Simultaneous and site-directed incorporation of an ester linkage and an azide group into a polypeptide by *in vitro* translation†

Martin Humenik, Yiwei Huang, Igor Safronov and Mathias Sprinzl*

Received 11th May 2009, Accepted 17th July 2009
First published as an Advance Article on the web 12th August 2009
DOI: 10.1039/b909188b

A method is presented by which an azide-containing side chain can be introduced into any internal position of a polypeptide chain by *in vitro* translation. For this, 2′-deoxy-cytidylyl-(3′→5′)-adenosine was acylated on the 3′(2′)-hydroxyl group of adenosine with 6-azido-2(*S*)-hydroxyhexanoic acid (AHHA), an α-hydroxy- and ε-azide derivative of L-lysine. The acylated dinucleotide was enzymatically ligated with a tRNA transcript to provide chemically stable *E. coli* suppressor AHHA-tRNA$^{Cys(CUA)}$. The esterase 2 gene from *Alicyclobacillus acidocaldarius* was modified by the amber stop codon (UAG) on position 118. Using AHHA-tRNA$^{Cys(CUA)}$ in an *E. coli in vitro* translation/transcription system, the site-directed introduction of an azide group linked to a backbone ester into the esterase polypeptide was achieved. The yield of the synthesized modified protein reached 80% compared to translation of the native esterase. Subsequently, azide coupling with an alkyne-modified oligodeoxynucleotide demonstrated the feasibility of this approach for conjugation of polypeptides.

Introduction

The nonsense suppression of the amber codon (UAG) by a suppressor tRNA$^{(CUA)}$ is used *in vivo* and *in vitro* for incorporation of unnatural amino acids (uAAs) into proteins (neoproteins).[1,2] Aminoacyl-tRNA synthetase (ARS) variants capable of aminoacylating orthogonal suppressor tRNA$^{(CUA)}$ with uAA are applied *in vivo*.[3] Many incorporated uAAs carried functional groups, which can be further modified *e.g.* with fluorescence labels, biotin and polyethylene glycols.[4] However, the protein concentrations obtained in this system (5–20 μg/mL) as well as the suppression efficiencies (25–50%, yield of neoprotein compared to the yield of a native protein) are low.[2,5,6] Several limitation complicate further improvements of the *in vivo* system. First, an unnatural amino acid has to adhere to requirements for cellular uptake (mainly lipophilic, uncharged residues are favored)[7] and cytotoxicity.[8] Second, the competition of suppressor tRNA$^{(CUA)}$ with release factor 1 (RF1) for the amber stop codon negatively influences the suppression efficiency. Improvements achieved by increasing tRNA$^{(CUA)}$ gene copy[9] or by using an orthogonal ribosome-mRNA pair,[6] were partially successful but a profound manipulation of involved translation factors can substantially disturb cell physiology.

In vitro translation (IVT) systems use an orthogonal aminoacyl-tRNA$^{(CUA)}$ prepared in a combination of organic and enzymatic synthesis.[10] Since the prerequisite of cell viability is irrelevant in this case, an IVT method offers broader control over the translation components. Several attempts to improve translation efficiency have been made. Mild heating of a thermolabile RF1 mutant-containing *E. coli* IVT mixture showed increased protein production due to decreased competition between RF1 and the suppressor tRNA.[11] However, yields were still less than 20 μg/mL and suppression efficiencies (24–68%) depended on the incorporated uAA. Purified translation systems, with all the components required for proper translation, have also been developed.[12] Although this method is especially attractive for producing proteins with multiple uAAs, achieved yields were lower than 80 μg/mL. The *in vivo* approach of engineered tRNA and ARS pairs was also adopted for cell-free translations with suppression efficiencies of 30–50%.[13-15]

Recently, we reported on esterase 2 from thermophilic bacteria *Alicyclobacillus acidocaldarius*, which can be produced at up to 200 μg/mL by an IVT system derived from *Escherichia coli*.[16] Deactivation of RF1 by specific antibodies allows translation of the UAG amber codon by seryl-tRNA$^{Ser(CUA)}$ in high yield.[17]

