

Published in final edited form as:

Chembiochem. 2014 August 18; 15(12): 1750–1754. doi:10.1002/cbic.201400075.

Toward Reassigning the Rare AGG Codon in *Escherichia coli*

Yu Zeng^a, Wei Wang^a, and Wenshe R. Liu^a

Wenshe R. Liu: wliu@chem.tamu.edu

^aDepartment of Chemistry, Texas A&M University, College Station, TX 77843 (USA)

Abstract

The rare AGG codon in *Escherichia coli* has been reassigned to code non-canonical amino acids using the PylRS–tRNA^{Pyl}_{CCU} pair. When *N*^ε-alloc-lysine was used as a PylRS substrate, close to quantitative occupancy of *N*^ε-alloc-lysine at an AGG codon site was achieved in minimal medium.

Keywords

Genetic code expansion; AGGA codon; AGG codon; sense codon reassignment; non-canonical amino acids

Pioneered by Schultz and others, using orthogonal aminoacyl-tRNA synthetase (aaRS)-suppressor tRNA pairs for the delivery of unique amino acids at nonsense and four-base codons have allowed the incorporation of more than 100 non-canonical amino acids (ncAAs) into proteins in living cells of *Escherichia coli*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, and mammals.¹ These ncAAs consist of diverse unique functionalities that can be applied for both basic studies and biotechnological development. Several broadly used orthogonal pairs have been constructed for this purpose, including one derived from the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS)-tRNA^{Tyr} pair for use in *E. coli*, two derived from the *E. coli* TyrRS-tRNA^{Tyr} and leucyl-tRNA synthetase-tRNA^{Leu} pairs for use in eukaryotic cells, and the naturally existing pyrrolysyl-tRNA synthetase (PylRS)-tRNA^{Pyl} pair for use in both bacterial and eukaryotic cells.^{1e, 2} All four aaRS-tRNA pairs are primarily used for reassigning the amber UAG codon and two have been extended to reassign the opal UGA, ochre UAA, and four-base AGGA codons.³ Using all four codons to code ncAAs typically suffers low ncAA incorporation yields. For three nonsense codons, suppressor tRNAs have to compete against release factors (RFs) for their recognition during translation; for the four-base AGGA codon, two endogenous arginyl-tRNA^{Pyl} variants, arginyl-tRNA^{Arg}_{CCU} and arginyl-tRNA^{Arg}_{UCU}, recognize the first three nucleotides and potentially lead to misreading during translation. To improve amber suppression efficiency, cells with an orthogonal ribosome, a truncated ribosome with slow RF1 binding, and RF1 knockout have been engineered.⁴ In these cells, the fact that translation inclines to stop at amber codon still remains. Previously we reported that the PylRS-tRNA^{Pyl} pair could be reprogrammed for reassigning three nonsense

Correspondence to: Wenshe R. Liu, wliu@chem.tamu.edu.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.20xxxxxxx>.

codons.⁵ Here, we wish to show that the same pair can be engineered to reassign the rare AGG codon almost quantitatively to code ncAAs, resolving the low ncAA incorporation efficiency issue that is typically observed in codon reassignment experiments.

Flexible tRNA anticodon recognition of PylRS was previously reported.⁶ By simply mutating the anticodon of tRNA^{Pyl}, we showed that the pair could be successfully reprogrammed to reassign opal and ochre codons for coding ncAAs such as N^ε-(t-butyloxycarbonyl)-lysine (BocK), N^ε-(allyloxycarbonyl)-lysine (AllocK), and N^ε-(t-propargyloxycarbonyl)-lysine (ProK) (Scheme 1).⁵ However, when we attempted to use the same strategy to reassign a four-base AGGA codon at the 134th amino acid site of a sequence-optimized superfolder green fluorescent protein (sfGFP) gene, a full-length protein with Arg134 was predominantly produced, indicating strong frameshift suppression from endogenous arginyl-tRNA^{Arg}. The presence of 5 mM BocK in the growth medium did not lead to a significant level of BocK incorporation at the AGGA codon site.^{6c} However, when a different GFP strain, GFP_{UV} with no sequence optimization and an AGGA mutation at its 149th amino acid site was used for a similar test, GFP_{UV} with BocK149 was successfully produced in cells that expressed the PylRS-tRNA^{Pyl}_{UCCU} pair and were grown in the presence of BocK (Figure 1A). The molecular weight of the purified protein determined by electrospray ionization mass spectrometry (ESI-MS) agreed well with its theoretical value (Figure 1B). In the absence of BocK, only a minimal amount of full-length GFP_{UV} could be detected, indicating a very low basal level of frameshift suppression from endogenous arginyl-tRNA^{Arg}. To investigate effects of arginine on the AGGA suppression, we prepared an arginine-dropout auto-induction (AI) medium (see supplementary information for recipe) and used it to test BocK incorporation efficiencies at AGGA in the presence of different concentrations of arginine. As shown in Figure 1A, adding 0.1 mM arginine in the medium improved the expression yield of GFP_{UV}, but the yield continued to drop when more arginine was provided. We suspect that the expression increase in the presence of 0.1 mM arginine was due to the relief of translation restriction at regular arginine codons and the decrease with further addition of arginine was caused by increasing misreading of the first three nucleotides of AGGA by endogenous arginyl-tRNA^{Arg}. These data strongly suggest that misreading at the first three nucleotides of AGGA is a competing process during the translational reassignment of AGGA to code a ncAA.

