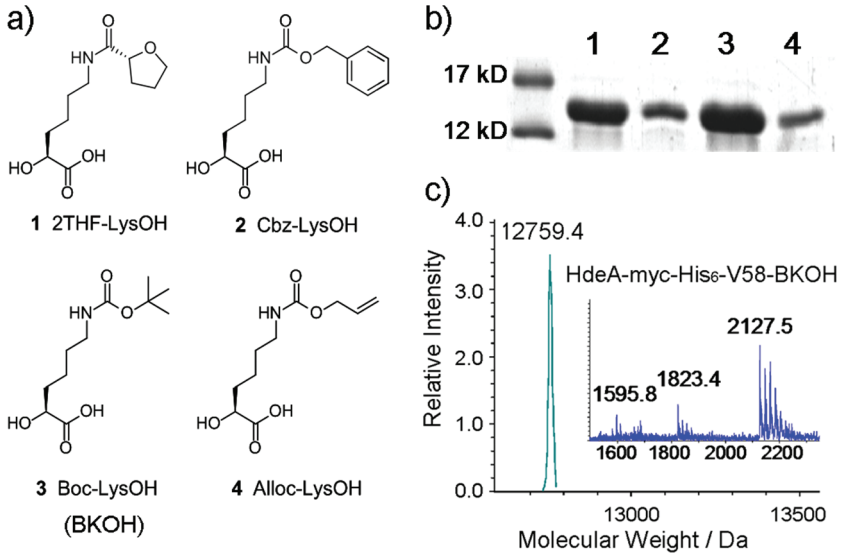


Figure 1. (a) Structures of the tested α-hydroxy acids. (b) SDS-PAGE analysis showing the efficiency for incorporation of compounds **1**–**4** into HdeA in *E. coli*. (c) ESI-MS analysis of the full-length HdeA-myc-His_6-V58BKOH with BKOH incorporated at residue V58 (theoretical molecular mass = 12,760 Da).

This yield appears to be higher than that of BKOH incorporated into GST protein as reported in the previous study.[38] The purity of expressed HdeA-V58BKOH was confirmed by HPLC and ESI-MS analysis (Supplementary Figure S1). Note that HdeA is a relatively small protein, and the expression yield of wild type HdeA is about 30 mg L^{-1}. Thus, the relatively high incorporation efficiency of BKOH can be attributed to two reasons: use of an evolved mutant PylRS and the high expression yield of wide type HdeA protein.

Hydrazinolysis of Protein Oxoester. Having confirmed the feasibility of obtaining recombinant protein oxoesters, we next examined the conversion of the protein oxoester to its corresponding α-hydrazide.[25−28] For this purpose, we treated the aforementioned HdeA-myc-His_6-V58BKOH protein with 4% (v/v) aqueous hydrazine (pH 9.0). After 1 h of incubation at RT, both HPLC (Supporting Information) and SDS-PAGE analyses show that the cleavage reaction was completed, producing two fragments (which can be readily separated by reverse-phase HPLC) with ESI-MS masses of 6109 and 6684 Da, respectively (Figure 2). These observations indicate that a clean and site-specific hydrozinolysis occurs at the 58 residue producing an α-hydrazide fragment HdeA(1−57)-$NHNH_2$ and a BKOH-capped peptide HdeA(58−89)-myc-His_6. The yield of the purified peptide hydrazide (as lyophilized powder) was measured to be about 1−2 mg L^{-1}.

Ligation of Protein α-Hydrazide. To test the ligation of the expressed protein α-hydrazide (*i.e.*, HdeA(1−57)-$NHNH_2$), we used the standard Fmoc solid-phase synthesis method to make a 32-mer peptide HdeA(V58C-89). In this peptide the V58 residue was mutated to Cys to facilitate the ligation. Subsequently, the recombinant HdeA(1−57)-$NHNH_2$ (1.0 equiv) and synthetic HdeA(V58C-89) (1.4 equiv) fragments were added together to the aqueous phosphate (0.2 M) buffer containing 6.0 M guanidinium chloride. At pH 3.0 and −10 °C, $NaNO_2$ (10 mM in final concentration) was added to the ligation mixture. After 20 min, MPAA (4-mercaptophenylacetic acid) was added (100 mM in final concentration) and the pH value was adjusted to 7.0. The ligation was then allowed to proceed for 4 h at RT. As shown in Figure 3, the full-length HdeA(1−89) was successfully obtained in 80% HPLC yield with a correct ESI-MS mass. Note that we also observed the formation of a small amount of

