Electronic Supplementary Material (ESI) for Molecular BioSystems

Supplementary Information

Wide-range protein photo-crosslinking achieved by a genetically encoded N^{ε}-(benzyloxycarbonyl)lysine derivative with a diazirinyl moiety

Tatsuo Yanagisawa[a]*, Nobumasa Hino[a]*, Fumie Iraha[a], Takahito Mukai[a], Kensaku Sakamoto[a], and Shigeyuki Yokoyama[a, b]

[a]*RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan*. [b]*Laboratory of Structural Biology, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*.

Supplementary Methods

Materials and methods

Biochemical and molecular biological procedures were performed using commercially available materials, enzymes, and chemicals. *p*NO_2ZLys was purchased from Bachem (Switzerland). 4-(3-(Trifluoromethyl)-3*H*-diazirin-3-yl)benzylalcohol was purchased from Teika Pharmaceuticals (Toyama, Japan). ESI-IT-TOF MS and MS/MS spectra were measured using an LCQ-DECA XP spectrometer (Thermofisher Scientific). ESI-Q-TOF MS and MS/MS spectra were measured using an Applied Biosystems QSTAR Elite spectrometer. NMR experiments were performed on a Bruker AV800

spectrometer.

Synthesis of N^{ε}-[((4-(3-Trifluoromethyl-3*H*-diazirin-3-yl)benzyl)oxy)carbonyl]-L-lysine (TmdZLys)

4-(3-(Trifluoromethyl)-3*H*-diazirin-3-yl)benzylalcohol (300 mg, 13.9 mmol) was activated with 1,1'-carbonyldiimidazole. The obtained 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzyloxycarbonylimidazole carbonate (45 mg, 1.45 mmol) was coupled with N^{α}-*t*-butyloxycarbonyl (Boc)-L-lysine (362 mg, 14.7 mmol) in the presence of DIPEA. The obtained N^{α}-Boc-(N^{ε}-[((4-(3-trifluoromethyl-3*H*-diazirin-3-yl)benzyl)oxy)carbonyl]-L-lysine was deprotected with 20% trifluoroacetic acid (TFA) to yield TmdZLys (180 mg, 0.46 mmol). The chemical structure of TmdZLys was confirmed by ^{1}H-, ^{13}C-, and ^{19}F-NMR and mass spectrometry. ^{1}H NMR (800 MHz, DMSO-d6): δ (ppm) 1.380−1.525 (m, 2H), 1.4496−1.562 (m, 2H), 1.774−1.907 (m, 2H), 3.8230 (s, 1H), 3.076−3.123 (m, 2H), 5.183 (s, 2H), 7.4275 (d, J = 8.3 Hz, 2H), 7.4474 (t, J = 6.4 Hz, 1H), 7.5290 (d, J = 8.3 Hz, 2H), 8.156 (bs, 1H). ^{13}C-NMR (200 MHz, DMSO-d6): δ (ppm) 22.28, 28.495 (d, J = 40.69Hz), 29.377, 30.389, 40.543, 52.857, 64.735, 122.176 (d, J = 273.73 Hz), 127.047, 127.304, 128.757, 140.377, 156.355, 171.358. ^{19}F-NMR (376 MHz, DMSO-d6): δ (ppm) -64.665. MALDI-TOF Mass: Calcd, m/z = 388.35 $[M+H]^{+}$, Obsd, m/z = 389.10 $[M+H]^{+}$.