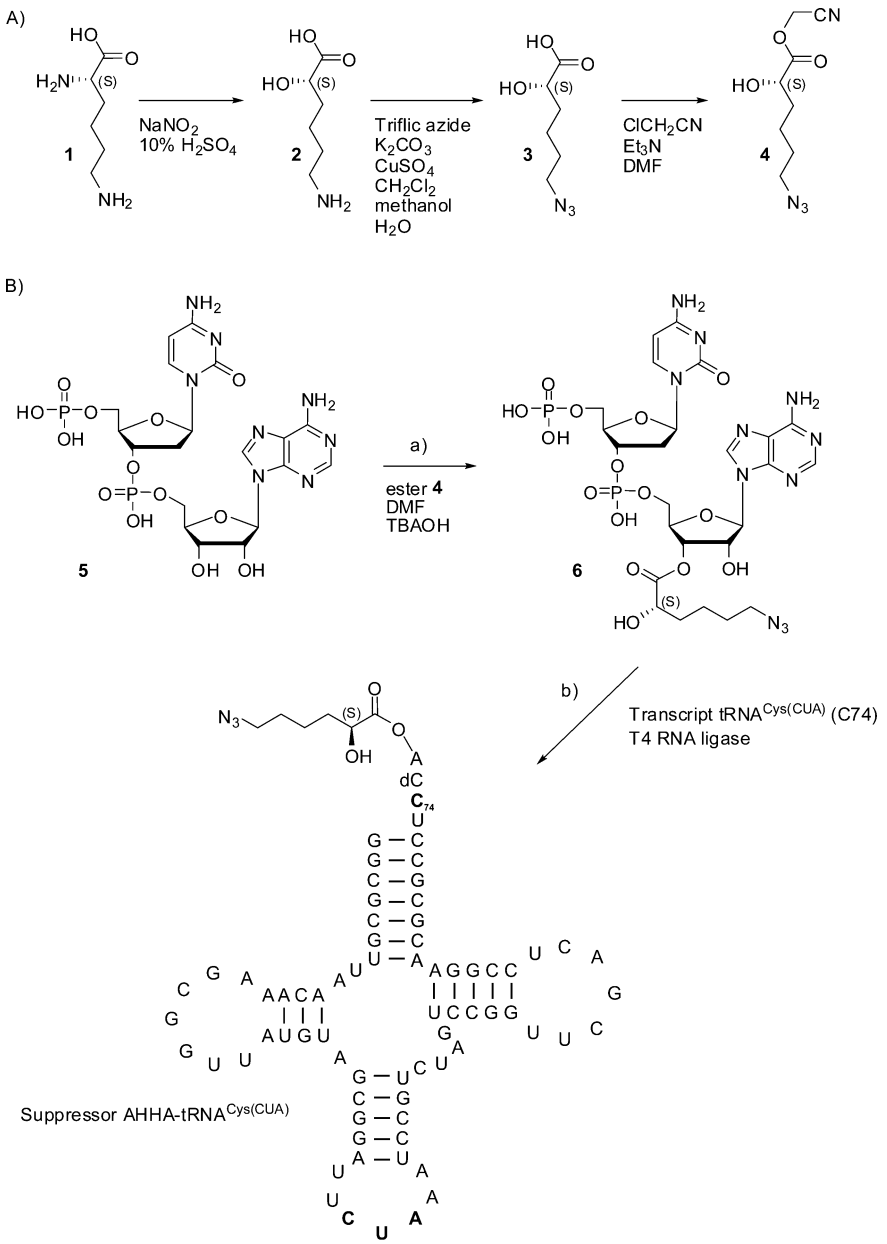
Here, we describe the application of this system for incorporation of potentially any uAA into polypeptides. As an example, suppression of the amber codon with orthogonal suppressor tRNA$^{Cys(CUA)}$ acylated by a α-hydroxy-ε-azido-derivative of L-lysine (**4**, Scheme 1) provided an azide group linked to a backbone ester of the neoprotein. Combination of these components leads to an *in vitro* system that affords an exceptional suppression efficiency and high fidelity of the neoprotein synthesis.

Results and discussion

L-Lysine **1** was modified in three steps, desamination,[18] amine–azide nucleophilic substitution[19] and carboxylic acid activation[20] (Scheme 1A). The cyanomethyl ester of 6-azido-2(*S*)-hydroxyhexanoic acid (**4**) was prepared in 52% overall yield. Dinucleotide 2′-deoxy-cytidylyl-(3′→5′)-adenosine (pdCpA, **5**) was prepared using phosphoramidite chemistry, acylated with cyanomethyl ester **4**[20,21] (Scheme 1B), purified by HPLC and characterized by NMR and MALDI-TOF spectroscopy (ESI†).

Laboratorium für Biochemie, Universität Bayreuth, Universitätstrasse 30, 95440 Bayreuth, Germany. E-mail: mathias.sprinzl@uni-bayreuth.de; Fax: +49-921-55-2066

† Electronic supplementary information (ESI) available: TOF MS ES+ spectrum, NMR structural assignment, ^{1}H NMR spectrum, HH-COSY NMR spectrum and ^{31}P NMR spectrum of compound **6**. See DOI: 10.1039/b909188b



Scheme 1 A) Preparation of cyanomethyl ester of α-hydroxy-ε-azido-modified L-lysine **4**. B) Preparation of suppressor AHHA-tRNA$^{Cys(CUA)}$ in two steps: (a) chemical acylation of dinucleotide pdCpA **5** on the 3′(2′) position of adenosine using activated ester **4** and (b) enzymatic ligation of acylated dinucleotide **6** to cytidine 74 of truncated transcript tRNA$^{Cys(CUA)}$ (C74) using T4 RNA ligase. The position of ligation (C_{74}) as well as the changed anticodon loop (CUA) are shown in bold letters in the secondary structure of the suppressor tRNA.

The ratio of 3′-O and 2′-O esters on the adenosine moiety of **6** was 2.3:1. tRNA$^{Cys(CUA)}$ charged with 6-azido-2(*S*)-hydroxyhexanoic acid (AHHA, **3**) was prepared in almost quantitative yield by enzymatic ligation of the truncated transcript of *E. coli* suppressor tRNA$^{Cys(CUA)}$ (C74) with pdCpA **6** (Scheme 1B, Fig. 1A). The suppressor AHHA-tRNA$^{Cys(CUA)}$ possesses several useful attributes for use in the *in vitro* translation. First, the ester bond of the α-hydroxycarboxylic acid attached to dinucleotide **6** or to the related tRNA is considerably more stable than the corresponding aminoacyl derivative.[22] This simplifies the synthetic protocols and increases the yield of the acylated suppressor tRNA.[20] Second, endogenous suppressor tRNAs used in *E. coli* IVT systems are more active in *in vitro* protein synthesis, as compared to exogenous suppressors.[23] AHHA-tRNA$^{Cys(CUA)}$ was derived from *E. coli* tRNA$^{Cys(GCA)}$. Substitutions in the anticodon loop (GCA for CUA) decrease the cysteinylation efficiency by three orders of magnitude.[24] The crystal structure of the complex between *E. coli* tRNA$^{Cys(GCA)}$ and CysRS revealed that the anticodon represents the essential recognition element.[25] This makes AHHA-tRNA$^{Cys(CUA)}$ (Scheme 1B) a suitable orthogonal candidate for an *E. coli* IVT system, *i.e.* it is not recognized by *E. coli* CysRS.