Our AGGA suppression results with GFP_{UV} prompted us to investigate the AGGA codon reassignment in sfGFP further. Since the sfGFP gene was codon-optimized, a potentially fast translation rate of its mRNA transcript might have favoured frameshift suppression at AGGA from arginyl-tRNA^{Arg}. To inquire this possibility, codons of five residues at the immediate N-terminal side of the 134th position of sfGFP were mutated to relatively rarely used codons. However, growing cells harbouring this gene, PylRS, and tRNA^{Pyl}_{UCCU} still led to the predominant production of sfGFP with Arg134 both in the absence and in the presence of BocK in LB medium (Figure 2). Using low temperature and GMML (M9 minimal medium with 1% glycerol and 0.3 mM leucine) for slow translation afforded similar results. At the current stage, we are still puzzled by conflicting AGGA suppression results for GFP_{UV} and sfGFP. We suggest precautions when using the AGGA codon to code ncAAs. Although this further study did not unveil underlying reasons for frameshift

suppression at AGGA from arginyl-tRNA^{Arg}, a close examination of the ESI-MS spectrum of sfGFP produced in the presence of BocK revealed two small peaks that did not match the theoretical molecular weights of sfGFP with BocK134 (Figure 2C). Therefore, tRNA^{Pyl}_{UCCU} was indeed able to deliver BocK at the AGGA mutation site but could not efficiently compete against endogenous arginyl-tRNA^{Arg} for binding to this particular site.

Since ribosome is a catalytic machinery for translating three-nucleotide codons, its binding to BocK-tRNA^{Pyl}_{UCCU} may be disfavoured, making BocK-tRNA^{Pyl}_{UCCU} potentially disadvantageous in competing against arginyl-tRNA^{Arg} for binding to AGGA. Deleting the first nucleotide from the anticodon of tRNA^{Pyl}_{UCCU} to form tRNA^{Pyl}_{CCU} will presumably improve its binding to ribosome. This new tRNA^{Pyl} mutant may be well charged by PylRS and efficiently compete against arginyl-tRNA^{Arg} in the recognition of the rare AGG codon, leading to its reassignment to code ncAAs. To examine this potential, a plasmid pET-AGG-GFP_{UV}149AGG that carried genes coding PylRS, tRNA^{Pyl}_{CCU} and GFP_{UV} with a single AGG codon at its 149th position was constructed. In this plasmid, PylRS and GFP_{UV}149AGG are under control of an IPTG-inducible T7 promoter and tRNA^{Pyl}_{CCU} is under control of a constitutive *lpp* promoter. Growing *E. coli* BL21(DE3) cells transformed with pET-AGG-GFP_{UV}149AGG in 2×YT medium expressed GFP_{UV} with Arg149 (Figure 3A). Providing 10 mM BocK in the medium led to the synthesis of two GFP_{UV} variants, one with Arg149 and the other with BocK149 (Figure 3B). Integrating ESI-MS peaks of two variants indicated that GFP_{UV} with BocK149 was roughly 24% of total GFP_{UV} variants compared to 15% when 5 mM BocK was provided in the medium (Supplementary Figure 1). Therefore PylRS is active toward tRNA^{Pyl}_{CCU} that does competitively recognize AGG during translation. Since 2×YT medium has a high concentration of arginine that promotes the reading of the AGG codon by endogenous arginyl-tRNA^{Arg}, we shifted to GMML to do further test. Growing *E. coli* BL21(DE3) cells transformed with pET-AGG-GFP_{UV}149AGG in GMML led to a dominant production of GFP_{UV} with BocK149 (Figure 3C). Among the total GFP_{UV} variants, GFP_{UV} with BocK149 accounts for 80%. To further increase the incorporation of BocK at the AGG codon site, an argW antisense RNA was coexpressed in *E. coli* BL21(DE3) cells that were transformed with pET-AGG-GFP_{UV}149AGG. ArgW codes tRNA^{Arg}_{CCU}. Expression of its antisense RNA potentially reduces the recognition of AGG by endogenous arginyl-tRNA^{Arg}_{CCU}. Shown in Figure 3D, coexpressing argW antisense RNA did increase the proportion of GFP_{UV} with BocK149 in totally expressed GFP_{UV} variants to 83% but was not significantly better than in cells with no argW antisense RNA expressed.

In experiments described in the previous paragraph, both IPTG and BocK were added into the medium at the same time to induce protein expression. Since PylRS was under control of an IPTG-inducible promoter, it surely took some time for PylRS to be expressed and achieve steady activity to charge tRNA^{Pyl}_{CCU} with BocK after induction. During this time, some GFP_{UV} with Arg149 had already been expressed since its expression was also induced by IPTG. One way to reduce this undesired expression of GFP_{UV} with Arg149 is to achieve a