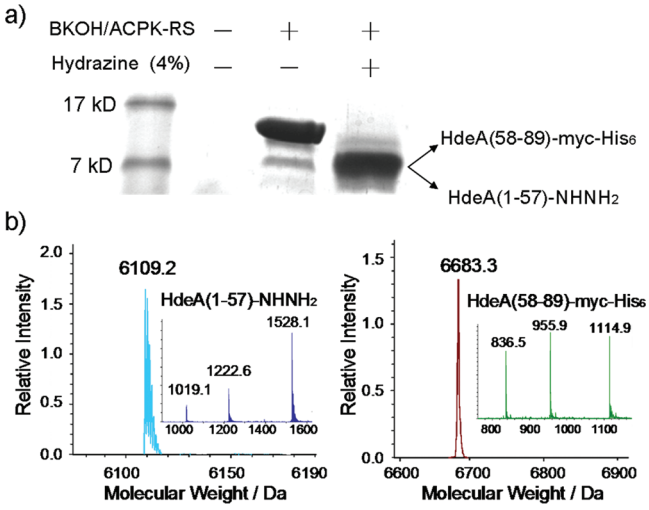


Figure 2. (a) SDS-PAGE analysis of the expression of HdeA-myc-His_6-V58BKOH and hydrazinolysis of the protein in the absence and presence of 4% (v/v) aqueous hydrazine for 60 min at pH 9.0. (b) ESI-MS spectra of the two fragments of HdeA-myc-His_6-V58-BKOH after hydrazinolysis. The theoretical molecular mass values of the two fragments are 6109 and 6684 Da, respectively.

a thiolactone byproduct. Thiolactone is produced by to the intramolecular reaction of the Cys thiol group of a peptide-thioester with its own thioester end. Thus thiolactone is a cyclic thioester, but it is usually less reactive in native chemical ligation.[47] Because HdeA(1−58)−$NHNH_2$ contains a Cys residue at position 18, it is not surprising that a thiolactone byproduct is formed (see Supplementary Figure S12).

Folding and Bioactivity. The folding of the synthetic HdeA(1−89) peptide was found to occur automatically in PBS buffer at pH 7. The Native PAGE gel in Figure 3 shows that the synthetic HdeA(1−89) forms a dimer at pH 7.0 in the same way as the recombinant HdeA-myc-His_6. On the other hand, the SDS-PAGE gel shows that both the synthetic HdeA(1−89) and recombinant HdeA-myc-His_6 change back to the monomer form under the denaturing conditions (Figure 3). Furthermore, the CD spectra of both synthetic HdeA(1−89) and recombinant HdeA-myc-His_6 superimpose fairly well (Figure 4),

dx.doi.org/10.1021/cb300020s | *ACS Chem. Biol.* 2012, 7, 1015−1022

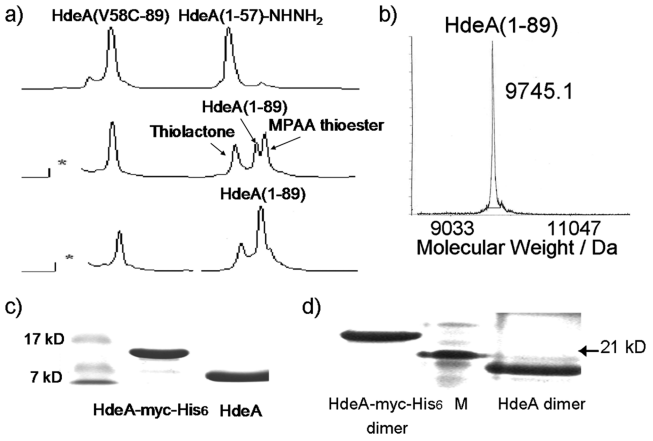


Figure 3. (a) Analytical HPLC trace (215 nm) for reaction mixture before and after the ligation. (b) MALDI spectra analysis of ligation product HdeA(1−89). (c) SDS-PAGE of recombinant HdeA-myc-His_6 and HdeA(1−89). (d) Native PAGE gel confirming the formation of HdeA-myc-His_6 and HdeA(1−89) dimer.

indicating that they fold to a similar three-dimensional structure.

To determine the biological activities of synthetic HdeA(1−89), we examined its interaction with the SurA protein, which has recently been identified as the *in vivo* client protein of HdeA.[48] Such interaction can be evaluated by measuring SurA's aggregation propensity in the presence and absence of synthetic HdeA(1−89) as compared to the recombinant HdeA-myc-His_6. As shown in Figure 4, in the absence of either synthetic HdeA(1−89) or recombinant HdeA-myc-His_6, a low pH value (pH = 2) causes SurA to fully aggregate. On the other hand, when either synthetic HdeA(1−89) or recombinant HdeA-myc-His_6 is added, a significant fraction of SurA becomes soluble both at pH 2 and after subsequent pH neutralization. These observations demonstrate that the synthetic HdeA(1−89) has similar biological activity as its recombinant counterpart. Besides, to demonstrate that Cys mutation at position 58 does not alter the structure or function of HdeA, HdeA(V58C) was expressed and tested by Native gel and SurA analysis (Supplementary Figures S12 and S13). The results indicate that the structure and biological function of the V58C mutant are similar to that of the wide type HdeA.

