Published in final edited form as:

Chembiochem. 2014 August 18; 15(12): 1805–1809. doi:10.1002/cbic.201402083.

Exploring the Substrate Range of Wild-type Aminoacyl-tRNA Synthetases

Dr. Chenguang Fan^{#[a]}, Joanne M.L. Ho^{#[a]}, Napon Chirathivat^[a], Dr. Prof. Dieter Söll^{*,[a],[b]}, and Dr. Yane-Shih Wang^{*,[a]}

^[a]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520 (USA)

^[b]Department of Chemistry, Yale University, New Haven, CT 06520 (USA)

[#] These authors contributed equally to this work.

Abstract

We tested the substrate range of four wild-type *Escherichia coli* aminoacyl-tRNA synthetases (AARSs) with a library of non-standard amino acids (nsAAs). While these AARSs could discriminate efficiently against the other canonical amino acids, they were able to use many nsAAs as substrates. Our results also showed *E. coli* tryptophanyl-tRNA synthetase (TrpRS) and tyrosyl-tRNA synthetase to have overlapping substrate ranges. In addition, we found that the nature of the anticodon sequence of tRNA^{Trp} altered the nsAA substrate range of TrpRS; this implies that the sequence of the anticodon affects the TrpRS amino acid binding pocket. These results highlight again that inherent AARS polyspecificity will be a major challenge to the goal of incorporating multiple different amino acids site-specifically into proteins.

Keywords

aminoacyl-tRNA synthetase; natural nonsense tRNA suppressor; non-standard amino acid; substrate specificity; synthetic biology

Introduction

Aminoacyl-tRNA synthetases (AARSs) are thought to be highly specific for their cognate amino acid substrates^[1]. However, studies over the past fifty years indicated that some non-standard amino acids (nsAAs) are readily incorporated into bacterial proteins^[2]. *In vitro* experiments with a mixture of all *Escherichia coli* AARSs showed that more than 90 nsAAs can be acylated onto tRNA^[3]. Protein engineering directed toward expansion of the genetic code during the past decade^[4] led to many variant AARS enzymes with substrate polyspecificity for some of the nsAAs employed^[4a, 5]. Some variants of *Methanococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS)^[5i] and of pyrrolysyl-tRNA synthetase can acylate more than ten different compounds^[5e, 5f]. Therefore, the high specificity observed

*dieter.soll@yale.edu; yane-shih.wang@yale.edu.

by the AARSs against the other canonical amino acids does not extend to a large number of nsAAs.

We decided to examine a number of canonical wild-type *E. coli* AARSs for their substrate specificity with a large library of nsAAs. As test enzymes we used *E. coli* leucyl-tRNA synthetase (LeuRS), seryl-tRNA synthetase (SerRS), tryptophanyl-tRNA synthetase (TrpRS) and TyrRS. Each of these enzymes has a natural nonsense suppressor tRNA^[6] that, when acylated, can serve to insert its amino acid in response to a stop codon (*e.g.* the amber (UAG) and opal (UGA) codons) engineered into the second codon of the superfolder green fluorescent protein (sfGFP) mRNA. In order to decrease the competition of the cognate amino acid with the nsAAs to be tested, the experiments were performed in *E. coli* auxotrophic strains for the cognate amino acid^[7]. Our results lend further support to our view of natural AARSs being polyspecific for nsAAs.

Results

Experimental scheme

We used the *E. coli* wild-type AARS enzymes with their native amber or opal suppressor tRNAs^[6]. They were *supU* (tRNA^{Trp}_{CUA}), *trpT* (tRNA^{Trp}_{Tyr}_{UCA}), *supF* (tRNA^{Leu}_{CUA} CUA) and *supD* (tRNA^{Ser}_{CUA}). The TyrRS/*supF* pair was used to optimize the fluorescence signal arising from read-through of the *sfGFP-UAG2* mRNA in a *tyrA*[−] auxotrophic strain (Figure S1). The auxotrophic strains (Table S1) transformed with the genes for the AARS/suppressor tRNA pair (Figure S2 and Table S2) were grown at 37°C for 16 h in minimal medium with 1% glycerol supplemented with the 20 canonical amino acids (1 mM). After washing to remove the amino acids and transfer to minimal medium, the strains were cultured in a 384-well plate format where each well contained one member of our nsAA library at 1 mM concentration. The nsAAs library contained 313 members, of which 272 were previously reported^[5b] (Tables S3-S5). The fluorescence readout, which was contingent on suppression of the *sfGFP-UAG2* mRNA, was used as a measure of the ability of the corresponding endogenous AARS (LeuRS, SerRS, TrpRS and TyrRS) to recognize the nsAA substrates. This allowed a measurement of the range of amino acid substrate specificity for the AARS.