stable level of Bock-tRNA^{Pyl}_{CCU} before GFP_{UV} expression is initiated. A separate plasmid pEVOL-AGG was constructed for this purpose. This plasmid contains genes coding PylRS under control of an arabinose-regulated pBAD promoter and tRNA^{Pyl}_{CCU} under control of a constitutive *proK* promoter. The *E. coli* BL21-AI strain was chosen for the following protein expression studies due to the fact that this strain allows T7 RNA polymerase and a target gene to be expressed separately. *E. coli* BL21-AI has a T7 RNA polymerase gene under control of an arabinose-regulated pBAD promoter in its chromosome. As such, in this strain, T7 RNA polymerase can be expressed to a stable level by the addition of arabinose before IPTG is provided to induce expression of a target protein. Since AGG codon reassignment may potentially affect the translation of T7 RNA polymerase itself, stably expressed T7 RNA polymerase before the induction with IPTG and Bock will provide enough functional enzyme for expressing a target protein. *E. coli* BL21-AI cells were co-transformed with pEVOL-AGG and pET-AGG-GFP_{UV}149AGG. A single transformant was then used to express GFP_{UV}. Cells were grown in GML to OD 0.5. arabinose, Bock, and IPTG were added sequentially in 60 min apart to induce the expression of GFP_{UV}. ESI-MS analysis of purified GFP_{UV} indicated a predominant GFP_{UV} variant with Bock149 that accounted for 89% of total expressed GFP_{UV} that had an expression level of 30 mg/L (Figure 4A). This represents a significant increase from 80% that was observed in the previous study with BL21(DE3) transformed with pET-AGG-GFP_{UV}134AGG. We also tested another two PylRS substrates, AllocK and ProK. For AllocK, 92% of total expressed GFP_{UV} with an expression level of 36 mg/L was the variant with AllocK149 (Figure 4B). This high incorporation efficiency at AGG is probably due to AllocK as a better substrate of PylRS than Bock. For ProK, a significant proportion of GFP_{UV} with Arg149 (32%) was in the finally expressed GFP_{UV} that had an expression level of 33 mg/L, indicating ProK is probably a weaker substrate of PylRS than Bock (Figure 4C). The GFP_{UV} expression level in the absence of a ncAA was 24 mg/L. We also tested the toxic effects of the AGG reassignment. *E. coli* BL21-AI cells co-transformed with pEVOL-AGG and pET-AGG-GFP_{UV}149AGG were grown with or without 10 mM ProK and induced with the addition of arabinose and IPTG. Cells in media with Bock grew faster than those in media without the addition of Bock, indicating a very low toxicity of the AGG reassignment to cells (Supplementary Figure 2). A faster growth pattern for cells in the presence of Bock might be due to that Bock could be metabolized to provide nutrients for cell growth. We also investigated the global incorporation of a ncAA at proteins whose genes naturally have AGG codons. Lysates of cells grown in the presence of 10 mM ProK were labelled with 3-azido-7-hydroxyl-coumarin using the copper-catalyzed click reaction. There were only a few proteins which could be fluorescently labelled (Supplementary Figure 3). This is not a surprise given that AGG is a rare codon and arginine has important structural roles in protein folding. When ProK replaces arginine, it may likely causes a protein not properly folded and then cleared out in cells. Since we previously observed strong arginine incorporation at the AGGA mutation site of the 134th amino acid position of sfGFP, we also tested whether a similar observation is associated with an AGG mutation at this site. A plasmid pET-AGG-sfGFP134AGG was constructed by swapping the GFP_{UV} gene in pET-AGG-GFP_{UV}149AGG with the sequence-optimized sfGFP gene with an AGG codon at its 134th amino acid position. This plasmid together with pEVOL-AGG was used to transform *E. coli*

BL21-AI cells. A single transformant was then employed to produce sfGFP by following the same procedure for the production of GFP_{UV}. BocK was used in this investigation. As presented in Figure 4D, sfGFP with BocK134 was the major sfGFP variant and accounted for 70% of total expressed sfGFP. Therefore, BocK-tRNA^{Arg}_{CCU} is competitive in recognizing this site. The low incorporation of BocK at the AGGA mutation of the same site by tRNA^{Arg}_{UCCU} must be due to the weak binding of tRNA^{Arg}_{UCCU} to ribosome, though we do not fully understand how arginyl-tRNA^{Pyl} could induce so strong frameshift suppression at an AGGA mutation site.

ProK has a terminal alkyne. Its genetic incorporation followed by click labelling with azides have been used for site-selective protein labelling. This strategy was also applied to label GFP_{UV} that was expressed in the conditions shown in Figure 4C. GFP_{UV} expressed in the presence of ProK could be selectively labelled with a fluorescein azide (Figure 5).⁷ However, a same labelling reaction was not able to label GFP_{UV} that was expressed in the absence of ProK. The labelling data and the ESI-MS analysis shown in Figure 4C collectively support the genetic incorporation of ProK at the AGG codon site.

In summary, we have shown that the rare AGG codon can be reassigned to code ncAAs using the PylRS-tRNA^{Pyl}_{CCU} pair. When AllocK was used as a PylRS substrate, close to quantitative occupancy at the AGG codon was achieved in minimal medium. The other two PylRS substrates, BocK and ProK were also successfully incorporated at the AGG codon with high efficiencies. So far, PylRS and its engineered mutants have been shown to recognize close to more than 90 ncAAs.^{1i, 2d, 8} These ncAAs can be potentially incorporated at the AGG codon with varying efficiencies depending on their activities toward corresponding enzymes. Since AGG is a sense codon, the reported approach resolves the typical low ncAA incorporation issue that has been pegged to the ncAA mutagenesis and therefore allow bulk preparation of proteins with site-selectively incorporated ncAAs for applications such as therapeutic protein production. One unresolved obstacle of the current method is the production of a protein mixture with either a ncAA or an arginine at a designated AGG site. However, a ncAA such as ProK can be easily modified such as with polyethylene glycol. The modified protein can then be separated from the arginine version. Church *et al.* recently reported an *E. coli* cell strain in which RF1 was removed and all 321 TAG codons in the *E. coli* genome were replaced with TAA codons.⁹ The same technique can be applied to construct an *E. coli* cell strain that has all AGG codons replaced with other arginine codons and tRNA^{Pyl}_{CCU} removed. This new cell strain will allow the exclusive incorporation of a ncAA at an AGG codon site. A cell strain that has both TAG and AGG reassigned can also be potentially constructed for the easy production of proteins with two different ncAAs.^{3a, 10}

Experimental Section

Details of plasmid construction, protein expression, ESI-MS analysis, and protein labelling are presented in the supplementary materials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