Discussion. At this point we have demonstrated that the ligation of expressed protein α-hydrazides can be used to synthesize proteins with correctly folded structures and full biological activity. The same synthetic target can be handled by the intein-based technology. Nonetheless, as shown below, the new hydrazide-based EPL strategy may sometimes provide advantages that are complementary to the intein-based method.

First, it has been reported that in the application of intein-based EPL, some C-terminal residues such as Pro should be avoided because they could lead to a low yield of protein α-thioesters.[49] This particular problem can be solved by the hydrazide-based EPL method because the hydrazinolysis reaction is not found to be sensitive to the chemical nature or steric hindrance of the cleavage site. For instance, we can easily incorporate BKOH at the A61 position in the HdeA protein (Figure 5). Treatment of the recombinant HdeA-A61BKOH with aqueous hydrazine (4%) (v/v) produces HdeA(1-P60)-$NHNH_2$ with a C-terminal Pro residue in *ca.* 90% yield. Our previous study already showed that the C-terminal Pro residue would not cause any problem in the hydrazide-based ligation reaction.[20]

Second, in the application of the intein-based EPL the ligation site is usually not chosen within the secondary or tertiary structure of the protein. The reason is that protein fragments dissected between modular domains often show poor handling properties (*e.g.*, being insoluble) and therefore are difficult to express.[50] This potential problem can be solved using the hydrazide-based EPL method, which can always rely on the expression of full-length and therefore well-behaved proteins. As a good illustration of this problem, we have attempted to express the HdeA(1−37) peptide fused to the C-terminal gyrA intein-chitin binding domain (CBD). Because the A37 residue locates at the middle of the protein's tertiary structure, the target HdeA(1−37)-intein-CBD was not obtained as a soluble protein as

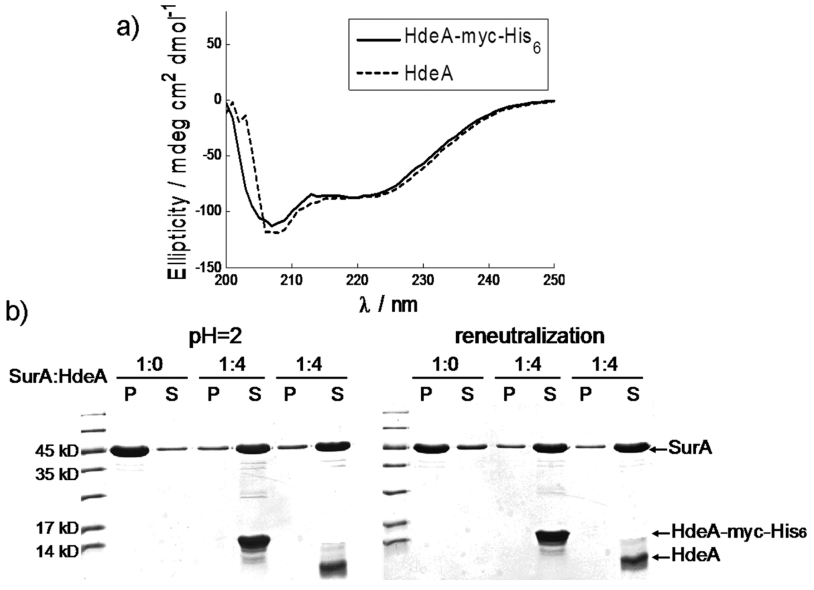


Figure 4. (a) CD spectrum analysis of synthetic HdeA compared with recombinant HdeA-myc-His_6. (b) HdeA suppresses the aggregation of its client protein-SurA at a low pH and after the pH neutralization. SDS-PAGE analysis was followed by coomassie blue staining of the soluble supernatant (S) and aggregated pellet (P) protein in each sample.

dx.doi.org/10.1021/cb300020s | *ACS Chem. Biol.* 2012, 7, 1015−1022

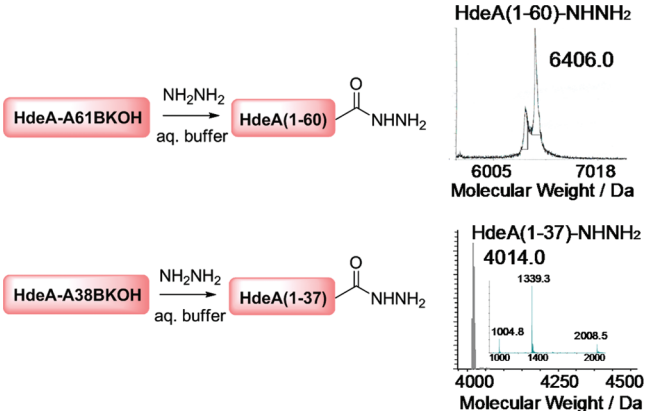


Figure 5. ESI-MS analysis of the HdeA fragments with C-terminal Pro-$NHNH_2$ and Glu-$NHNH_2$. The theoretical molecular mass values of the two fragments are 6406 and 4014 Da, respectively.

it ended up in inclusion bodies. In contrast, HdeA(1-A37)-$NHNH_2$ was easily obtained by the expression and hydrazinolysis of HdeA-A38BKOH as a well-behaved protein (Figure 5). Thus, the hydrazide-based EPL method allows for the selection of ligation site with more flexibility.