The wild-type protein sequence of sfGFP contains a single Trp (**1**) residue (W57). Under our incorporation conditions with wild-type tRNA^{Trp}_{CCA} this W57 could also be replaced by Trp analogs (**2-4**), Cys analog **5** and Phe analog **6** (Scheme 1; Figure 1A). The fluorescence intensities generated by nsAAs **4** and **5** were comparable to that of Trp. A similar substrate preference was observed with the TrpRS/*supU* pair (Figure 1B), although the pattern of suppression efficiencies was different. This change in substrate range stemmed purely from the different anticodon sequence, since the *supU* tRNA differs only by a single base change from the endogenous tRNA^{Trp}_{CCA} isoacceptor.^[6] We also tested the TrpRS substrate range with the natural opal suppressor tRNA^{Trp}_{UCA} encoded by *trpT*. As expected, the TrpRS/*trpT* pair incorporated Trp (**1**) into sfGFP but at a lower efficiency (Figure 1C). Interestingly, a Trp lactic acid analog (**2**) (Figure 1C) was a better substrate than Trp for TrpRS/*trpT* pair. Thus, different tRNAs gave somewhat different suppression profiles suggesting that the nature of the anticodon may affect the amino acid recognition, a property that was also seen

with glutamyl-tRNA synthetase^[8]. This suggests that in the strategy to develop orthogonal translation systems for genetic code expansion, modulations in amino acid recognition should be considered when tRNA anticodons are altered.

The substrate specificity profile of TyrRS/*supF* pair indicated that eight aromatic ring-containing nsAAs (**3-4**, **7-12**) were substrates with appreciable signals but had low suppression efficiencies, albeit comparable to that of Tyr (**13**) (Figure 2A). The LeuRS/*supP* pair revealed three major fluorescence signals that indicated efficient incorporation of Leu (**14**) and two related nsAAs (**15-16**) that contained γ -vinyl and α -hydroxyl functional groups (Figure 2B). Previous studies showed γ -trifluoromethyl-leucine (**17**) to be a substrate for LeuRS [9]; our nsAA library screening also gave a small positive signal.

Our auxotrophic strain strategy provided sufficient amounts of Trp, Tyr, and Leu, but a much smaller concentration of Ser. Serine is involved in multiple cellular pathways, and is subject to high metabolic flux^[10], and the higher conversion rate of serine may have led to a lower amino acid level. Thus, the SerRS/*supD* pair gave a low fluorescence signal with Ser (Figure 2C; **18**), but high signals with two Ser analogs (Figure 2C; **19** and **20**). Compared to the wider substrate range of the TyrRS (Figure 2A) and TrpRS (Figure 1B), the smaller substrate range of LeuRS (Figure 2B) and SerRS (Figure 2C) may be explained that the latter two synthetases have an editing function, while the former two do not^[1a, 11].

LC-MS/MS analysis of amino acids incorporated at positions 2 and 57 of sfGFP-UAG2

We performed liquid chromatography with tandem mass spectrometry (LC-MS/MS) to confirm the results by fluorescence screening mentioned above. For the TrpRS/*supU* pair, LC-MS/MS analysis identified nsAA **6** incorporation at position 2 and Trp at position 57 of sfGFP-**6** protein produced from adding 1 mM nsAA **6** (Figure 3). Cross-talk events, *i.e.* Trp incorporation at position 2 and nsAA **6** at position 57, were not detected from the LC-MS/MS analysis. New peaks that appear may be caused by recognition of the introduced tRNA suppressor by AARSs other than the intended AARS (*e.g.* **6** in Figure 1A and 1B). Such unintended cross-talk may remain undetected; Trp incorporation was not detected at position 2 in the LC-MS/MS analysis of the sfGFP. Therefore, we are not able to rule out the possibility of cross-talk from other AARSs although the amino acid purity had been established.