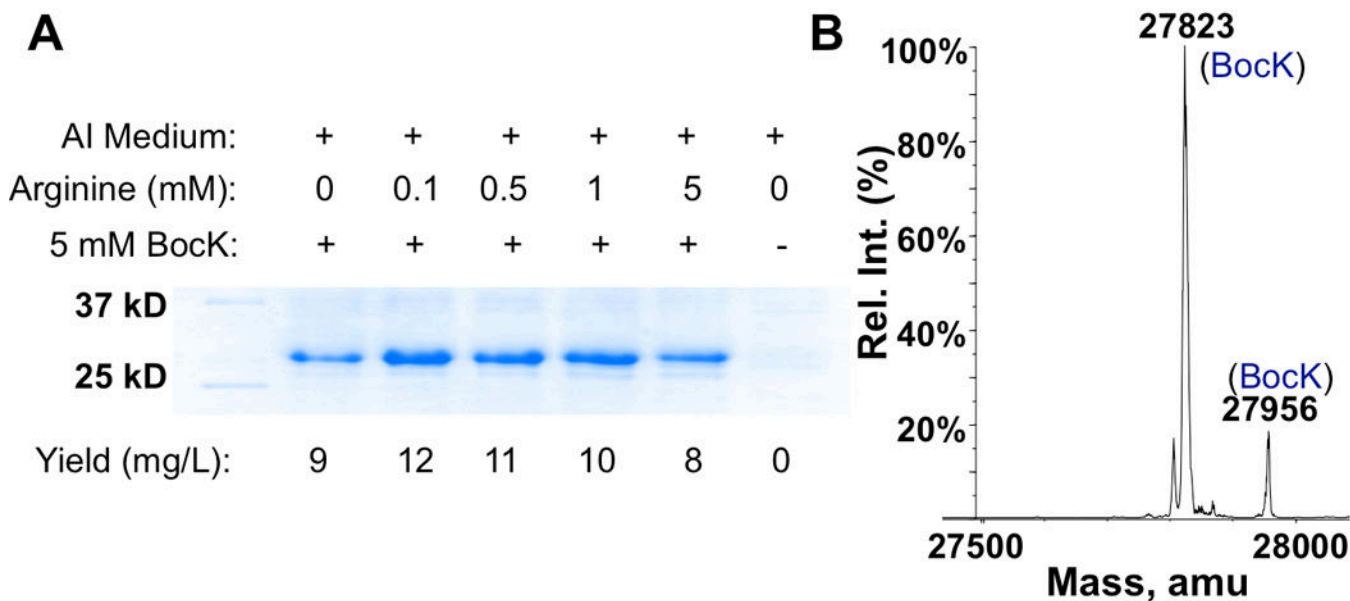
Acknowledgments

We thank Dr. Yohannes H. Reznom from the Laboratory for Biological Mass Spectrometry at Texas A&M University for characterizing proteins with electrospray-ionization mass spectrometry. This work was supported in part by the National Institutes of Health (grant 1R01CA161158), the National Science Foundation (grant CHEM-1148684), and the Welch Foundation (grant A-1715).

References

1. a) Furter R. *Protein Sci.* 1998; 7:419–426. [PubMed: 9521119] b) Wang L, Brock A, Herberich B, Schultz PG. *Science.* 2001; 292:498–500. [PubMed: 11313494] c) Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG. *Science.* 2003; 301:964–967. [PubMed: 12920298] d) Liu W, Brock A, Chen S, Chen S, Schultz PG. *Nat. Methods.* 2007; 4:239–244. [PubMed: 17322890] e) Wu N, Deiters A, Cropp TA, King D, Schultz PG. *J. Am. Chem. Soc.* 2004; 126:14306–14307. [PubMed: 15521721] f) Hancock SM, Uprety R, Deiters A, Chin JW. *J. Am. Chem. Soc.* 2010; 132:14819–14824. [PubMed: 20925334] g) Bianco A, Townsley FM, Greiss S, Lang K, Chin JW. *Nat. Chem. Biol.* 2012; 8:748–750. [PubMed: 22864544] h) Greiss S, Chin JW. *J. Am. Chem. Soc.* 2011; 133:14196–14199. [PubMed: 21819153] i) Mukai T, Kobayashi T, Hino N, Yanagisawa T, Sakamoto K, Yokoyama S. *Biochem. Biophys. Res. Commun.* 2008; 371:818–822. [PubMed: 18471995] j) Liu CC, Schultz PG. *Annu. Rev. Biochem.* 2010; 79:413–414. [PubMed: 20307192] k) Wan W, Sharp JM, Liu WR. *Biochim. Biophys. Acta.* 2014; 1844:1059–1070. [PubMed: 24631543]
2. a) Wang L, Schultz PG. *Chem. Biol.* 2001; 8:883–890. [PubMed: 11564556] b) Chin JW, Cropp TA, Chu S, Meggers E, Schultz PG. *Chem. Biol.* 2003; 10:511–519. [PubMed: 12837384] c) Blight SK, Larue RC, Mahapatra A, Longstaff DG, Chang E, Zhao G, Kang PT, Green-Church KB, Chan MK, Krzycki JA. *Nature.* 2004; 431:333–335. [PubMed: 15329732] d) Neumann H, Peak-Chew SY, Chin JW. *Nat. Chem. Biol.* 2008; 4:232–234. [PubMed: 18278036]
3. a) Wan W, Huang Y, Wang Z, Russell WK, Pai PJ, Russell DH, Liu WR. *Angew. Chem. Int. Ed. Engl.* 2010; 49:3211–3214. [PubMed: 20340150] b) Neumann H, Slusarczyk AL, Chin JW. *J. Am. Chem. Soc.* 2010; 132:2142–2144. [PubMed: 20121121] c) Niu W, Schultz PG, Guo J. *ACS Chem. Biol.* 2013; 8:1640–1645. [PubMed: 23662731]
4. a) Rackham O, Chin JW. *J. Am. Chem. Soc.* 2005; 127:17584–17585. [PubMed: 16351070] b) Huang Y, Russell WK, Wan W, Pai PJ, Russell DH, Liu W. *Mol. Biosyst.* 2010; 6:683–686. [PubMed: 20237646] c) Johnson DB, Xu J, Shen Z, Takimoto JK, Schultz MD, Schmitz RJ, Xiang Z, Ecker JR, Briggs SP, Wang L. *Nat. Chem. Biol.* 2011; 7:779–786. [PubMed: 21926996]
5. Odoi KA, Huang Y, Rezenom YH, Liu WR. *PLoS One.* 2013; 8:357035.
6. a) Ambrogelly A, Gundllapalli S, Herring S, Polycarpo C, Frauer C, Soll D. *Proc. Natl. Acad. Sci. USA.* 2007; 104:3141–3146. [PubMed: 17360621] b) Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. *J. Mol. Biol.* 2008; 378:634–652. [PubMed: 18387634] c) O'Donoghue P, Prat L, Heinemann IU, Ling J, Odoi K, Liu WR, Soll D. *FEBS Lett.* 2012; 586:3931–3937. [PubMed: 23036644]
7. Wu B, Wang Z, Huang Y, Liu WR. *Chembiochem.* 2012; 13:1405–1408. [PubMed: 22628069]
8. a) Polycarpo CR, Herring S, Berube A, Wood JL, Soll D, Ambrogelly A. *FEBS Lett.* 2006; 580:6695–6700. [PubMed: 17126325] b) Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. *Chem. Biol.* 2008; 15:1187–1197. [PubMed: 19022179] c) Chen PR, Groff D, Guo J, Ou W, Cellitti S, Geierstanger BH, Schultz PG. *Angew. Chem. Int. Ed. Engl.* 2009; 48:4052–4055. [PubMed: 19378306] d) Fekner T, Li X, Lee MM, Chan MK. *Angew. Chem. Int. Ed. Engl.* 2009; 48:1633–1635. [PubMed: 19156778] e) Li WT, Mahapatra A, Longstaff DG, Bechtel J, Zhao G, Kang PT, Chan MK, Krzycki JA. *J. Mol. Biol.* 2009; 385:1156–1164. [PubMed: 19063902] f) Li X, Fekner T, Ottesen JJ, Chan MK. *Angew. Chem. Int. Ed. Engl.* 2009; 48:9184–9187. [PubMed: 19882608] g) Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, Chin