Admittedly, it should be noted that in the intein-based EPL method the ligation sites are often intentionally placed between protein domains or in flexible loops due to protein folding considerations. The reason is that when preparing analogues of large proteins, it is often difficult to find conditions for efficient *in vitro* refolding. To circumvent this problem two protein modules are expressed and folded *in vivo* and subsequently ligated under non-denaturing conditions, thus obviating the refolding step. Thus a key advantage of the intein-based EPL approach is that the reaction can be performed under non-denaturing conditions, and this is crucial when working with large proteins. In contrast, in the method reported here, the protein fragments have to be repeatedly exposed to denaturing conditions. Such conditions may create problems for the synthesis of proteins with multiple domains.

In summary, we have reported a new alternative version of expressed protein ligation by combining the recently developed technologies of hydrazide-based protein ligation and genetic code expansion. This method does not involve protein splicing and generates recombinant protein α-hydrazides as ligation intermediates that are more chemically stable than protein α-thioesters. By using HdeA as a model protein we demonstrated that the hydrazide-based EPL method can be used to synthesize proteins with correctly folded structures and full biological activity. Because this PylRS-$tRNA^{Pyl}_{CUA}$ system is compatible with both prokaryotic and eukaryotic cells, the strategy we presented here may be readily expanded to manipulate proteins produced in mammalian cells. Moreover, the new hydrazide-based EPL method supplements the intein-based EPL by allowing for a more flexible selection of ligation site. Thus the availability of a new alternative version of EPL technology is expected to improve the flexibility and power of protein semisynthesis.

METHODS

General Information. All reagents and solvents were purchased commercially and were purified when necessary. THF and Et_2O were distilled from sodium/diphenyl ketone immediately prior to use. DMF was distilled under reduced pressure from sodium sulfate and stored over 4 Å molecular sieves. CH_2Cl_2, pyridine, and Et_3N were distilled from calcium hydride immediately prior to use. All other commercially available reagents and solvents were used as received without further purification unless otherwise indicated. All organic extracts were dried over sodium sulfate or magnesium sulfate and concentrated using a rotary evaporator. TLC was carried out on the plates precoated with silica gel 60 F254 (250 layer thickness). Visualization was accomplished using UV light, iodine vapors, ninhydrin solution, permanganate solution and/or phosphomolybdic acid solution. Flash column chromatographic purification of products was accomplished using forced-flow chromatography on silica gel (300−400 mesh on large-scale or 200−300 mesh on small-scale).

Analytical HPLC was run on a SHIMADZU (Prominence LC-20AT) instrument using an analytical column (Grace Vydac Protein C18, 250 mm × 4.6 mm, 5 μm particle size, flow rate 1.0 mL min^{-1}, rt). Analytical injections were monitored at 214 and 254 nm. Semipreparative HPLC was run on a SHIMADZU (Prominence LC-20AT) instrument using a semipreparative column (Grace Vydac Peptide C18, 250 mm × 10 mm, 10 μm particle size, flow rate 3.0 mL min^{-1}). Solvent A was 0.08−0.1% (v/v) TFA in acetonitrile, and solvent B was 0.08−0.1% (v/v) TFA in water. Both solvents were filtered through 0.22 μm filter paper and sonicated for 20 min before use. MALDI-TOF mass spectra were measured on an Applied Biosystems 4700 Proteomics Analyzer 283. A solution of 10 mg mL^{-1} matrix α-cyano-4-hydroxy cinnamic acid containing 1:1 v/v (0.1% TFA in acetonitrile/0.1% (v/v) TFA in water) was used for generating the probe−matrix mixture. High-resolution ESI mass spectra were measured on a Bruker APEX IV Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. Normal ESI mass spectra were measured on a Bruker Daltonics DataAnalysis 3.0 workstation. ^1H and ^{13}C NMR spectra were recorded on an Oxford 300 MHz spectrometer in deuteriochloroform ($CDCl_3$) or deuterium oxide (D_2O) with the solvent residual peak ($CDCl_3$, 7.26 ppm (^1H), 77.23 ppm (^{13}H); D_2O, 4.79 ppm (^1H)) as internal reference unless otherwise stated. Data are reported in the following order: chemical shifts (δ); multiplicities indicated as br (broadened), s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), app (apparent); coupling constants, *J*, reported (Hz); integration is provided.

For molecular biology and biochemistry experiments, primers were ordered from Sangon Biotech (Shanghai) Co., Ltd. Enzymes were ordered from New England Biolabs (NEB). Bacterial cells were grown in LB (Luria−Bertani) broth or on LB agar medium (Sigma). Antibiotics were used at final concentrations of 100 μg mL^{-1} for ampicillin and kanamycin and 50 μg mL^{-1} chloramphenical (Sigma). All pictures of protein gels including comassie stained SDS-PAGE gel and Western blotting membrane were taken on a ChemDocXRS+ (Bio-Rad).