Discussion

The active sites of TrpRS, TyrRS, LeuRS and SerRS

While the nsAA library contained analogs of the cognate substrates of all four tested AARSs, TrpRS and TyrRS displayed a much larger productive substrate range than did LeuRS and SerRS. The four AARSs subjected to the nsAA library screening displayed diverse substrate specificities. TyrRS incorporated two Trp analogs (**3** and **4**) and Tyr analogs whose phenyl groups were substituted at the *para* and *meta* positions (**7-13**), although the nsAA incorporation efficiency was low (Figure 2A). TyrRS accommodated Tyr analogs as well as larger compounds like *N*ⁱⁿ-formyl-Trp (**3**) and *N*ⁱⁿ-*tert*-boc-Trp (**4**). This implies that the TyrRS amino acid binding pocket is wider and deeper than necessary for

accommodating Tyr. Sequence and structure based evolutionary analyses show the relationship of TrpRS and TyrRS.^[12] There is high backbone conservation between TrpRS and TyrRS suggesting that side chain variations in the amino acid binding pocket are the main source of discrimination between Trp and Tyr.^[12a] Thus the use of nsAAs containing aromatic side chains (**3** and **4**) by both enzymes is plausible. AARSs with narrow amino acid substrate range (SerRS and LeuRS) tend to recognize nsAAs similar to their cognate amino acid. Intrinsic editing mechanism that prevent formation of misacylated tRNA by SerRS and LeuRS may be responsible for the limited substrate response.^[1a]

tRNA sequence and amino acid selection

The tRNA anticodon affects the amino acid substrate range of its cognate AARS. Given this differential aminoacylation property, select tRNA anticodon variants may permit specific incorporation of nsAAs at targeted codons. For instance, nsAA **6** was efficiently charged onto tRNA^{Trp}_{CUA} (*supU*) but not tRNA^{Trp}_{CCA} (Figure 1A and 1B); this observation was further corroborated by LC-MS/MS analyses that revealed incorporations of nsAA **6** at position 2 and Trp at W57 position, respectively. Since wild-type AARSs differentially aminoacylate their tRNA substrates based on their anticodon sequences, synonymous codons may encode the incorporation of distinct palettes of nsAAs. Superposition of the substrate range profile of tRNA anticodon variants can reveal outliers that may be candidates for specific nsAA incorporation. This approach may allow incorporation of more nsAAs by reducing the degeneracy of the canonical genetic code *in vivo*.

Conclusion

As suggested by early data^[2a] wild-type AARSs can mediate the incorporation of nsAAs into *E. coli* proteins. This has been borne out in our library screens presented here. While AARSs are exceedingly specific in their substrate responses to the constituents of the normal cell^[13], evolution did not provide any defenses against the legions of nsAAs that have been used in the genetic code expansion studies^[4a, 14]. This polyspecificity^[5f, 5i] by wild-type AARSs and their variants will be the great challenge to the site-specific introduction of multiple nsAAs into proteins during attempts to expand the genetic code^[15]. Overcoming these challenges will require more sophisticated genome engineering^[16] and inspired, labor-intensive enzyme engineering.

Experimental Section

System optimization and screening of incorporation of canonical or non-standard amino acids by sfGFP production

E. coli strain BW25113 or its auxotrophic strain JW2581-1 (*tyrA*[−]) was co-transformed with pBAD-*sfGFP-UAG2* and pTech-*supF* (Figure S1, Table S1 and S2). BW25113 cells were cultured in 10 mL M9 medium (M9 salt supplemented with 1% glycerol, 2 mM MgSO₄, 0.1 mM CaCl₂) supplemented with 1 mM of the 20 canonical amino acids, 34 µg/mL chloramphenicol (Cm) and 100 µg/mL ampicillin (Amp) at 37 °C. JW2581-1 cells were cultured in 10 mL M9 medium supplemented with 1 mM of the 20 canonical amino acids, 50 µg/mL kanamycin (Kan), 34 µg/mL Cm and 100 µg/mL Amp at 37 °C. When cell A₆₀₀

reached 0.7 to 1.0, the cells were pelleted and washed thrice with M9 medium, and resuspended in M9 medium (to an A_{600} of 1.0) supplemented with 1 mM of the appropriate nsAAs without 20 natural amino acids or 1mM Tyr, and arabinose (0.2 %) was added to induce sfGFP protein production for 12 h at 37 °C. Cells were collected by centrifugation and re-suspended in 1× phosphate buffered saline (PBS) with same culture volume. The fluorescence emission intensity of cells was measured at excitation wavelength 485/20 nm and emission wavelength 528/20 nm. The suppression experiments on the five natural tRNA suppressors (pTech-*supU*, pTech-*trpT*, pTech-*supD* and pTech-*supP*) were performed with the same conditions but in different auxotrophic strains (Figure S2 and Table S1).