- JW. *J. Am. Chem. Soc.* 2009; 131:8720–8721. [PubMed: 19514718] h) Ai HW, Lee JW, Schultz PG. *Chem. Commun. (Camb)*. 2010; 46:5506–5508. [PubMed: 20571694] i) Gautier A, Nguyen DP, Lusic H, An W, Deiters A, Chin JW. *J. Am. Chem. Soc.* 2010; 132:4086–4088. [PubMed: 20218600] j) Groff D, Chen PR, Peters FB, Schultz PG. *Chembiochem*. 2010; 11:1066–1068. [PubMed: 20422671] k) Huang Y, Wan W, Russell WK, Pai PJ, Wang Z, Russell DH, Liu W. *Bioorg. Med. Chem. Lett.* 2010; 20:878–880. [PubMed: 20074948] l) Li X, Fekner T, Chan MK. *Chem. Asian. J.* 2010; 5:1765–1769. [PubMed: 20544791] m) 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] n) Chou CJ, Uprety R, Davis L, Chin JW, Deiters A. *Chemical Science*. 2011; 2:480–483. o) Gautier A, Deiters A, Chin JW. *J. Am. Chem. Soc.* 2011; 133:2124–2127. [PubMed: 21271704]
9. a) Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. *Nature*. 2009; 460:894–898. [PubMed: 19633652] b) 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]
10. a) Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW. *Nature*. 2010; 464:441–444. [PubMed: 20154731] b) Xiao H, Chatterjee A, Choi SH, Bajjuri KM, Sinha SC, Schultz PG. *Angew. Chem. Int. Ed. Engl.* 2013; 52:14080–14083. [PubMed: 24353230] c) Chatterjee A, Xiao H, Bollong M, Ai HW, Schultz PG. *Proc. Natl. Acad. Sci. USA*. 2013; 110:11803–11808. [PubMed: 23818609]

**Figure 1.**

(A) Effects of arginine on the AGGA suppression. *E. coli* BL21(DE3) cells coding PylRS, tRNA^{Pyl}_{UCCU}, and GFP_{UV} with an AGGA mutation at the 149 position and under control of an inducible T7 promoter were grown in AI medium supplemented with varying concentrations of arginine and BocK. Protein expression was induced by the addition of 0.5 mM IPTG. (B) The ESI-MS spectrum of purified GFP_{UV} with BocK149. Theoretical molecular weights of full-length protein and full-length protein without Met1 are 27824 Da and 27955 Da, respectively.