Fmoc SPPS. Screw-cap glass peptide synthesis reaction vessels were attained from commercial sources. Wang resin was initially swelled with DCM/DMF (1/1) for about 3 h. For preactivation of the first protected amino acid, 3.6 equiv of HBTU, 4 equiv HOBt, 8 equiv of DIEA, and 0.1 equiv of DMAP were added to a solution of 4 equiv of protected amino acid (0.4 M) in DMF. After preactivation for 2 min, the mixture was added to the resin. After 8 h the resin was washed by 5 × DMF, 5 × DCM, and 5 × DMF. Capping reagent was acetic anhydride/DIEA/DMF (1:1:8) (2 × 5 min). The resin was then washed by 5 × DMF, 5 × DCM, and 5 × DMF. The deprotection reagent was 20% (v/v) piperidine/DMF (5 min +10 min). The resin was washed with 5 × DMF, 5 × DCM, and 5 × DMF. The cleavage reagent chosen was Reagent K. Normally a TFA cocktail of TFA/phenol/water/thioanisole/EDT (82.5/5/5/5/2.5) was added to the dry resin prewashed with DCM under reduced pressure. After 2 h, the resin was washed with an equal volume of TFA once. The combined solutions were concentrated by blowing with N_2. The crude peptides were obtained by precipitation with cold ether and centrifugation at 5000 rpm. The residue was dissolved in 0.1% (v/v) TFA containing cosolvent of acetonitrile and water (1:1), analyzed by analytical HPLC plus MALDI-TOF/MS, purified by semipreparative HPLC, and subsequently lyophilized.

Protein Expression and Purification. HdeA protein expression was carried out in *E. coli* DH10B cells co-transformed with plasmids expressing PylRS-$tRNA^{Pyl}_{CUA}$ pair and HdeA-V58TAG. Cells were grown in LB medium containing ampicillin (100 μg mL^{-1}) and

dx.doi.org/10.1021/cb300020s | *ACS Chem. Biol.* 2012, 7, 1015−1022

chloramphenicol (50 μg mL^{-1}) with shaking overnight at 37 °C. After 1:100 dilution in LB medium containing ampicillin (100 μg mL^{-1}) and chloramphenicol (50 μg mL^{-1}), the culture was grown at 37 °C to an OD_{600}~0.5. Compound **3** or other three hydroxy acids (100 mM for stock solution) was added by 1:100 dilution to a final concentration of 1 mM and incubated with cell culture for 30 min. Protein expression was induced by the addition of arabinose and IPTG to the final concentration of 0.2% (v/v) and 0.5 mM, respectively. After expression for 12 h, cells were harvested by centrifugation (10,000*g*, 10 min) and resuspended in lysis buffer (20 mM Tris-HCl, 300 mM NaCl, pH 7.4). Bacterial lysate after sonication was centrifugated (16,000*g*) for 30 min, and the supernatant was loaded onto a Ni-NTA column (Histrap 5 mL, GE Healthcare). The column was washed with 30 mL of washing buffer (20 mM Tris-HCl, 300 mM NaCl, pH 7.4 with 35 mM imidazole) and then eluted with 15 mL of elution buffer (20 mM Tris-HCl, 300 mM NaCl, pH 7.4 with 250 mM imidazole). Then the eluted protein was concentrated by using centrifugal filters (10,000 MW cutoff) and further purified process RP-HPLC. The yield of the protein (as solid powder) was quantified after lyophilization.

For HdeA(1−37)-intein-CBD cloning and expression, the PCR amplified DNA was purified by agarose gel extraction (QIAquick kit), digested with NdeI and BspQI, and then ligated into a NdeI, BspQI digested pTXB1 bacterial expression vector (New England Biolab, Beverly, MA). Then, *E. coli* ER2566 (NEB) cells transformed with the pTXB1-HdeA(1−37) plasmid were grown to midlog phase (OD_{600} = 0.6) in 1 L of LB medium and induced with 0.2 mM IPTG at 16 °C overnight or 0.5 mM IPTG at 30 °C for 8 h. Cells were harvested by ultrasonification and resuspended in lysis buffer containing 20 mM HEPES (pH 7.4), 500 mM NaCl, 0.1% Triton X-100 (v/v), and 1 mM EDTA. After cell lysis, the supernatant and aggregated pellets were analyzed by Western blotting.

Hydrazinolysis. The hydrazinolysis buffer was prepared by dilution of 80% (v/v) hydrazine hydrate in PBS (6 M Gn-HCl, 200 mM Na_2HPO_4, pH 9.0). pH was adjusted with concentrated HCl. The recombinant peptide ester was purified by HPLC, lyophilized, and dissolved in hydrazinolysis buffer. After 2 h, an aliquot of the reaction buffer was mixed with an equal volume of pH-neutral 5% TCEP and analyzed by HPLC. The hydrazinolysis condition was further optimized as mentioned above.