For the substrate screening, we used a library that contained 313 nsAAs (containing 41 new entry nsAAs) (Table S3-S5), and its details were previously reported^[5b]. *E. coli* auxotrophic strains were transformed with either pBAD-*sfGFP* (JW2580 *trpC*⁻); pBAD-*sfGFP*-*UGA2* in conjunction with pTech-*trpT* (JW2580 *trpC*⁻); or pBAD-*sfGFP*-*UAG2* in conjunction with pTech-*supD* (JW2880 *serA*⁻), pTech-*supP* (JW5807 *leuB*⁻), pTech-*supF* (JW2581 *tyrA*⁻) or pTech-*supU* (JW2580 *trpC*⁻) plasmids (Table S1 and S2). Cells were cultured in 100 mL M9 medium supplemented with 1 mM of the 20 natural amino acids, 34 µg/mL Cm, 50 µg/mL Kan, and 100 µg/mL Amp at 37 °C. The JW2580 *trpC*⁻ cell strain containing pBAD-*sfGFP* was similarly cultured but without Cm in the medium. The cells were pelleted and washed twice with M9 medium, and resuspended in M9 medium (to an A_{600} of 1.0) supplemented with 0.2% arabinose was added. Aliquots (50 µL) of the cell suspension were transferred into the wells of a 384-well plate containing a different nsAA (1 mM) in each well. Cells were continuously shaken for 12 h at 37 °C, with monitoring of fluorescence intensity (excitation wavelength 485/20 nm; emission wavelength 528/20 nm) and optical density (A_{600} value) every 10 minutes. Four wells (A1, A2, I1 and I2) without added nsAAs were used as controls to measure the background signals. The nsAA screening experiments were performed in triplicate. For the annotated peaks generated by **1-20** the standard deviations of fluorescence intensities are listed in Table S8.

LC-MS/MS analyses of digested peptides of sfGFP proteins

The sfGFP sample purified from cells expressing *supU* was analyzed via LC-MS/MS. The full-length protein was trypsin digested by a standard in-gel digestion protocol, and analyzed via LC-MS/MS performed on an LTQ Orbitrap XL equipped with a Waters nanoACQUITY™ UPLC™ system. A Waters Symmetry® C18 180 µm × 20 mm trap column and a 1.7 µm, 100 µm × 250 mm nanoAcquity™ UPLC™ column (35°) were utilized for peptide separation. Trapping was done at 15 µL/min, 99% Buffer A (100% water, 0.1% formic acid) for 1 min. Peptide separation was performed at 300 nL/min with Buffer A and Buffer B (100% CH₃CN containing 0.1% formic acid). A linear gradient (51 min) was run with 5% buffer B at initial conditions, 50% B at 50 min, and 85% B at 51 min. MS was acquired in the Orbitrap using 1 microscan, and a maximum inject time of 900 ms followed by data dependent MS/MS acquisitions in the ion trap (via Collision Induced Dissociation, CID). The Mascot search algorithm was used to search for the appropriate non-canonical substitution (Matrix Science).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institute for General Medical Sciences (GM22854) and the Defense Advanced Research Projects Agency (contract N66001-12-C-4211). We thank Mary LoPresti for preparing the MS samples and Dr. TuKiet Lam and Jean Kanyo for collecting and processing the LC-MS/MS data. We are grateful to the members of the Söll lab (Markus Englert, Li-Tao Guo, Noah Reynolds, Akiyoshi Nakamura, and Carla Polycarpo) for critical discussions and helpful insights.