The formation of an aa-tRNA·EF-Tu·GTP ternary complex is essential in protein translation since this step controls the entry of aminoacyl-tRNAs to the ribosomal A-site.[26] The complex

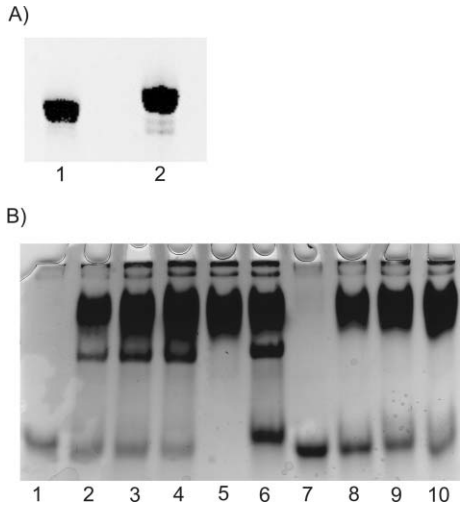


Fig. 1 A) Ligation of the truncated transcript $tRNA^{Cys(CUA)}$ (C74) with acylated dinucleotide **6**; lane 1: tRNA transcript; lane 2: ligation of truncated tRNA and **6** by T4 RNA ligase affording quantitative yield of suppressor AHHA-$tRNA^{Cys(CUA)}$; analyzed by urea-PAGE. B) EF-Tu·GTP complex formation with studied tRNAs; lane 1: suppressor AHHA-$tRNA^{Cys(CUA)}$, lanes 2, 3 and 4: AHHA-$tRNA^{Cys(CUA)}$ (3 μM) in the presence of 1, 2, and 3 μM EF-Tu·GTP, respectively, lane 5: EF-Tu, lane 6: native Phe-$tRNA^{Phe}$ (3 μM) and EF-Tu·GTP (3 μM), lane 7: uncharged $tRNA^{Cys(CUA)}$ pdCpA, lanes 8, 9 and 10: uncharged $tRNA^{Cys(CUA)}$ pdCpA (3 μM) in the presence of 1, 2 and 3 μM EF-Tu·GTP, respectively; analyzed by Native PAGE.

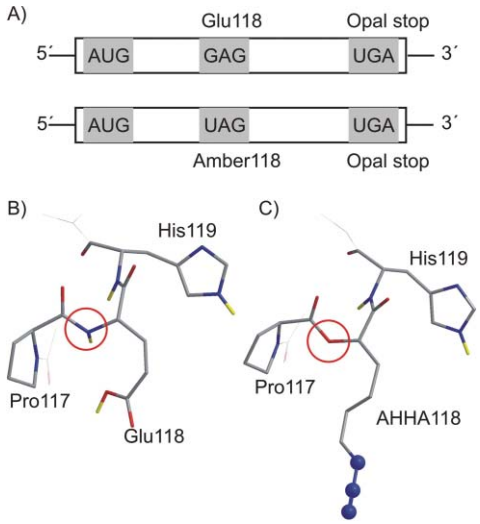


Fig. 2 A) Substitution of the sense glutamine codon (GAG) with the amber codon (UAG) in the position 118 of esterase 2 mRNA. B) Structure of Pro117, Glu118 and His119 in the esterase 2 (data taken from PDB ID 1EVQ). C) Possible structure of the esterase-neoprotein with introduced AHHA **3** on position 118. The substitution of a peptide bond with an ester bond in the protein backbone is highlighted by a red circle. The azide group linked to the backbone ester is depicted in blue. Other colors code for oxygen (red), carbon (grey), nitrogen (blue) and hydrogen on heteroatoms (yellow).

formation of acyl-tRNAs with EF-Tu·GTP was analyzed by polyacrylamide gel electrophoresis (PAGE), as shown in Fig. 1B. The AHHA-$tRNA^{Cys(CUA)}$ treated with different concentrations of EF-Tu·GTP (lanes 2–4) formed a complex comparable with that obtained with Phe-$tRNA^{Phe}$ (lane 6). As a control, nonacylated $tRNA^{Cys(CUA)}$pdCpA, prepared by ligation of dinucleotide **5** with truncated $tRNA^{Cys(CUA)}$ (C74), was used (lane 7). No binding of EF-Tu·GTP to uncharged $tRNA^{Cys(CUA)}$ (lanes 8–10) was observed.

The activity of the AHHA-$tRNA^{Cys(CUA)}$ was then tested in an *E. coli in vitro* transcription/translation system. For this, a plasmid pIVEX-Est2_Amb118 with the amber stop codon UAG 118 instead of GAG (Glu118) coding for esterase 2 from *Alicyclobacillus acidocaldarius* was prepared (Fig. 2). The amber stop codon, which is normally recognized by RF1, was assigned for suppressor AHHA-$tRNA^{Cys(CUA)}$. For the termination of translation, the opal codon (UGA), recognized by RF2, was used.

The formerly estimated optimal concentration for suppressor Ser-$tRNA^{Ser(CUA)}$, obtained by enzymatic aminoacylation with endogenous SerRS, was 2.5 μM in the presence of RF1. The inactivation of RF1 by specific antibody led to full suppression of the activity of Ser-$tRNA^{Ser(CUA)}$ and the optimal concentration was found to be 45 nM.[17] Using the same IVT with depleted RF1 and plasmid pIVEX-Est2_Amb118, addition of 10 μM AHHA-$tRNA^{Cys(CUA)}$ led to the highest yield of the esterase-neoprotein. The difference in the optimal concentration of Ser-$tRNA^{Ser(CUA)}$ and AHHA-$tRNA^{Cys(CUA)}$ can be explained by two factors. First, orthogonal suppressor AHHA-$tRNA^{Cys(CUA)}$ cannot be recycled as native aa-tRNAs are. Second, higher dissociation constants are observed for interaction between EF-Tu·GTP and α-hydroxyacyl-tRNAs as compared to α-aminoacyl-tRNAs.[27] An esterase activity test[28] was used to compare yields of the esterase-neoprotein synthesized under optimized conditions (5 μM) and the unmodified esterase 2 translated in the presence of RF1 (6.2 μM). The suppression efficiency reached 80%, producing 160 μg/mL of active protein.

The ester bond and the azide group can be specifically modified without side reactions on the rest of the esterase polypeptide. This orthogonality was exploited for: (a) test of the stability of the backbone ester under alkaline conditions, (b) availability of the linked azide group for a copper catalyzed conjugation and (c) determination of translation fidelity—the ratio between amounts of the protein with incorporated AHHA **3** and total amount of full length protein.