Ligation of Protein α-Hydrazide. Synthetic peptide HdeA-(V58C-89) (3.1 mg, 0.95 μmol, 1.4 equiv) and expressed protein HdeA(1−57)-$NHNH_2$ (4.3 mg, 0.70 μmol, 1 equiv) were dissolved in 400 μL of ligation buffer containing 40 μL of internal standard buffer. The reaction mixture was held in 1.5 mL EP tube at −10 °C in an ice−salt bath. Then, 40 μL of oxidative solution (200 mM) was added dropwise, and the reaction mixture was stirred for 20 min in ice−salt bath. Next, 400 μL of thiol solution (200 mM) was added, and pH was carefully adjusted to 7.0 with NaOH (2.0 M). The reaction mixture was stirred at RT then. The reaction reached completion within 5 h (HPLC yield 72%). The ligation products were collected by semipreparative HPLC and lyophilized to white solid (1.9 mg, 28%). Ligation buffer: 6.0 M Gn-HCl, 200 mM Na_2HPO_4, carefully adjusted to pH 3.0 with NaOH (2.0 M) or concentrated HCl. Internal standard buffer: 9.9 mg $BzNH_2$, 2.0 mL ligation buffer, pH 7.0. Oxidative solution: $NaNO_2$ (41.4 mg, 0.6 mmol) in neat water (3.0 mL). Thiol solution: MPAA (33.6 mg, 0.2 mmol) in neutral ligation buffer (1 mL). All reactions were monitored by HPLC, and the ligation yield was determined according to the internal standard, $BzNH_2$. Before analysis, the reaction solution was reduced by TCEP (30 mM, pH 7.0) for 3 min.

SDS-PAGE, Native PAGE Gel, and Western Blotting Analysis. For SDS-PAGE, samples were loaded onto 15% SDS-PAGE gels and electrophoresed for 30 min at 80 V and 50 min at 150 V. The Native PAGE gel was prepared using Bio-Rad Mini-PROTEIN Tetra Electrophoresis System. SDS was removed from the ingredients of both the stacking gel (pH 6.8, 4%) and the resolving gel (pH 8.8, 15%). All protein samples were prepared in Tris buffer containing bromophenol blue. Native running buffer (1 L) consisted of 14.4 g glycine and 3.03 g Tris base dissolved in ddH_2O. Samples were run under ice-cold condition (150 V, 400 mA, 60 min). Soybean trypsin inhibitor (from Beijing Biodee Biotechnology Co., Ltd., pI = 4.6, MW ~21 kDa) was used as the marker protein (pI of HdeA = 5.2). Semisynthesis HdeA(1−89) was confirmed to form dimer in solution by Native PAGE gel. For Western blotting analysis, the separated proteins were transferred onto polyvinylidene difluoride membranes (Bio-Rad) at 250 mA, 4 °C 2 h. After blocking at RT for 1 h in Blotto (5% v/v nonfat dry milk in 1×TBST), analysis of the membranes was performed using 1:1000 dilutions of monoclonal anti His-tag (Santa Cruz) or anti-CBD (NEB) as primary antibody overnight at 4 °C followed by incubation with horseradish peroxidase (HRP)-conjugated goat anti-mouse secondary antibody (Santa Cruz) at RT for 2 h. Antibodies were detected with ECL reagents (Pierce).

CD Spectroscopy. CD spectra were recorded on a Applied Photophysics Chirascan spectrometer from 250 to 200 nm in a 0.2 cm path length cell at RT. Each protein sample was dissolved in 20 mM phosphate buffer at pH 7.2 with a final concentration of 0.4 mg mL^{-1}. The result was recorded by averaging three scans and plotted as mean residue ellipticity [θ] (mdeg cm^2 dmol^{-1}).

SurA Assay. SurA protein was incubated at a concentration of 8 μM in buffer A (8 mM H_3PO_4, 150 mM KCl, and 150 mM $(NH_4)_2SO_4$, pH 2.0 at 37 °C for 30 min in the absence or presence of 32 μM HdeA. After 30 min, the pH was neutralized by adding 0.133 vol of buffer B (0.5 M sodium phosphate, pH 8). After 30 min at 37 °C, the soluble supernatant (S) and aggregated pellet (P) protein in each sample at the indicated HdeA:SurA ratios were separated by spinning at 12,000*g* at 4 °C using a SIGMA 3-18K centrifuge and then analyzed by SDS-PAGE.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional details about the experiments and compound spectra. This material is available free of charge *via* the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: pengchen@pku.edu.cn; lliu@mail.tsinghua.edu.cn.

Author Contributions

§These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the national "973" grants from the Ministry of Science and Technology (No. 2011CB965300 to L.L.; and No. 2010CB912300 to P.R.C.); and NSFC (No. 20932006 to L.L.; No. 91013005 to P.R.C.; No. 21102083 to Y.M.L.).

■ REFERENCES

(1) Kent, S. B. H. (2009) Total chemical synthesis of proteins. *Chem. Soc. Rev. 38*, 338−351.