Plasmid construction and site-directed mutagenesis

The tyrosine codon at position 25 of the wild-type GST gene in the plasmid pET-GSTwt was mutagenized to an amber codon by using a QuikChange site-directed mutagenesis kit (Stratagene), to create the plasmid pET-GST(25Am)[S1]. A pACYC184-derived plasmid, pTK2.1-PylRS(Y306G-Y384F)-tRNA^{Pyl}, containing the *M. mazei* PylRS(Y306G-Y384F) gene under the control of the tryptophanyl-tRNA synthetase (*trpS*) promoter, and the tRNA^{Pyl} gene under the control of the prolipoprotein (*lpp*) promoter, was PCR-mutagenized from the plasmid pTK2.1-PylRS(Y306A-Y384F)-tRNA^{Pyl} [S1]. The following mammalian expression plasmids were constructed as described[S2]: pcDNA4/TO-GRB2GFP, containing the GRB2-EGFP gene: pOriP-GRB2, containing the GRB2 gene with the sequence encoding a C-terminal FLAG tag (DYKDDDDK): pcDNA4/TO-SHC-EGFR, containing the SHC and EGFR genes: pOriP9xU6tRNA^{Pyl}, containing nine copies of the tRNA^{Pyl} gene: and pOriP-PylRS, containing the PylRS(Y306A-Y384F) [or PylRS(Y306G-Y384F)] gene. The mutagenesis to the wild-type GRB2 gene was performed with a QuikChange kit, to create a series of clones with an amber codon at positions 103–112 [S2].

Expression and purification of PylRS and tRNA^{Pyl}

The *M. mazei* PylRS mutant PylRS(Y306G-Y384F) was overexpressed in *E. coli* BL21-Gold(DE3) cells harboring pET28-PylRS(Y306G-Y384F), and was purified by three column chromatography steps on HisTrap (GE Healthcare), HiTrap Q (GE

Healthcare), and Superdex 75 HiLoad 16/60 (GE Healthcare) as described[S3], with some modifications. The purified PylRS protein was dialyzed against 20 mM potassium phosphate buffer (pH 7.4), containing 300 mM NaCl and 10 mM β-mercaptoethanol, and was concentrated to 3.5 mg•ml^{-1}. *M. mazei* tRNA^{Pyl} was transcribed and purified as described previously[S3].

Aminoacylation assay

The aminoacylation assay by acidic urea PAGE was performed as described previously[S1,S3,S4]. The aminoacylation reactions were incubated at 37 °C for 1 hr. The standard aminoacylation assay solution (20 μl) contained 2.8 μM purified *M. mazei* PylRS(Y306G/Y384F), 10 mM $MgCl_2$, 2 mM ATP, 4 mM DTT, 2.1 μM *M. mazei* tRNA^{Pyl} transcript, and the appropriate amino acid concentrations, in 100 mM Na-Hepes buffer (pH 7.2). Unaminoacylated and aminoacylated tRNA^{Pyl}s were subjected to electrophoresis on a 10% denaturing urea polyacrylamide gel under acidic conditions (pH 5.0) at 4 °C for 18 hr, and were stained with 2 % toluidine blue.

Site-specific incorporation of TmdZLys into proteins, and *in vivo* photo-crosslinking

Cotransformation of the plasmids pET-GST(25Am) and pTK2.1-PylRS(Y306A-Y384F)-tRNA^{Pyl} [or pTK2.1-PylRS(Y306G-Y384F)-tRNA^{Pyl}] into *E. coli* BL21-Gold(DE3) (Novagen) was performed as described previously[S1]. The incorporation of TmdZLys into GST in *E. coli* cells, using the

PylRS(Y306A-Y384F)/tRNA^{Pyl} and PylRS(Y306G-Y384F)/tRNA^{Pyl} systems, the purification of GST proteins, and the MALD-TOF MS and ESI-MS/MS spectrometry analyses were performed as described previously[S1,S5], with some modifications. The incorporation of TmdZLys into GST was performed in M9 medium[S6] containing 7.5% DMSO. To estimate the incorporation rates of TmdZLys into proteins in mammalian cells, T-Rex CHO cells (Invitrogen) and HEK293 c-18 cells (ATCC) were transfected with the gene encoding the GRB2-EGFP fusion protein or its amber mutant GRB2(111Am)-EGFP, together with the genes encoding PylRS(Y306A-Y384F) and tRNA^{Pyl}, as described previously[S7]. Four hours after transfection, the cell culture medium was replaced with DMEM/F-12 (Invitrogen) or DMEM (Invitrogen) supplemented with 0–0.5mM of ZLys or TmdZLys. After a 16 hr incubation, the fluorescence intensity from the cell lysates was measured with a Fusion microplate analyzer (Perkin-Elmer). Protein photo-crosslinking in the cells, immunoprecipitation, and western blotting were performed as described previously[S2,S8]. HEK293 cells expressing each of the ten GRB2(TmdZLys) variants [GRB2(103TmdZLys) to GRB2(112TmdZLys)] (20 kDa), together with EGFR (170 kDa) and SHC (50–55 kDa) were irradiated with 365-nm UV light at 0 °C for 15 min.