An ester bond, representing an isostere of the amide group in the polypeptides, is prone to hydrolysis under mild basic conditions, while peptide bonds remain stable.[29] We tested this in experiments presented in Fig. 3A. The *in vitro* synthesis of native esterase and esterase-neoprotein was performed in the presence of [^{14}C]Leu. Native esterase 2 (Fig. 3A, lanes 1–3) and esterase-neoprotein (lanes 4–6) were incubated under alkaline conditions in Tris/HCl buffer, pH 9.5 or 0.1 M NaOH. Only neoprotein with ester bond-118 was hydrolyzed, under formation of a 192 amino acid long polypeptide with apparent mass of 20 kDa (Fig. 3A, lane 6). However, strong base (0.1 M NaOH) was required to draw hydrolysis to completion. This unusually high stability of the backbone ester is, however, consistent with a previous observation, where branched and sterically demanding residues at the position preceding the ester bond influence the rate of ester bond cleavage.[29] Pro117 of the modified esterase possesses such sterical requirements to hinder ester group accessibility. This enhanced stability of the incorporated backbone ester allows convenient handling, PAGE analyses and enzymatic assaying of the neoprotein using common buffers and pH <9.

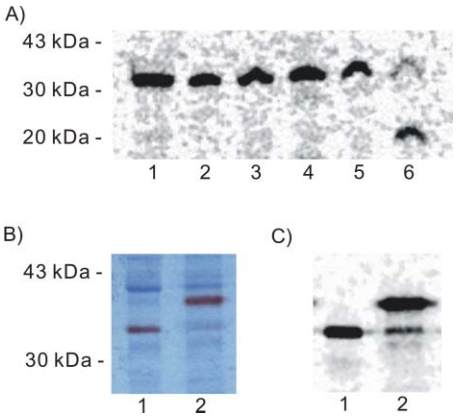
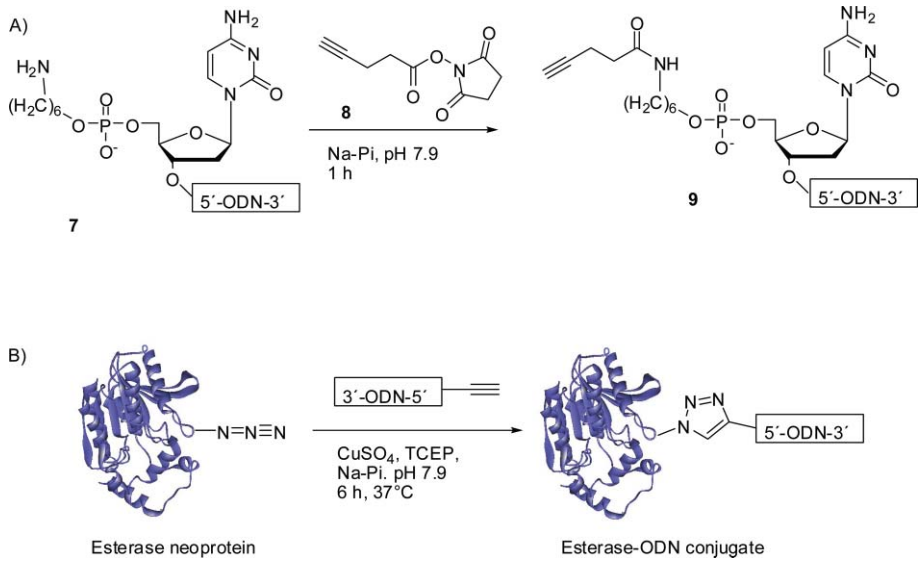


Fig. 3 SDS-PAGE of *in vitro* translation products. A) [^{14}C] Autoradiography of the products after alkaline cleavage; lanes 1–3: native esterase 2; lanes 4–6: esterase-neoprotein; lanes 1 and 4: samples analyzed directly after *in vitro* synthesis, lanes 2 and 5: samples analyzed after hydrolysis in Tris/HCl buffer pH 9.5 at 37 °C, for 30 min, lanes 3 and 6: samples analyzed after hydrolysis in 0.1 N NaOH at 37 °C, for 30 min. B) Activity staining of esterase-neoprotein (lane 1) and esterase-neoprotein conjugated with 5-alkyne-ODN (lane 2); esterase activity staining results in dark red bands of active enzyme. In addition the gels were stained by Coomassie Blue to visualize other proteins in the IVT mixture. C) [^{14}C] Autoradiography of the gel shown in B.

Cycloaddition reaction between alkynes and azides in the presence of Cu^{I} ions do not affect other reactive groups such as amines, carboxyls and sulfhydryls, commonly present in biological systems.[30,31] Thus, the presence of an azide group in position 118 of the esterase was used for conjugation with an alkyne-oligodeoxynucleotide. The 5′-alkyne-oligodeoxynucleotide (ODN) was prepared from a 10 nucleotides long 5′-amino-ODN **7** and 4-pentynoic acid succinimidyl ester (**8**) (Scheme 2A) as described previously.[32] The 5′-alkyne-ODN **9** was conjugated with azide modified esterase in the presence of whole IVT system, $CuSO_4$ and tris(2-carboxyethyl)phosphine (TCEP) in sodium phosphate buffer (Na-Pi), pH 7.9 (Scheme 2B).[32] The formation of a product with higher apparent molecular mass can be conveniently visualized by SDS-PAGE. The gel (Fig. 3B) was first stained for esterase activity[33] showing an active enzyme as a dark red band on a blue background (Coomassie Blue stained proteins of IVT mixture). Thus both, the esterase-neoprotein (lane 1) and the esterase-ODN conjugate (lane 2) were enzymatically active. The same gel was analyzed by [^{14}C] autoradiography (Fig. 3C) to calculate the ratio between conjugation product and unreacted polypeptide (Fig. 3C, lane 2). The evaluation of the band intensities revealed 80% conversion of the azide-modified esterase into esterase-ODN conjugate. The yield of the conjugation could be assumed to be quantitative since the same part of the neoprotein also underwent alkaline hydrolysis (Fig. 3A, lane 6). In accord with these results, it is plausible that the unconjugated and non-cleavable part of the synthesized protein represents polypeptides consisting of natural amino acids, probably resulting from frame-shifts or misincorporation during translation. These results mean also that the translation fidelity, with which the α-hydroxy ε-azido L-lysine **3** was introduced into the neoprotein by the presented IVT system, reached 80%.