(2) Nilsson, B. L., Soellner, M. B., and Raines, R. T. (2005) Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct. 34*, 91−118.

(3) Casi, G., and Hilvert, D. (2003) Convergent protein synthesis. *Curr. Opin. Struct. Biol. 13*, 589−594.

(4) Dawson, P. E., and Kent, S. B. H. (2000) Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem. 69*, 923−960.

(5) Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. H. (1994) Synthesis of proteins by native chemical ligation. *Science 266*, 776−779.

(6) Muir, T. W., Sondhi, D., and Cole, P. A. (1998) Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci. U.S.A. 95*, 6705−6710.

(7) Severinov, K., and Muir, T. W. (1998) Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *J. Biol. Chem. 273*, 16205−16209.
(8) Hahn, M. E., Pellois, J. P., Vila-Perello, M., and Muir, T. W. (2007) Tunable photoactivation of a post-translationally modified signaling protein and its unmodified counterpart in live cells. *ChemBioChem 8*, 2100−2105.
(9) Hahn, M. E., and Muir, T. W. (2004) Photocontrol of Smad2, a multiphosphorylated cell-signaling protein, through caging of activating phosphoserines. *Angew. Chem., Int. Ed. 43*, 5800−5803.
(10) Lee, H. M., Larson, D. R., and Lawrence, D. S. (2009) Illuminating the chemistry of life: design, synthesis, and applications of "caged" and related photoresponsive compounds. *ACS Chem. Biol. 4*, 409−427.
(11) Vila-Perello, M., Pratt, M. R., Tulin, F., and Muir, T. W. (2007) Covalent capture of phospho-dependent protein oligomerization by site-specific incorporation of a diazirine photo-cross-linker. *J. Am. Chem. Soc. 129*, 8068−8069.
(12) Triola, G., Waldmann, H., and Hedberg, C. (2012) Chemical biology of lipidated proteins. *ACS Chem. Biol. 7*, 87−99.
(13) Nikolov, D. B., Singla, N., Erdjument-Bromage, H., Himanen, J. P., and Muir, T. W. (2011) A semisynthetic Eph receptor tyrosine kinase provides insight into ligand-induced kinase activation. *Chem. Biol. 18*, 361−371.
(14) Chatterjee, C., McGinty, R. K., Fierz, B., and Muir, T. W. (2010) Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation. *Nat. Chem. Biol. 6*, 267−269.
(15) McGinty, R. K., Kohn, M., Chatterjee, C., Chiang, K. P., Pratt, M. R., and Muir, T. W. (2009) Structure−activity analysis of semisynthetic nucleosomes: Mechanistic insights into the stimulation of Dot1L by ubipuitylated histone H2B. *ACS Chem. Biol. 4*, 958−968.
(16) Wu, Y. W., Oesterlin, L. K., Tan, K. T., Waldmann, H., Alexandrov, K., and Goody, R. (2010) Membrane targeting mechanism of Rab GTPases elucidated by semisynthetic protein probes. *Nat. Chem. Biol. 6*, 534−540.
(17) McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G., and Muir, T. W. (2008) Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature 453*, 812−816.
(18) Fierz, B., Chatterjee, C., McGinty, R. K., Bar-Dagan, M., Raleigh, D. P., and Muir, T. W. (2011) Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nat. Chem. Biol. 7*, 113−119.
(19) Scheuermann, J. C., Alonso, A. G. D., Oktaba, K., Ly-Hartig, N., McGinty, R. K., Fraterman, S., Wilm, M., Muir, T. W., and Muller, J. (2010) Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. *Nature 465*, 243−247.
(20) Fang, G. M., Li, Y. M., Shen, F., Huang, Y. C., Li, J. B., Lin, Y., Cui, H. K., and Liu, L. (2011) Protein chemical synthesis by ligation of peptide hydrazides. *Angew. Chem., Int. Ed. 50*, 7645−7649.
(21) Yanaihara, N., Yanaihara, C., Dupuis, G., Beacham, J., Camble, R., and Hofmann, K. (1969) Studies on polypeptides. XLII. Synthesis and characterization of seven fragments spanning the entire sequence of ribonuclease T1. *J. Am. Chem. Soc. 69*, 2184−2185.
(22) Felix, A. M., and Merrifield, R. B. (1970) Azide solid phase peptide synthesis. *J. Am. Chem. Soc. 92*, 1385−1391.
(23) Wang, P., Layfield, R., Landon, M., Mayer, R. J., and Ramage, R. (1998) Synthesis of peptide C-terminal derivatives using the transfer active ester condensation technique. *Tetrahedron Lett. 39*, 8711−8714.
(24) Liao, Y., Kong, Y., Hu, N., Jin, Z., and Wang, P. (2010) Selective coupling at the alpha-amino group of cysteine using transfer active-ester-condensation technology to synthesize a linear octadecapeptide. *Chem. Lett. 39*, 196−197.
(25) Geiger, T., and Clarke, S. (1987) Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J. Biol. Chem. 262*, 785−794.