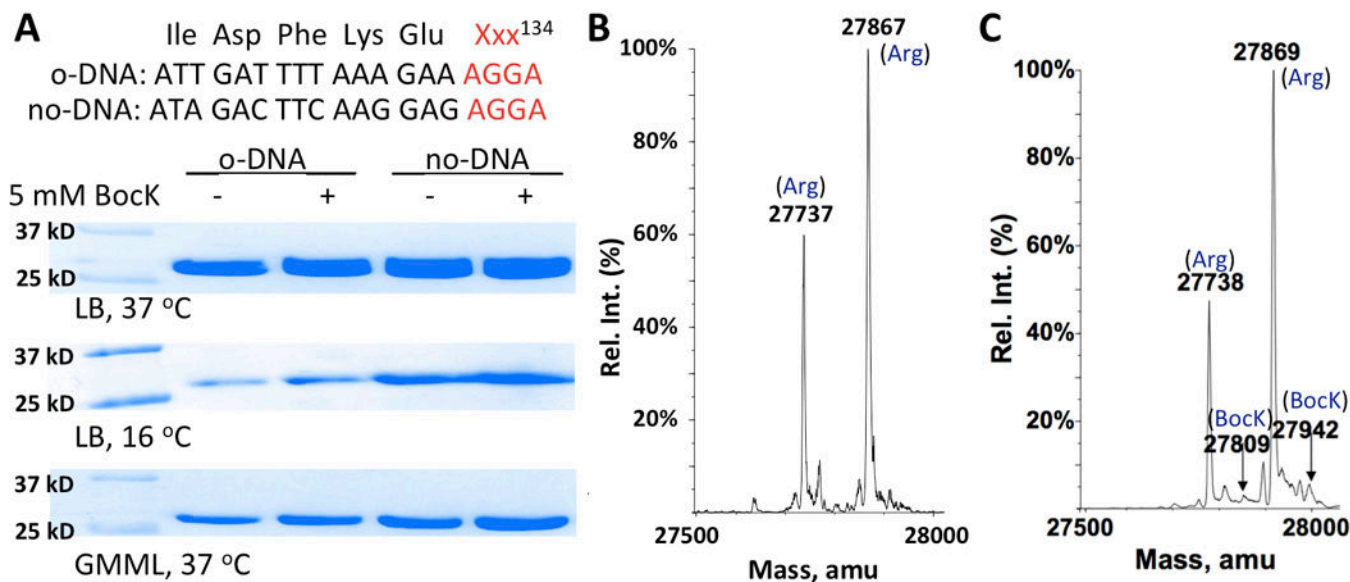


Figure 2.

(A) Frameshift suppression at AGGA of sequence-optimized (opt) and sequence-unoptimized (unopt) sfGFP genes. *E. coli* BL21(DE3) cells coding PyIRS, tRNA^{Pyl}_{UCCU}, and a sfGFP gene (opt or unopt) with an AGGA mutation at the 134th position and under control of an inducible T7 promoter were used for the test. Protein expression was induced with the addition of 0.5 mM IPTG. (B) The ESI-MS spectrum of sfGFP expressed in *E. coli* BL21(DE3) cells that coded PyIRS, tRNA^{Pyl}_{UCCU}, and sequence-optimized sfGFP with an AGGA mutation at the 134th position and were grown at 37 °C in LB. Theoretic molecular weights are 27867 Da for sfGFP with Arg134 and 27735 Da for sfGFP with Arg134 but without Met1. (C) The ESI-MS spectrum of sfGFP expressed in *E. coli* BL21(DE3) cells that coded PyIRS, tRNA^{Pyl}_{UCCU}, and sequence-unoptimized sfGFP with an AGGA mutation at the 134th position and were grown at 37 °C in LB with 5 mM Bock. Theoretic molecular weights are 27940 Da for sfGFP with Bock134 and 27809 Da for sfGFP with Bock134 but without Met1.