Conclusion

We established a new orthogonal AHHA-tRNA$^{Cys(CUA)}$ for site-specific protein modifications in an *E. coli in vitro* translation system. Under conditions of RF1 depletion a high yield of *in vitro* synthesis of an azide-modified neoprotein was achieved. High suppression efficiency, together with the general applicability of the azide product for a click reaction makes the described approach superior to previously known methods for posttranslational modification of polypeptides. The presented *in vitro* synthesis of neoproteins necessitates the incorporation of an ester bond into the polypeptide chain, which is, however, sufficiently stable and does not affect the activity of the neoprotein.



Scheme 2 A) Preparation of 5′-alkyne modified ODN. B) Conjugation of the azide modified esterase and 5′-alkyne-ODN using Cu^{I} catalyzed cycloaddition.

Experimental

All reagents were of analytical grade and were purchased from Sigma-Aldrich (Taufkirchen, Germany). DNase I was from Peqlab Biotechnologie GmbH (Erlangen, Germany), T4 RNA ligase, T4 DNA ligase and restriction enzymes were from New England Biolabs (Frankfurt, Germany). 5′-Amine modified ODN was purchased from Biomers (Ulm, Germany). The transcription/translation kits were a gift from RiNA (Berlin, Germany). ^{1}H and ^{31}P NMR spectra were recorded on Bruker Avance-360 NMR (Bruker GmbH, Rheinstetten, Germany) in $CDCl_3$ using tetramethylsilane as internal standard. The autoradiography of dry gels was performed on the PhosphorImager SI (Molecular Dynamics, Sunnyvale, USA). The MALDI-TOF measurements were carried out with a Bruker Reflex III mass spectrometer (Bremen, Germany).

Preparation of 6-amino-2*(S)*-hydroxyhexanoic acid (2)

L-Lysine·H_2O **1** (3.28 g, 20 mmol) was dissolved in 50 mL of 10% H_2SO_4, heated to 45 °C and an aqueous solution of $NaNO_2$ (5.18 g, 72 mmol, 20 ml) was added dropwise during 2 h and stirred additionally at 45 °C for 7 h. The reaction was stopped by addition of an aqueous solution of urea (7.0 g, 0.12 mol, 20 ml) The reaction mixture was stirred further for 30 min, diluted with 140 mL H_2O and applied onto an ion exchange column of Dowex 50W × 8 (Serva, Heidelberg, Germany) in NH_4^+-form, (22 × 2 cm, 70 ml). The column was washed with 50 mM NH_4HCO_3 (100 ml) and product was eluted stepwise with NH_4HCO_3 (75 mM, 150 mL; 100 mM, 250 mL; 150 mM, 250 ml). Fractions containing compound **2** were combined, concentrated under reduced pressure and co-evaporated 3 times with water to remove NH_4HCO_3. Product **2** (2.36 g) was obtained in 80% yield. ^{1}H-NMR (270 MHz, $CDCl_3$) δ_H 1.30–1.75 (m, 6H, -CH_2- CH_2- CH_2-), 2.95–3.00 (m, 2H, -C*H*$_2$-NH_2), 3.95–4.00 (m, 1H,-C*H*(OH)).

Preparation of 6-azido-2*(S)*-hydroxyhexanoic acid (3)

A suspension of NaN_3 (6.8 g, 95.4 mmol) in 17 mL H_2O and 9 mL of CH_2Cl_2 was cooled on ice and triflic (Tf) anhydride (7.42 g, 5.6 mL, 26.3 mmol) was added dropwise to a vigorously stirred suspension. The reaction mixture was stirred at room temperature for 2 h. The organic layer was separated and the aqueous layer was extracted with 15 mL of CH_2Cl_2. The organic solutions were combined and washed with saturated aqueous Na_2CO_3. The resulting solution of TfN_3 in CH_2Cl_2 was added to the mixture of compound **2** (1.93 g, 13.1 mmol), K_2CO_3 (2.17 g, 15.7 mmol) and $CuSO_4$·5H_2O (26 mg, 1.1 mmol) in H_2O/methanol (100 mL, 2:3). The reaction mixture was stirred for 18 h and organic solvents were evaporated under reduced pressure. Residual aqueous solution was acidified with concentrated HCl to pH 1.7 and water was removed under reduced pressure. The crude product was purified on a silica gel column (15 × 5 cm, 300 ml) equilibrated with 30% ethyl acetate in n-hexane. Stepwise elution with 30% (300 ml), 50% (400 ml) and 60% (600 ml) ethyl acetate in n-hexane was performed and fractions containing product **3** were combined, concentrated under reduced pressure and dried over P_2O_5, affording 1.59 g (78%) of **3**. ^{1}H-NMR: (270 MHz, $CDCl_3$) δ_H 1.30–1.95 (m, 6H, -CH_2- CH_2- CH_2-), 3.25–3.30 (m, 2H, -CH_2-N_3), 4.25–4.30 (m, 1H,-C*H*(OH)); IR (ν (cm^{-1})): 3350 (O-H), 2912 and 2870 (C-H), 2120 (N_3), 1728 (C=O), 1458, 1350, 1262 (C(=O)-OH), 1116 (>C-OH), 897, 735, 665.