Screening of a PylRS mutant that incorporates *p*NO$_2$ZLys into a protein

If the Tmd group is to be applied as an efficient photo-crosslinking probe, then the *para* substitution on the benzene ring of ZLys is much better than the *ortho* position. To create a PylRS active site suitable for a *para*-substituted derivative of ZLys, we

screened the PylRS mutants with a ZLys derivative bearing a nitro group at the *para*-position, N^{ε}-(*p*-nitrobenzyloxycarbonyl)-L-lysine (*p*NO_2ZLys). PylRS(Y384F) was used as the template for mutagenesis, because PylRS(Y384F) has higher aminoacylation activities for bulky lysine derivatives than those of the wild-type PylRS[S1]. On the basis of the PylRS•pyrrolysine complex structure[S3], a library of PylRS mutants was generated, in which the active site residues Tyr306, Leu309, and Ile413 of PylRS(Y384F) were randomized by overlap extension PCR with synthetic oligonucleotide primers (the intended mutations were encoded by NNK; N =A+G+C+T, K= G+T), and screened by two rounds of alternating positive and negative selections[S1]. The mutant PylRS library was subjected to the positive selection, based on the suppression of an amber stop codon in the chloramphenicol acetyltransferase (CAT) gene. Cells transformed with the mutant PylRS library and the $tRNA^{Pyl}$ gene were grown in LB medium containing 1 mM *p*NO_2ZLys, and were screened for their survival in the presence of chloramphenicol. The PylRS mutant library was then subjected to the negative selection in the absence of *p*NO_2ZLys, based on the suppression of two amber stop codons in the toxic barnase gene. After two rounds of the positive and negative selections, we obtained the PylRS(Y306G-Y384F) mutant.

Supplementary References

[S1] T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, and S. Yokoyama, *Chem. Biol.* 2008, **15**, 1187-1197.

[S2] N. Hino, Y. Okazaki, T. Kobayashi, A. Hayashi, K. Sakamoto, and S. Yokoyama, *Nat. Methods* 2005, **2**, 201-206.

[S3] T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, and S. Yokoyama, *J. Mol. Biol.* 2008, **378**, 634-652.

[S4] U. Varshney, C.P. Lee, and U.L. RajBhandary. *J Biol Chem*. 1991, **266**, 24712-24718.

[S5] T. Yanagisawa, T. Sumida, R. Ishii, C. Takemoto, and S. Yokoyama, *Nat. Struct. Mol. Biol.* 2010, **17**, 1136-1143.

[S6] K. Kuwasako, F. He, M. Inoue, A. Tanaka, S. Sugano, P. Güntert, Y. Muto, and S. Yokoyama. *Structure* 2006, **14**, 1677-1689.

[S7] N. Hino, K. Sakamoto, and S. Yokoyama, *Methods Mol. Biol.* 2012, **794**, 215-228.

[S8] N. Hino, M. Oyama, A. Sato, T. Mukai, F. Iraha, A. Hayashi, H. Kozuka-Hata, T. Yamamoto, S. Yokoyama, and K. Sakamoto, *J. Mol. Biol.* 2011, **406**, 343-353.