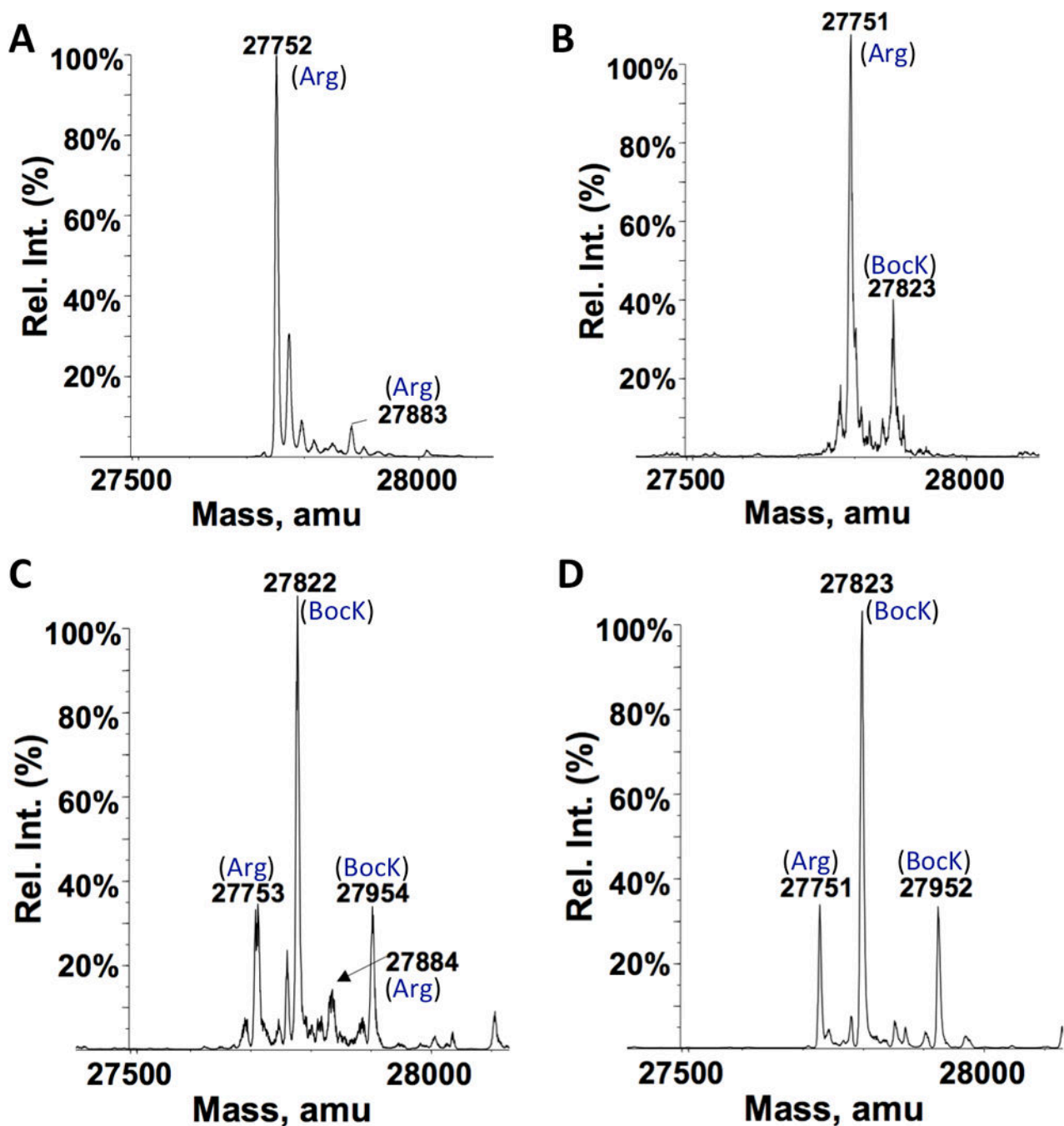


Figure 3.

ESI-MS spectra of GFP_{UV} variants expressed in different conditions. (A) *E. coli* BL21(DE3) cells with pET-AGG-GFP_{UV}149AGG were grown in 2×YT medium and induced with the addition of 0.5 mM IPTG. (B) *E. coli* BL21(DE3) cells with pET-AGG-GFP_{UV}149AGG were grown in 2×YT medium and induced with the addition of 0.5 mM IPTG and 10 mM BockK. (C) *E. coli* BL21(DE3) cells with pET-AGG-GFP_{UV}149AGG were grown in GMML and induced with the addition of 0.5 mM IPTG and 10 mM BockK. (D) *E. coli* BL21(DE3) cells co-transformed with pET-AGG-GFP_{UV}149AGG and another

plasmid containing an *argW* antisense RNA gene were grown in GMMML and induced with the addition of 0.5 mM IPTG and 10 mM Bock.