Cyanomethyl ester of 6-azido-2*(S)*-hydroxyhexanoic acid (4)

Compound **3** (0.43 g, 2.5 mmol) was coevaporated with dry benzene and dry toluene and redissolved in dry DMF (3 ml) under N_2. The solution was placed on ice and treated with chloroacetonitrile (1 mL, 1.2 g, 17.6 mmol), triethylamine (0.4 mL, 0.3 g, 3 mmol). The solution was stirred at room temperature for 3 days, diluted with diethyl ether (20 ml), washed with an aqueous solution of 0.1 M HCl (60 ml) and a saturated aqueous solution of NaCl. The aqueous layer was extracted with two portion of diethyl ether (2 × 20 ml). The combined organic solutions were washed with saturated aqueous solutions of $NaHCO_3$ (30 ml) and NaCl. The organic layer was dried over $MgSO_4$, concentrated under reduced pressure and the residue was purified on a silica gel column (11 × 3 cm, 80 ml) equilibrated with CH_2Cl_2. Elution with diethyl ether (10% in CH_2Cl_2) was applied to the column and fractions containing product **4** were combined, concentrated under reduced pressure and dried over P_2O_5. Compound **4** (0.45 g) was obtained in 84% yield. ^{1}H-NMR (270 MHz, $CDCl_3$) δ_H 1.40–1.90 (m, 6H, -CH_2- CH_2- CH_2-), 2.64 (d, 1H, J = 5.8 Hz, -CH(O*H*)), 3.29 (t, 2H, J = 6.5 Hz, -C*H*$_2$-N_3), 4.25–4.35 (m, 1H,-C*H*(OH)), 4.03 (s, 2H, -CH_2-CN); IR (ν (cm^{-1})): 3450 (O-H), 2912 and 2870 (C-H), 2120 (N_3 and C≡N), 1800 (C=O), 1460, 1430, 1363, 1257 and 1160 (C(=O)-OR), 1120 (>C-OH), 1048, 1008, 897, 735.

Tetrabutylammonium salt of pdCpA dinucleotide 5

Dinucleotide **5** (3090 OD_{260}, 0.134 mmol) in 100 mL of 10 mM acetic acid was treated with 1.5 mM tetrabutylammonium hydroxide. The solution was lyophilized, the residue was redissolved in 15 mL of water and the pH of the resulting solution was adjusted to 5.5 with 0.1 M tetrabutylammonium hydroxide. The lyophilization was repeated and the residue was co-evaporated with absolute ethanol under reduced pressure. The tetrabutylammonium salt of dinucleotide **5** was dried over P_2O_5 for 16 hours.

Preparation of acylated dinucleotide pdCpA 6

The tetrabutylammonium salt of dinucleotide **5** (0.022 mmol, ESI†), dissolved in 0.52 mL of 40 mM solution of tetrabutylammonium hydroxide in dry dimethylformamide, was concentrated under reduced pressure and co-evaporated two times with dry dimethylformamide. The residue was treated with a solution of cyanomethyl ester **4** (0.3 mL, 0.24 mmol) in dry dimethylformamide. The reaction was stirred under a protective atmosphere of N_2 for 5 h. The course of the reaction was followed by HPLC. The reaction product was isolated by preparative HPLC affording acylated dinucleotide **6** in 10% overall yield (52 OD_{260nm}, 2.25 μmol). The product was concentrated under reduced pressure and the resulting ammonium salt of **6** was lyophilized three times from 25 mL of 10 mM acetic acid to obtain the H^+-form of compound **6**. For mixture of both 3′- and 2′-acyl isomers in ratio 2.3 : 1; calculated: $C_{25}H_{35}N_{11}O_{15}P_2$, MW = 791.56; found: TOF MS ES+, $[M+H]^+$ 792.23; ^{1}H-NMR (360 MHz, D_2O) δ_H (tetrabutylammonium and H_2O signals are excluded) 1.20–1.40 (m, 3H), 1.60–1.95 (m, 4H), 2.30–2.40 (m, 1H), 3.13 (t, 0.6H,

J = 6,5 Hz), 3.23 (t, 1H, J = 6,5 Hz), 3.88 (s, 2H), 3.95–4.05 (m, 2H), 4.18 (s, 1H), 4.25–4.35 (m, 0.6H), 4.35–4.45 (m, 1.3H), 4.98 (t, 0.7H, J = 6.5 Hz), 5.41–5.44 (m, 0.7H), 5.59 (t, 0.3H, J = 4.5 Hz), 5.95–6.05 (m, 2.7H), 6.18 (d, 0.3H, J = 4.5 Hz), 7.83 (d, 1H, J = 7,5 Hz), 8.17 (s, 1H), 8.38 (s, 0.3H), 8.45 (s, 0.7H); ^{31}P-NMR (145 MHz, D_2O) δ_p 0.25 (s, 3′-isomer), 0.94 (s, 2′-isomer).^{1}H, ^{13}C, HH-COSY, ^{31}P NMR spectra, structural assignment and MALDI-TOF spectra of compound **6** are presented in the ESI† (Figs. 1S–5S).

Analytical HPLC

Column SP 50/10 Nucleosil 100-5 C18 (Macherey-Nagel, Düren, Germany), buffer A: 50 mM ammonium acetate, pH 4.5, buffer B: mixture of 50 mM ammonium acetate pH 4.5 and CH_3CN in the ratio of 1:9, flow 1 ml/min, detection at 260 and 280 nm, linear gradient of buffer B from 0–100% over 30 min. Retention times of eluted samples: t_R = 8.5 min (dinucleotide **5**), t_R = 11.5 min (acylated product **6**).

Preparative HPLC

Column SP 250/10 Nucleosil 100-5 C18 (Macherey-Nagel), buffer A: 50 mM ammonium acetate, pH 4.5, buffer B: mixture of 50 mM ammonium acetate pH 4.5 and CH_3CN in the ratio of 3:7, flow 1 ml/min, detection at 260 and 280 nm, isocratic elution with buffer A over 25 min and then linear gradient of buffer B from 0–100% over 80 min. Retention times of eluted samples: t_R = 59 min (dinucleotide **5**), t_R = 72 min (acylated product **6**).