Electronic Supplementary Material (ESI) for Molecular BioSystems

Yanagisawa *et al.* Figure S1

A

1 25

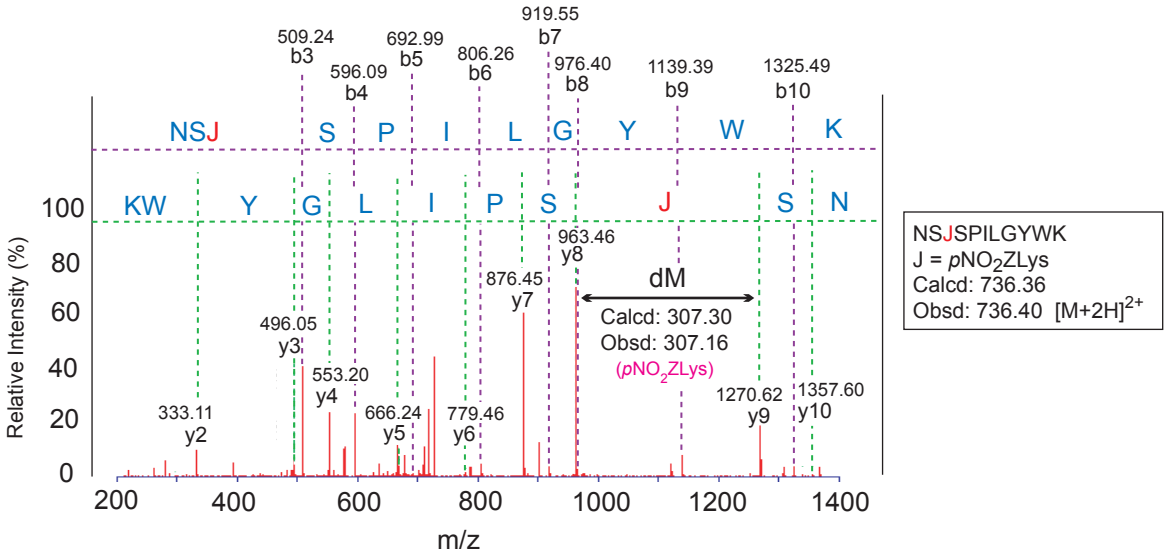
MASMTGGQQMGRDPGANSGVTKNSYSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDE

GDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGA

VLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDAL

DVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSD

B



Supplementary Figures

Figure S1.

(A) Amino acid sequence of *S. japonicum* GST with a 24-residue T7-peptide tag at the N-terminus. The GST residue 25 is highlighted in green. (B) The incorporation of *p*NO_2ZLys into GST at position 25 was confirmed by an ESI-IT-TOF MS/MS analysis of the tryptic peptide (calcd: 736.36, obsd: 736.40 $[M+2H]^{2+}$) from GST(25*p*NO_2ZLys). The molecular mass of J was 307.16, which agrees well with the calculated mass of *p*NO_2ZLys (m/z = 307.30 $[M+H]^{+}$).

Electronic Supplementary Material (ESI) for Molecular BioSystems

Yanagisawa *et al*. Figure S2

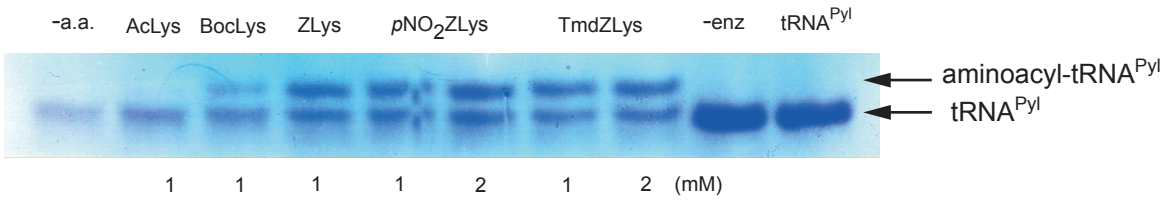


Figure S2.

PylRS(Y306G-Y384F) esterifies $tRNA^{Pyl}$ with ZLys derivatives containing nitro- and Tmd groups at the *para*-position. The aminoacylation of $tRNA^{Pyl}$ was monitored by acidic urea PAGE analysis. Aminoacylation assay conditions are described in the Supplementary methods. Each lane shows a reaction in the presence of PylRS(Y306G-Y384F), $tRNA^{Pyl}$, and the following: no amino acid; 1 mM N^{ε}-acetyl-L-lysine (AcLys); 1 mM N^{ε}-(*tert*-butyloxycarbonyl)-L-lysine (BocLys); 1 mM ZLys; 1 mM *p*NO_2ZLys, 2 mM *p*NO_2ZLys, 1 mM TmdZLys; 2 mM TmdZLys, no enzyme, and control $tRNA^{Pyl}$.