References

- [1]. a) Ling J, Reynolds N, Ibba M. *Annu Rev Microbiol.* 2009; 63:61–78. [PubMed: 19379069] b) Reynolds NM, Lazazzera BA, Ibba M. *Nature Reviews Microbiology.* 2010; 8:849–856.
- [2]. a) Richmond MH. *Bacteriol Rev.* 1962; 26:398–420. [PubMed: 13982167] b) Xu ZJ, Love ML, Ma LY, Blum M, Bronskill PM, Bernstein J, Grey AA, Hofmann T, Camerman N, Wong JT. *J Biol Chem.* 1989; 264:4304–4311. [PubMed: 2494170] c) Li L, Boniecki MT, Jaffe JD, Imai BS, Yau PM, Luthey-Schulten ZA, Martinis SA. *Proc Natl Acad Sci USA.* 2011; 108:9378–9383. [PubMed: 21606343]
- [3]. Hartman MC, Josephson K, Szostak JW. *Proc Natl Acad Sci U S A.* 2006; 103:4356–4361. [PubMed: 16537388]
- [4]. a) Liu CC, Schultz PG. *Annu Rev Biochem.* 2010; 79:413–444. [PubMed: 20307192] b) Johnson JA, Lu YY, Van Deventer JA, Tirrell DA. *Curr Opin Chem Biol.* 2010; 14:774–780. [PubMed: 21071259] c) Lepthien S, Hoesl MG, Merkel L, Budisa N. *Proc Natl Acad Sci U S A.* 2008; 105:16095–16100. [PubMed: 18854410]
- [5]. a) Kavran JM, Gundllapalli S, O'Donoghue P, Englert M, Söll D, Steitz TA. *Proc Natl Acad Sci U S A.* 2007; 104:11268–11273. [PubMed: 17592110] b) Ko JH, Wang YS, Nakamura A, Guo LT, Söll D, Umehara T. *FEBS Lett.* 2013; 587:3243–3248. [PubMed: 23994531] c) Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, Chin JW. *J Am Chem Soc.* 2009; 131:8720–8721. [PubMed: 19514718] d) Takimoto JK, Dellas N, Noel JP, Wang L. *ACS Chem Biol.* 2011; 6:733–743. [PubMed: 21545173] e) Wang YS, Fang X, Chen HY, Wu B, Wang ZU, Hilty C, Liu WR. *ACS Chem Biol.* 2013; 8:405–415. [PubMed: 23138887] f) Wang YS, Fang X, Wallace AL, Wu B, Liu WR. *J Am Chem Soc.* 2012; 134:2950–2953. [PubMed: 22289053] g) Wang YS, Russell WK, Wang Z, Wan W, Dodd LE, Pai PJ, Russell DH, Liu WR. *Mol Biosyst.* 2011; 7:714–717. [PubMed: 21234492] h) Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. *Chem Biol.* 2008; 15:1187–1197. [PubMed: 19022179] i) Young DD, Young TS, Jahnz M, Ahmad I, Spraggon G, Schultz PG. *Biochemistry.* 2011; 50:1894–1900. [PubMed: 21280675] j) Wang ZU, Wang YS, Pai PJ, Russell WK, Russell DH, Liu WR. *Biochemistry.* 2012; 51:5232–5234. [PubMed: 22697363] k) Wang YS, Wu B, Wang Z, Huang Y, Wan W, Russell WK, Pai PJ, Moe YN, Russell DH, Liu WR. *Mol Biosyst.* 2010; 6:1557–1560. [PubMed: 20711534] l) Fekner T, Li X, Lee MM, Chan MK. *Angew Chem Int Ed Engl.* 2009; 48:1633–1635. [PubMed: 19156778]
- [6]. Eggertsson G, Söll D. *Microbiol Rev.* 1988; 52
- [7]. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. *Mol Syst Biol.* 2006; 2:2006–0008.
- [8]. a) Hong KW, Ibba M, Weygand-Durasevic I, Rogers MJ, Thomann HU, Söll D. *EMBO J.* 1996; 15:1983–1991. [PubMed: 8617245] b) Ibba M, Hong KW, Sherman JM, Sever S, Söll D. *Proc. Natl. Acad. Sci. USA.* 1996; 93:6953–6958. [PubMed: 8692925]
- [9]. Rennert OM, Anker HS. *Biochemistry.* 1963; 2:471–476. [PubMed: 14069531]
- [10]. Kanehisa M, Goto S. *Nucleic Acids Res.* 2000; 28:27–30. [PubMed: 10592173]
- [11]. a) Englisch S, Englisch U, von der Haar F, Cramer F. *Nucleic Acids Res.* 1986; 14:7529–7539. [PubMed: 3534789] b) Gruic-Sovulj I, Rokov-Plavec J, Weygand-Durasevic I. *FEBS Lett.* 2007;