Preparation of suppressor AHHA-tRNA$^{Cys(CUA)}$

The plasmid ptCys_amber for *in vitro* transcription of tRNA$^{Cys(CUA)}$ (C74) was constructed by overlap extension PCR. The forward primers (1) 5′-*GAATTCT*AATACGACTCACTATAG-3′, (2) 5′-TAATACGACTCACTATAGGCGCGTTAACAAAGC-3′, (3) 5′-GGTTATGTAGCGGATTCTAAATCCGTCTAGTCC-3′, (4) 5′ - GGTTCGACTCCGGAACGCGCCTCCAGGTGATCC - 3′ and the reverse primers (5) 5′-ATCCGCTACATAACCGCTTTG-TTAACGCGCCTA-3′, (6) 5′-TTCCGGAGTCGAACCGGA-CTAGACGGATTTAGA-3′, (7) 5′-*AAGCTT*GGATGGAT-CACCTGGAGGCGCG-3′ were used for the PCR. The PCR product was cloned into pGEM-T vector. The resulting plasmid was sequenced. The correct gene of tRNA$^{Cys(CUA)}$ (C74) was then cloned into pUC19 vector *via* the cleavage sites *Eco*RI and *Hind*III, italicized in the primer sequences. The resulting plasmid was used for *in vitro* transcription of tRNA$^{Cys(CUA)}$ (C74). The DNA template of *in vitro* transcription was obtained by PCR with the forward primer (1) and reverse primer (8) 5′-GAGGCGCGTTCCGGAGTC-3′. Run off transcripts of tRNA$^{Cys(CUA)}$ (C74) were obtained by incubation of 400 µL mixtures containing 40 mM Tris/HCl, pH 8.1, 22 mM $MgCl_2$, 1 mM spermidine, 5 mM DTE, 0.01% Triton X-100, 4 mM of ATP, CTP, GTP and UTP, 15 ng/mL T7 RNA polymerase and DNA template at 37 °C for 3 hours. After transcription, 4 µL DNase I (1 U/µL) as added to the reaction mixtures which were incubated 1 hour at 37 °C to remove DNA template. Reaction mixtures were purified by preparative 20% Urea-PAGE (40 × 20 × 0.1 cm). The transcription mixture (400 µL) was loaded onto the gel in a 12 cm wide pocket. Resolution of one nucleotide was obtained after electrophoresis at 40 mA for 12–15 hours. Transcripts were eluted from gel slices using 0.3 M NaOAc (pH 5.2). Concentrations of transcripts were calculated from absorbance measurements at 260 nm, assuming that 1 absorbance unit (1 cm pathway) corresponds to 40 µg/mL. Purified truncated transcript tRNA$^{Cys(CUA)}$ (C74) (30 µM) was ligated with acylated dinucleotide **6** (300 µM) using 1 U of T4 RNA ligase in 50 mM Tris/HCl, 10 mM MgCl2, 1 mM ATP, 10 mM dithiothreitol, 25 µg/mL BSA, pH 7.5 at 4 °C for 5 h. The aliquots of suppressor AHHA-tRNA$^{Cys(CUA)}$ were used directly in *in vitro* nonsense suppression experiments.

Construction of plasmid pIVEX-Est2_Amb118

The esterase gene in pIVEX_Est2 (pEst2) plasmid was prepared with a UGA stop codon at the end of the esterase gene, by the overlap extension method. Two separate PCRs were carried out using T7 promoter primer (5′-TAATACGACTCAC-TTATAGGG-3′) and EstE118x_rev (5′-CAGGGAACTT-GTG*CTA*CGGCGCCAGGCGGTAGTC-3′); EstE118x_for (5′-CCGCCTGGCGCCG*TAG*CACAAGTTCCCTGCCGC-3′) and T7 terminator primer (5′-CTAGTTATTGCTCAGCGGTG-3′). The mutated codons are italicized. Position 118 (Glu) of the esterase amino acid sequence was replaced by RF1 stop codon (TAG). The PCR fragments were fused in a second PCR using T7 promoter and T7 terminator primers. The fused PCR product was digested with *Nco*I/*Sac*I and ligated into *Nco*I/*Sac*I digested pIVEX vector. The ligation mixture was transformed into *Escherichia coli* strain XL-1 Blue. The plasmid DNA was isolated from clones and sequenced. The resulting plasmid pEst2_amb118 was used for *in vitro* translation.

Preparation of azide modified esterase-neoprotein by *in vitro* translation

The translation mixture was incubated with 0.5 mg/mL purified rabbit polyclonal antibody against RF1, obtained as described,[34] at 37 °C for 5 minutes to deactivate endogenous RF1. The reaction was started by adding 5 nM of template pIVEX-Est2_Amb118 and 10 µM of AHHA-tRNA$^{Cys(CUA)}$ and incubated at 37 °C for 2 h. For [^{14}C]-labeling, 0.5 mM [^{14}C]leucine (17.3 mCi/mmol) was added. The reaction mixture was analyzed by 12.5% SDS-PAGE.[35] Polyacrylamide gels were first stained for esterase activity[33] and then proteins were fixed with 15% formaldehyde in 60% methanol and stained with Coomassie Brilliant Blue G-250. Radioactive probes were identified by PhosphorImager SI.

Alkaline treatment of *in vitro* synthesized native esterase and esterase-neoprotein

Tris-HCl buffer (1M, pH 9.5) or NaOH aqueous solution (1M) were added into IVT mixture (10 µL) of [^{14}C] labeled native esterase and esterase-neoprotein, respectively, to a final concentration of 100 mM. The mixture was incubated at 37 °C for 30 minutes. The samples were loaded on SDS-PAGE. The gel was stained with Coomassie Blue and the dried gel was visualized by [^{14}C] autoradiography.

Conjugation of 5′-alkyne-ODN with the azide group of esterase-neoprotein

An IVT mixture containing modified esterase (5 μM) was pre-treated with 20 mM EDTA, subjected to buffer exchange on Microcon filter device YM10 (Millipore, Schwalbach, Germany) against 0.1 M sodium phosphate buffer, pH 7.9 and eluted with the same buffer. The IVT mixture (20 μL, 5 μM of neoprotein) was added to dry 5′-alkyne ODN (2.0 nmol). The resulting mixture was treated with 0.5 μL of $CuSO_4$ (final concentration 2 mM) and 0.5 μL of TCEP (final concentration 3 mM). The reaction mixture was incubated in a sealed Eppendorf tube at 37 °C for 6 h. The prepared neoprotein-ODN conjugate was characterized by 12.5% SDS-PAGE (Fig. 3B, C).