(26) Stephenson, R. C., and Clarke, S. (1989) Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. *J. Biol. Chem. 264*, 6164−6170.
(27) Kalia, J., and Raines, R. T. (2006) Reactivity of intein thioesters: appending a functional group to a protein. *ChemBioChem 7*, 1375−1383.
(28) Thom, J., Anderson, D., McGregor, J., and Cotton, G. (2011) Recombinant protein hydrazides: Application to site-specific protein PEGylation. *Bioconjugate Chem. 22*, 1017−1020.
(29) Wang, L., Xie, J. M., and Schultz, P. G. (2006) Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct. 35*, 255−249.
(30) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem. 79*, 413−444.
(31) Chin, J. W. (2011) Reprogramming the genetic code. *EMBO J. 30*, 2312−2324.
(32) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of *Escherichia coli*. *Science 292*, 498−500.
(33) Wang, Q., and Wang, L. (2008) New methods enabling efficient incorporation of unnatural amino acids in yeast. *J. Am. Chem. Soc. 130*, 6066−6067.
(34) Takimoto, J. K., Dellas, N., Noel, J. P., and Wang, L. (2011) Stereochemical basis for engineered pyrrolysyl-tRNA synthetase and the efficient in vivo incorporation of structurally divergent non-native amino acids. *ACS Chem. Biol. 6*, 733−743.
(35) Johnson, D. B., Xu, J., Shen, Z., Takimoto, J. K., Schultz, M. D., Schmitz, R. J., Xiang, Z., Ecker, J. R., Briggs, S. P., and Wang, L. (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol. 7*, 779−86.
(36) Neumann, H., Wang, K. H., Davis, L., Garcia-Alai, M., and Chin, J. W. (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature 464*, 441−444.
(37) Guo, J. T., Wang, J. Y., Anderson, J. C., and Schultz, P. G. (2008) Addition of an alpha-hydroxy acid to the genetic code of bacteria. *Angew. Chem., Int. Ed. 47*, 722−725.
(38) Kobayashi, T., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2009) Recognition of non-alpha-amino substrates by pyrrolysyl-tRNA synthetase. *J. Mol. Biol. 385*, 1352−1360.
(39) Chen, P. R., Groff, D., Guo, J. T., Ou, W. J., Cellitti, S., Geierstanger, B. H., and Schultz, P. G. (2009) A facile system for encoding unnatural amino acids in mammalian cells. *Angew. Chem., Int. Ed. 48*, 4052−4055.
(40) Hancock, S. M., Uprety, R., Deiters, A., and Chin, J. W. (2010) Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J. Am. Chem. Soc. 132*, 14819−14824.
(41) Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008) Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat. Chem. Biol. 4*, 232−234.
(42) Hao, Z. Y., Hong, S. L., Chen, X., and Chen, P. R. (2011) Introducing bioorthogonal functionalities into proteins in living cells. *Acc. Chem. Res. 9*, 742−751.
(43) Gautier, A., Nguyen, D. P., Lusic, H., An, W. A., Deiters, A., and Chin, J. W. (2010) Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc. 132*, 4086−4088.
(44) Hao, Z. Y., Song, Y. Q., Lin, S. X., Yang, M. Y., Liang, Y. J., Wang, J., and Chen, P. R. (2011) A readily synthesized cyclic pyrrolysine analogue for site-specific protein "click" labeling. *Chem. Commun. 47*, 4502−4504.
(45) Gajiwala, K. S., and Burley, S. K. (2000) HdeA, a periplasmic protein that supports acid resistance in pathogenic enteric bacteria. *J. Mol. Biol. 295*, 605−612.
(46) Hong, W. Z., Jiao, W., Hu, J., Zhang, J., Liu, C., Fu, X., Shen, D., Xia, B., and Chang, Z. Y. (2005) Periplasmic protein HdeA exhibits chaperone-like activity exclusively within stomach pH range by transforming into disordered conformation. *J. Biol. Chem. 280*, 27029−27034.
(47) Durek, T., Torbeev, V. Y., and Kent, S. B. (2007) Convergent chemical synthesis and high-resolution x-ray structure of human lysozyme. *Proc. Natl. Acad. Sci. U.S.A. 104*, 4846−4851.

(48) Wu, Y. E., Hong, W. Z., Liu, C., Zhang, L., and Chang, Z. Y. (2008) Conserved amphiphilic feature is essential for periplasmic chaperone HdeA to support acid resistance in enteric bacteria. *Biochem. J. 412*, 389−397.
(49) *Protein Expression Analysis Impact Kit Instruction Manual* (2009) New England Biolabs Inc., Ipswich, MA.
(50) Komarov, A. G., Linn, K. M., Devereaux, J. J., and Valiyaveetil, F. I. (2009) Modular strategy for the semisynthesis of a K^+ channel: Investigating interactions of the pore helix. *ACS Chem. Biol. 4*, 1029−1038.