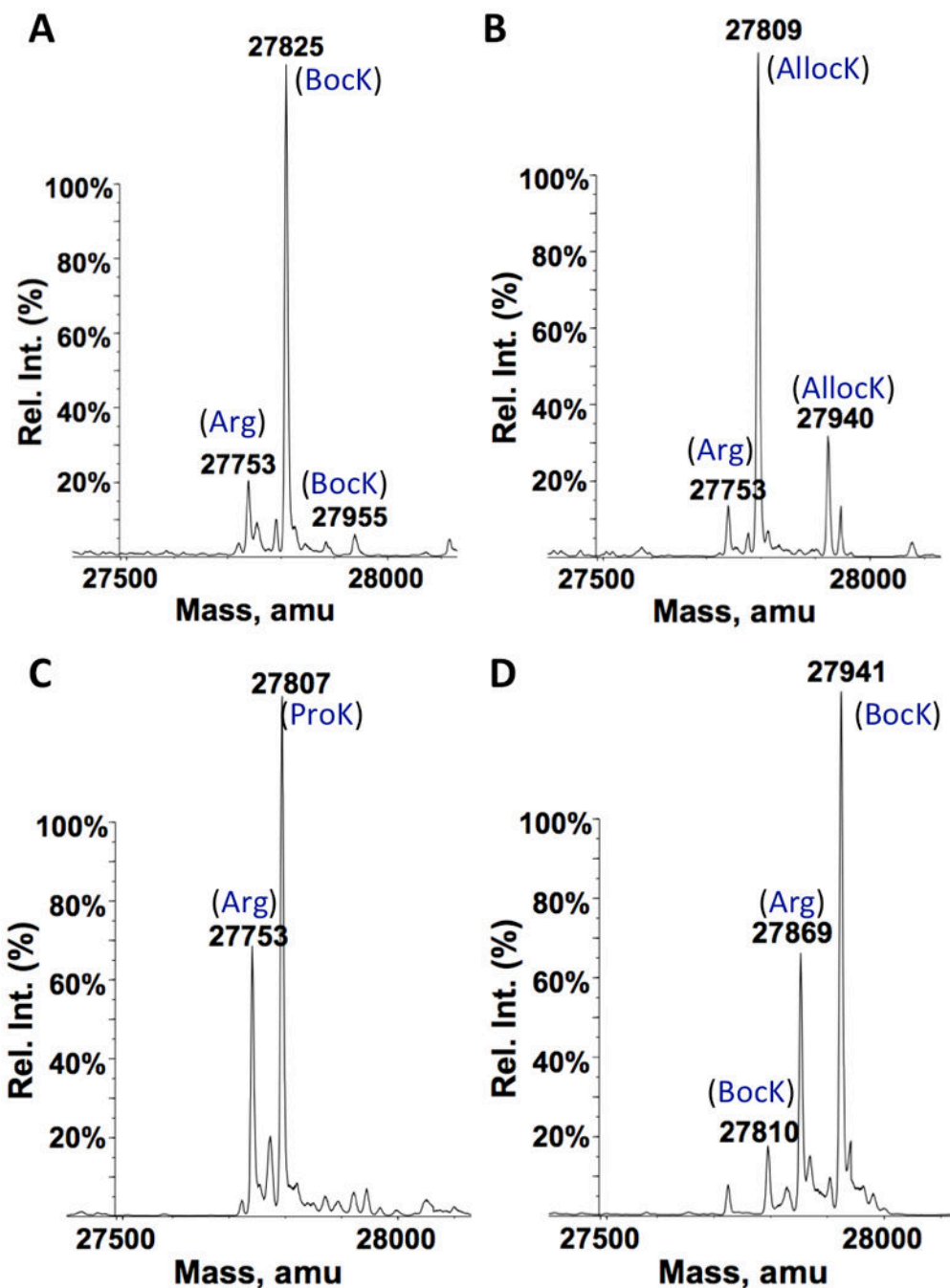


Figure 4. ESI-MS spectra of GFP_{UV} and sfGFP variants expressed in different conditions. For A–C, *E. coli* BL21-AI cells transformed with pEVOL-AGG and pET-AGG-GFP_{UV}149AGG were grown in GML to OD 0.5. 0.2% arabinose, 10 mM ncAA (BocK for A, AlloK for B, and ProK for C), and 0.5 mM IPTG was provided in a sequential order with 1 h apart to induce the expression of GFP_{UV}. For D, *E. coli* BL21-AI cells were transformed with pEVOL-AGG and pET-AGG-sfGFP134AGG. The expression condition of sfGFP was as

same as in **A**. Theoretic molecular weights are 27867 Da for sfGFP with Arg134, 27940 Da for sfGFP with BocK134, and 27809 Da for sfGFP with BocK134 but without Met1.

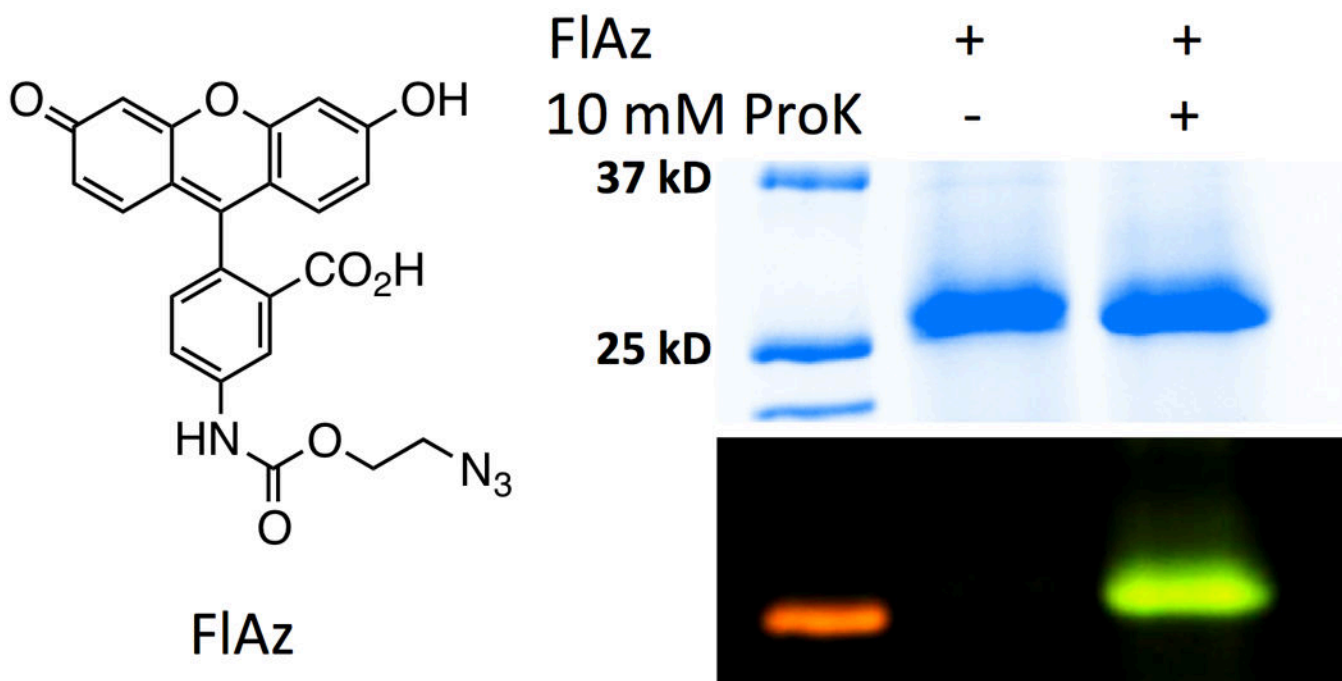
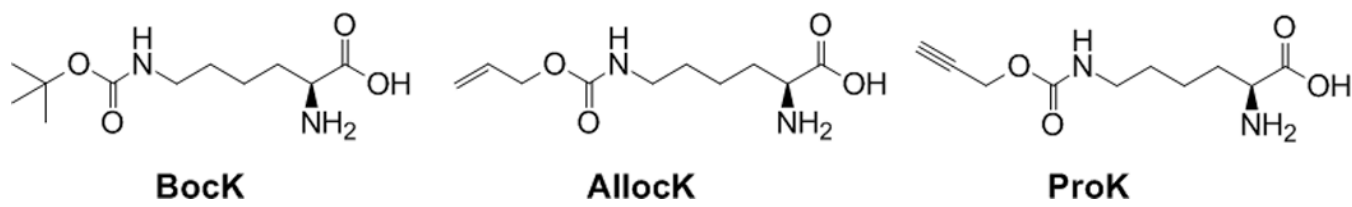


Figure 5.

Labelling of GFP_{UV} that were expressed in the absence and in the presence of ProK. Cells and expression conditions were as same as shown in Figure 4C except that GFP_{UV} shown in the second lane was produced in the absence of ProK. To label GFP_{UV}, Two purified GFP_{UV} samples (1mg/mL) was incubated with 100 μ M CuSO₄, 1 mM NiCl₂, 0.5 mM tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, and 2.5 mM FIAz, followed by 5 mM sodium ascorbate, RT for 3 h.



Scheme 1.
Three ncAAs used in the study.