Acknowledgements

This work was supported by Deutsche Forschungsgemeinschaft, DFG Sp 243/12-2. We thank RiNA GmbH for supplying us with the *in vitro* transcription/translation kits.

References

1 T. L. Hendrickson, V. de Crecy-Lagard and P. Schimmel, *Annu. Rev. Biochem.*, 2004, **73**, 147–176.
2 A. J. de Graaf, M. Kooijman, W. E. Hennink and E. Mastrobattista, *Bioconjug. Chem.*, 2009.
3 D. R. Liu, T. J. Magliery, M. Pasternak and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 10092–10097.
4 L. Wang and P. G. Schultz, *Angew. Chem., Int. Ed.*, 2005, **44**, 34–66.
5 J. M. Xie and P. G. Schultz, *Methods*, 2005, **36**, 227–238.
6 K. H. Wang, H. Neumann, S. Y. Peak-Chew and J. W. Chin, *Nat. Biotechnol.*, 2007, **25**, 770–777.
7 D. A. Dougherty, *Curr. Opin. Chem. Biol.*, 2000, **4**, 645–652.
8 B. Wiltschi and N. Budisa, *Appl. Microbiol. Biotechnol.*, 2007, **74**, 739–753.
9 K. Sakamoto, A. Hayashi, A. Sakamoto, D. Kiga, H. Nakayama, A. Soma, T. Kobayashi, M. Kitabatake, K. Takio, K. Saito, M. Shirouzu, I. Hirao and S. Yokoyama, *Nucleic Acids Res.*, 2002, **30**, 4692–4699.
10 D. Mendel, V. W. Cornish and P. G. Schultz, *Annu. Rev. Biophys. Biomol. Struct.*, 1995, **24**, 435–462.
11 G. F. Short, S. Y. Golovine and S. M. Hecht, *Biochemistry*, 1999, **38**, 8808–8819.
12 Y. Shimizu, T. Kanamori and T. Ueda, *Methods*, 2005, **36**, 299–304.
13 D. Kiga, K. Sakamoto, K. Kodama, T. Kigawa, T. Matsuda, T. Yabuki, M. Shirouzu, Y. Harada, H. Nakayama, K. Takio, Y. Hasegawa, Y. Endo, I. Hirao and S. Yokoyama, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9715–9720.
14 S. Ohno, M. Matsui, T. Yokogawa, M. Nakamura, T. Hosoya, T. Hiramatsu, M. Suzuki, N. Hayashi and K. Nishikawa, *J. Biochem.*, 2007, **141**, 335–343.
15 J. Fukunaga, T. Yokogawa, S. Ohno and K. Nishikawa, *J. Biochem.*, 2006, **139**, 689–696.
16 D. E. Agafonov, K. S. Rabe, M. Grote, Y. W. Huang and M. Sprinzl, *FEBS Lett.*, 2005, **579**, 2082–2086.
17 D. E. Agafonov, Y. W. Huang, M. Grote and M. Sprinzl, *FEBS Lett.*, 2005, **579**, 2156–2160.
18 D. S. Karanewsky, M. C. Badia, D. W. Cushman, J. M. Deforrest, T. Dejneka, M. J. Loots, M. G. Perri, E. W. Petrillo and J. R. Powell, *J. Med. Chem.*, 1988, **31**, 204–212.
19 A. J. Link, M. K. S. Vink and D. A. Tirrell, *J. Am. Chem. Soc.*, 2004, **126**, 10598–10602.
20 P. M. England, H. A. Lester and D. A. Dougherty, *Tetrahedron Lett.*, 1999, **40**, 6189–6192.
21 S. A. Robertson, C. J. Noren, S. J. Anthonycahill, M. C. Griffith and P. G. Schultz, *Nucleic Acids Res.*, 1989, **17**, 9649–9660.
22 D. Hentzen, P. Mandel and J. P. Garel, *Biochim. Biophys. Acta*, 1972, **281**, 228–232.
23 S. T. Cload, D. R. Liu, W. A. Froland and P. G. Schultz, *Chem. Biol.*, 1996, **3**, 1033–1038.
24 G. A. Komatsoulis and J. Abelson, *Biochemistry*, 1993, **32**, 7435–7444.
25 S. Hauenstein, C. M. Zhang, Y. M. Hou and J. J. Perona, *Nat. Struct. Mol. Biol.*, 2004, **11**, 1134–1141.
26 M. Sprinzl, *Trends Biochem. Sci.*, 1994, **19**, 245–250.
27 K. H. Derwenskus and M. Sprinzl, *FEBS Lett.*, 1983, **151**, 143–147.
28 G. Manco, E. Adinolfi, F. M. Pisani, G. Ottolina, G. Carrea and M. Rossi, *Biochem. J.*, 1998, **332**(Pt 1), 203–212.
29 P. M. England, H. A. Lester and D. A. Dougherty, *Biochemistry*, 1999, **38**, 14409–14415.
30 Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless and M. G. Finn, *J. Am. Chem. Soc.*, 2003, **125**, 3192–3193.
31 M. Meldal and C. W. Tornoe, *Chem. Rev.*, 2008, **108**, 2952–3015.
32 M. Humenik, Y. Huang, Y. Wang and M. Sprinzl, *ChemBioChem*, 2007, **8**, 1103–1106.
33 T. B. Higerd and J. Spizizen, *J. Bacteriol.*, 1973, **114**, 1184–1192.
34 D. E. Agafonov, K. S. Rabe, M. Grote, C. S. Voertler and M. Sprinzl, *ChemBioChem*, 2006, **7**, 330–336.
35 U. K. Laemmli, *Nature*, 1970, **227**, 680–685.