- 581:5110–5114. [PubMed: 17931630] c) Boniecki MT, Martinis SA. J Biol Chem. 2012; 287:11285–11289. [PubMed: 22334703]
- [12]. a) O'Donoghue P, Luthey-Schulten Z. Microbiol Mol Biol Rev. 2003; 67:550–573. [PubMed: 14665676] b) Woese CR, Olsen GJ, Ibba M, Söll D. Microbiol Mol Biol Rev. 2000; 64:202–236. [PubMed: 10704480]
- [13]. Ibba M, Söll D. Annu Rev Biochem. 2000; 69:617–650. [PubMed: 10966471]
- [14]. Hoesl MG, Budisa N. Curr Opin Biotechnol. 2012; 23:751–757. [PubMed: 22237016]
- [15]. O'Donoghue P, Ling J, Wang YS, Söll D. Nat Chem Biol. 2013; 9:594–598. [PubMed: 24045798]
- [16]. Lajoie MJ, Rovner AJ, Goodman DB, Aerni HR, Haimovich AD, Kuznetsov G, Mercer JA, Wang HH, Carr PA, Mosberg JA, Rohland N, Schultz PG, Jacobson JM, Rinehart J, Church GM, Isaacs FJ. Science. 2013; 342:357–360. [PubMed: 24136966]

Figure 1.

Range of substrate specificities of TrpRS. Suppression of the *sfGFP-UAG2* and *sfGFP-UGA2* genes by the library of nsAA-tRNA was measured by fluorescence intensity. Four wells (A1, A2, I1 and I2) were set as control experiments without adding any nsAA to detect the background signals, which were used to normalize the signals generated by supplied nsAAs. (A) The substrate specificity profile of TrpRS/ $tRNA^{Trp}_{CCA}$ by production of wild type *sfGFP* gene. (B) The substrate specificity profile of TrpRS/*supU* ($tRNA^{Trp}_{CUA}$) by suppression of the *sfGFP-UAG2* gene. (C) The substrate specificity profile of TrpRS/*trpT*

(tRNA^{Trp}_{UCA}) by suppression of the *sfGFP-UGA2* gene. The cell strain JW2580-1 (*trpC*⁻) was used in these experiments. The nsAA chemical structures are provided in our reported nsAA library^[5b]; Tables S4 and S5 describe the additional nsAAs not previously reported. The labeled nsAA chemical structures are given in Scheme 1.

Figure 2.

Range of substrate specificities of TyrRS, LeuRS and SerRS. Suppression of the *sfGFP-UAG2* gene by the library of nsAA-tRNA_{CUA} was measured by fluorescence intensity. Four wells (A1, A2, I1 and I2) were set as control experiments without adding any nsAA to detect the background signals, which were used to normalize the signals generated by supplied nsAAs. (A) The substrate specificity profile of TyrRS/*supF* (tRNA^{Tyr}_{CUA}) pair. (B) The substrate specificity profile of LeuRS/*supP* (tRNA^{Leu}_{CUA}) pair. (C) The substrate specificity profile of SerRS/*supD* (tRNA^{Ser}_{CUA}) pair. The nsAA chemical structures are

provided in our reported nsAA library^[5b]; Tables S4 and S5 describe the additional nsAAs not previously reported. The labeled nsAA chemical structures are given in Scheme 1.

Figure 3.

Mass spectrometric analyses of sfGFP proteins. The full length sfGFP-6 was produced by using tRNA^{Trp}_{CUA} gene in the Trp auxotrophic strain supplied with 1 mM nsAA 6 in the M9 medium. (A) The LC-MS/MS spectrum of the MXKGEELFTGVVPILVELDGDVNGHK (X denotes nsAA 6 incorporated at Ser2TAG position) fragment from sfGFP-6 protein. (B) The LC-MS/MS spectrum of the ICTTGKLPVPWPPTLVTTTL fragment from sfGFP-6 protein. b⁰ and y⁰ denote loss of water from the b or y fragments; y* denotes loss of

ammonia (NH₃) from y fragments respectively. The full list of m/z values of fragments is provided in Table S6 and S7.

Scheme 1.

Chemical structures of canonical and non-standard amino acids mentioned in this study.